
GAZİ ÜNİVERSİTESİ
Gazi
1926

EFFICIENT ESTIMATION OF SHRINKAGE PARAMETERS IN FUZZY RIDGE AND FUZZY LIU REGRESSION MODELS USING α-CUT-BASED METHODS UNDER MULTICOLLINEARITY

Ammar HOMAIDA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
IN STATISTICS DEPARTMENT

GAZİ UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

JUNE 2025

ETHICAL STATEMENT

I at this moment declare that in this thesis study, I prepared the thesis writing rules of Gazi University Graduate School of Natural and Applied Sciences;

- All data, information, and documents presented in this thesis have been obtained within the scope of academic rules and ethical conduct,
- All information, documents, assessments, and results have been presented by scientific ethical conduct and moral rules,
- All material used in this thesis that is not original to this work has been exhaustively cited and referenced,
- No change has been made in the data used,
- The work presented in this thesis is original,

or else, I admit all loss of rights to be incurred against me.

Ammar HOMAIDA

16/06/2025

EFFICIENT ESTIMATION OF SHRINKAGE PARAMETERS IN FUZZY RIDGE AND FUZZY LIU REGRESSION MODELS USING α-CUT-BASED METHODS UNDER MULTICOLLINEARITY

(Ph. D. Thesis)

Ammar HOMAIDA

GAZİ UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

June 2025

ABSTRACT

This study presents a novel hybrid fuzzy regression framework that integrates Ridge and Liu estimation techniques within the α-cut-based approach to effectively handle fuzzy datasets affected by multicollinearity. While previous studies relied on K-fold cross-validation to select the fuzzy Ridge bias parameter k, this research explores a broader range of well-established and newly developed formulas to compute k, offering a more efficient and interpretable alternative. Additionally, this work introduces fuzzy Liu estimation into the α-cut-based fuzzy regression context for the first time, applying 13 distinct methods for selecting the fuzzy Liu bias parameter d. Extensive evaluations were conducted using a wide variety of simulated scenarios and three real-world datasets to assess the robustness and accuracy of each method. The results demonstrate that formula-based approaches not only achieve performance comparable to K-fold cross-validation in fuzzy regression but also do so with significantly reduced computational cost. Moreover, the proposed α-cut-based fuzzy Ridge regression consistently outperformed both fuzzy Liu and fuzzy ordinary least squares methods in most evaluation settings. However, findings also indicate that each of the three estimation techniques—fuzzy Ridge, fuzzy Liu, and fuzzy ordinary least squares —exhibits unique strengths depending on the data characteristics and parameter configurations, highlighting their complementary roles within the proposed hybrid framework. These findings underscore the effectiveness of the hybrid model and emphasize the value of formula-driven bias parameter selection in advancing fuzzy regression techniques for complex, uncertain data environments.

Science Code	:	20512-20513
Key Words:	:	α-cut, fuzzy logic, multicollinearity, Ridge regression, Liu regression, ordinary least squares.
Page Number	:	106
Supervisor	:	Prof. Dr. Meral EBEGİL

ÇOKLU BAĞINTI DURUMUNDA BULANIK RİDGE VE BULANIK LİU REGRESYON MODELLERİNDE α-KESİM TABANLI YÖNTEMLER KULLANILARAK SHRİNKAGE PARAMETRELERİNİN ETKİN TAHMİNİ

(Doktora Tezi)

Ammar HOMAIDA

GAZİ ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

Haziran 2025

ÖZET

Bu çalışma, çoklu doğrusal bağlantıdan etkilenen bulanık veri kümelerini etkin bir şekilde işleyebilmek amacıyla, α-kesim-dayalı yaklaşım içinde Ridge ve Liu tahmin yöntemlerini birleştiren yeni bir hibrit bulanık regresyon çerçevesi sunmaktadır. Önceki çalışmalar, bulanık Ridge regresyonunda sapma parametresi k'nın belirlenmesinde K-katlı çapraz-doğrulamaya dayanırken, bu araştırma hem yerleşik hem de yeni geliştirilen çeşitli formülleri kullanarak k'yı hesaplamaya yönelik daha geniş kapsamlı ve yorumlanabilir bir alternatif sunmaktadır. Ayrıca, bu çalışmada bulanık Liu regresyonu ilk kez α-kesim-dayalı bulanık regresyon bağlamında uygulanmış ve sapma parametresi d'nin belirlenmesi için 13 farklı yöntem kullanılmıştır. Yöntemlerin sağlamlığı ve doğruluğunu değerlendirmek amacıyla çok sayıda simülasyon senaryosu ve üç gerçek dünya veri kümesi üzerinde kapsamlı analizler gerçekleştirilmiştir. Elde edilen sonuçlar, formül tabanlı yöntemlerin yalnızca K-katlı çapraz-doğrulama ile karşılaştırılabilir performans sağlamakla kalmayıp, aynı zamanda önemli ölçüde daha düşük hesaplama maliyetiyle bunu başardığını göstermektedir. Ayrıca, önerilen α-kesim-dayalı bulanık Ridge regresyonu, değerlendirme ortamlarının çoğunda hem bulanık Liu hem de bulanık en küçük kareler yöntemlerinden daha başarılı sonuçlar vermiştir. Bununla birlikte, bulgular; bulanık Ridge, bulanık Liu ve bulanık en küçük kareler olmak üzere üç tahmin yönteminin de, veri özellikleri ve parametre yapılandırmalarına bağlı olarak kendine özgü avantajlar sergilediğini ve önerilen hibrit çerçeve içinde tamamlayıcı roller üstlendiğini ortaya koymaktadır. Bu bulgular, hibrit modelin etkinliğini vurgulamakta ve sapma parametresinin formül tabanlı seçiminin, belirsizlik içeren karmaşık veri ortamlarında bulanık regresyon tekniklerinin geliştirilmesine katkı sağladığını göstermektedir.

1. Giriş

Bu çalışma, belirsizlik ve bulanıklık içeren veri yapılarının analizinde karşılaşılan temel zorlukları aşmayı hedefleyerek, özellikle çoklu bağlantı problemlerinin neden olduğu olumsuz etkileri minimize etmek amacıyla geliştirilmiş yeni bir hibrit bulanık regresyon modelini kapsamlı ve detaylı biçimde ele almaktadır. Günümüz veri analizi uygulamalarında, özellikle sosyal bilimler, ekonomi, mühendislik ve doğa bilimleri gibi alanlarda kullanılan veriler sıklıkla kesinlikten uzak, belirsiz ve bulanık özellikler taşımaktadır. Geleneksel istatistiksel yöntemler, bu tür verilerdeki karmaşıklığı ve belirsizliği tam olarak yansıtmakta yetersiz kalmakta, ayrıca çoklu doğrusal bağlantı

varlığında tahmin doğruluğu ve model kararlılığı önemli ölçüde azalmaktadır. Bu bağlamda, daha esnek, uyarlanabilir ve hesaplama açısından verimli modellerin geliştirilmesi ihtiyacı giderek artmıştır.

Çalışmada temel alınan α-kesim yaklaşımı, bulanık küme teorisinin güçlü bir bileşeni olarak, bulanık verilerin farklı kesimlerinde daha esnek ve hassas tahminlerin yapılmasını sağlamaktadır. Araştırma, bu yaklaşımı kullanarak Ridge ve Liu regresyon yöntemlerinin bulanık veri ortamına entegre edilmesini mümkün kılan yenilikçi bir hibrit model önerisi sunmaktadır. Böylece, model hem bulanık veri yapısının doğrudan analize dahil edilmesini sağlamakta hem de çoklu bağlantıdan kaynaklanan parametre sapmalarını etkili biçimde azaltarak tahminlerin güvenilirliğini artırmaktadır.

Bulanık mantıkla ilgili literatürde, regresyon analizindeki çoklu bağlantı problemlerine çözüm olarak sıklıkla K-katlı çapraz doğrulama gibi yoğun hesaplama gerektiren yöntemler kullanılmaktadır. Ancak bu yöntemlerin yüksek hesaplama maliyeti, özellikle büyük veri setleri veya yüksek boyutlu değişkenler söz konusu olduğunda pratikte uygulanabilirliğini kısıtlamaktadır. Bu noktada çalışma, yanlılık parametresi seçiminde formül tabanlı mekanizmalar geliştirmeyi amaçlamış ve böylelikle model doğruluğunu koruyup hesaplama verimliliğini önemli ölçüde artırmıştır. Bu yaklaşımla hem işlem süreleri kısalmış hem de modelin uygulamadaki pratikliği yükseltilmiştir. Ayrıca, çalışmanın önemli bir yeniliği olarak, bulanık Liu regresyonu α-kesim yaklaşımı ile kullanan ilk araştırma olması dikkat çekmektedir. Bu kapsamda, Liu regresyonunun yanlılık parametresi için 13 farklı seçim yöntemi titizlikle test edilmiş ve modelin farklı veri koşullarına uyum sağlama yeteneği kapsamlı biçimde incelenmiştir.

Modelin etkinliği, örneklem büyüklükleri, bağımlılık yapıları ve hata varyansları gibi parametreler sistematik biçimde değiştirilerek çok çeşitli simülasyon senaryoları oluşturularak test edilmiştir. Ayrıca, üç farklı gerçek veri seti üzerinde gerçekleştirilen uygulamalar ile modelin gerçek dünyadaki performansı ve pratik kullanımı da değerlendirilmiştir. Elde edilen sonuçlar, formül tabanlı yanlılık parametresi seçimlerinin yalnızca tahmin başarısı açısından değil, aynı zamanda işlem süresi ve hesaplama yükü bakımından da geleneksel yöntemlere göre açık bir üstünlük sunduğunu ortaya koymuştur. Özellikle α-kesim tabanlı bulanık Ridge regresyon yöntemi, çoklu bağlantının belirgin olduğu veri setlerinde birçok durumda bulanık Liu ve bulanık en küçük kareler yöntemlerine kıyasla daha başarılı ve kararlı sonuçlar vermiştir. Bununla birlikte, her bir yöntemin farklı veri yapıları ve koşullar altında farklı avantajlar ve sınırlamalar sergilediği görülmüş; bu durum, model seçiminde veri yapısının ve problem özelliklerinin dikkate alınmasının önemini vurgulamıştır.

Araştırma kapsamında farklı α-kesim dizilerinin kullanımı da modellenmiş ve yöntemler arasında performans farklılıklarının daha net ortaya çıkmasına olanak sağlanmıştır. α-kesim yaklaşımı, bulanık küme teorisinin temel bileşenlerinden biri olarak, bulanık verideki belirsizliklerin çoklu kesimler üzerinden incelenmesini mümkün kılarak tahminlerin duyarlılığını ve güvenilirliğini artırmaktadır. Bu, klasik yöntemlere göre çok daha esnek ve uyarlanabilir bir modelleme çerçevesi sunmakta, özellikle belirsizliklerin yoğun olduğu karmaşık veri setlerinde tahmin doğruluğunu yükseltmektedir.

Bulanık Ridge ve bulanık Liu regresyon yöntemleri, çoklu bağlantı problemi bulunan durumlarda α-kesim yöntemiyle birleştirildiğinde güçlü ve etkili araçlar olarak öne çıkmaktadır. Bu çalışmada bulanık Ridge regresyon için 77 farklı yanlılık parametresi

formülü, bulanık Liu regresyon için ise 13 farklı yanlılık parametresi formülü sistematik olarak test edilmiştir. Böyle geniş kapsamlı parametre denemeleri sayesinde, doğruluk ve hesaplama verimliliği arasında dengeli ve optimal çözümler geliştirilmiş; farklı veri yapıları ve koşullar altında en uygun parametrelerin nasıl seçileceğine ilişkin önemli bulgular elde edilmiştir. Bu durum, araştırmacılara parametre seçimi konusunda rehberlik etmekte ve modelin geniş bir kullanım alanı kazanmasına katkı sağlamaktadır.

Geleceğe dönük öneriler arasında, Gauss, trapez veya diğer üyelik fonksiyonu biçimleri gibi farklı bulanık sayı tiplerinin modellerle entegrasyonu ve bu sayede bulanık regresyon modellerinin kapsamının genişletilmesi yer almaktadır. Ayrıca, küçük örneklem büyüklüğüne sahip veya yüksek boyutlu değişken sayısına sahip karmaşık veri setlerinde α-kesim tabanlı parametre optimizasyonunun daha derinlemesine araştırılması ve optimize edilmesi, model performansının artırılması açısından kritik önemdedir. Bunun yanında, farklı disiplinlerden ve çeşitli yapısal özelliklere sahip gerçek veri setleri üzerinde modelin uygulanabilirliği, genellenebilirliği ve pratik faydalarının test edilmesi, hem akademik hem de uygulamalı alanlarda araştırmanın etkisini güçlendirecektir.

Sonuç olarak, bu çalışma hem teorik hem de uygulama açıdan bulanık regresyon yöntemlerinin geliştirilmesi ve çoklu bağlantı sorununun üstesinden gelinmesi hususunda önemli katkılar sunmaktadır. Aynı zamanda α-kesim yaklaşımının belirsizlik içeren veri analizlerinde güçlü, esnek ve hesaplama açısından verimli bir araç olduğunu ortaya koymaktadır. Önerilen hibrit model, bulanık verilerle çalışan araştırmacıların ve uygulayıcıların karşılaştığı temel problemleri aşmak için yeni perspektifler ve etkili çözümler sunmakta; istatistiksel modelleme ve veri bilimi alanlarında ileriye dönük araştırmalar için sağlam bir temel teşkil etmektedir.

2. Bulanık Mantık

Bulanık mantık, klasik mantığın katı ikili yapısının ötesine geçerek; belirsizlik, karmaşıklık ve kesinlikten uzak bilgilerin daha esnek ve gerçek yaşamı yansıtan biçimde modellenmesini sağlayan güçlü bir matematiksel mantık sistemidir.

Geleneksel ikili mantık sistemlerinde, herhangi bir önerme veya durum ya tamamen doğru (1) ya da tamamen yanlış (0) olarak değerlendirilir. Ancak insan düşünce yapısı ve günlük yaşamda karşılaşılan problemler çoğunlukla bu iki uç değer arasında yer alan, kesinlikten uzak, ara ve muğlak değerlere sahiptir. Örneğin, hava sıcaktır ifadesi mutlak olarak doğru ya da yanlış olmayabilir; sıcağın derecesi, kişinin algısına, bağlama ve zamana göre değişkenlik gösterebilir. İşte bulanık mantık, bu tür durumlarda, doğruluk değerlerini sadece 0 veya 1 olarak değil, 0 ile 1 arasında sürekli bir aralıkta ifade ederek ara değerlerin ve belirsizliklerin etkin bir şekilde temsil edilmesini mümkün kılar. Bu sayede, gerçek dünyadaki karmaşık ve belirsiz doğaya sahip problemler daha gerçekçi ve esnek biçimde modellenebilir.

Üyelik fonksiyonları, bulanık kümelerin tanımlanmasında merkezi bir rol oynar. Bu fonksiyonlar, üçgen, yamuk, Gauss, trapez gibi farklı şekillerde modellenebilir ve uygulamanın gereksinimlerine göre uyarlanabilir. Örneğin, üçgen üyelik fonksiyonu belirli bir noktada, genellikle 1 değerinde en yüksek üyelik derecesine ulaşırken, kenarlarda giderek azalan bir bağlılık gösterir. Gauss tipi üyelik fonksiyonları ise daha yumuşak geçişlere ve gerçek dünya problemine daha iyi uyum sağlar. Bu çeşitlilik, bulanık sistemlerin farklı veri tiplerine ve problem alanlarına esnek uyum sağlamasına olanak verir.

Gerçek dünya verileri genellikle kesin sayısal değerler olarak elde edilse de, bu değerlerin ölçüm hataları, belirsizlikleri veya doğası gereği taşıdığı muğlaklıklar vardır. Bulanık sistemlerde kesin veriler, bulanıklaştırma süreci ile belirli bir belirsizlik veya muğlaklık derecesi taşıyan bulanık değerlere dönüştürülür. Bulanıklaştırma, ham veriyi üyelik fonksiyonları ile temsil ederek, gerçek hayattaki doğrudan ölçülemeyen ya da kesin olarak ifade edilemeyen niteliklerin modellenmesini sağlar. Modelleme sürecinin sonunda, elde edilen bulanık sonuçlar ise netleştirme işlemiyle tekrar net, kesin sayılara çevrilir. Netleştirme teknikleri arasında ağırlıklı ortalama, maksimum üyelik derecesi veya maksimum üyeliklerin ortalaması gibi yöntemler bulunur. Doğru netleştirme yönteminin seçimi, modelin tahmin doğruluğu, karar verme kalitesi ve yorumlanabilirliği üzerinde doğrudan etkilidir.

Bulanık kümeler üzerinde gerçekleştirilen işlemler, klasik küme teorisindeki birleşim, kesişim ve fark işlemlerinden farklıdır; bu işlemler, elemanların üyelik derecelerine dayalı olarak tanımlanır. Örneğin, iki bulanık kümenin birleşimi, ilgili elemanların üyelik derecelerinin maksimumu ile, kesişimi ise minimumu ile ifade edilir. Ayrıca, bulanık sayılarla toplama, çıkarma, çarpma ve bölme gibi aritmetik işlemler genişletme prensibi ve α-kesim yöntemi gibi özel teknikler aracılığıyla uygulanır. Genişletme prensibi, bulanık sayıların işlemlerini üyelik fonksiyonlarının korunması koşuluyla klasik matematiksel işlemlere genişletir. α-kesim yöntemi ise bir bulanık sayıyı, üyelik fonksiyonunun belirli bir α seviyesi üzerinde keserek klasik küme kesimlerine dönüştürür ve böylece bu kesimler üzerinde klasik matematiksel işlemlerin yapılmasına imkân tanır. α-kesim yöntemi hem hesaplama kolaylığı sağlar hem de işlemlerin görselleştirilmesi ve yorumlanmasını kolaylaştırarak uygulamalarda önemli avantajlar sunar.

Bulanık mantık ve bulanık kümeler, karmaşık, kesinlikten uzak ve belirsiz düşünce süreçlerinin matematiksel olarak modellenebilmesi için geliştirilmiş kapsamlı araçlar olarak, günümüzde birçok bilim alanında yaygın bir şekilde kullanılmaktadır. Özellikle mühendislik, kontrol sistemleri, yapay zeka, karar destek sistemleri, ekonomi, finans, imaj işleme, robotik, veri madenciliği, makine öğrenmesi ve doğal dil işleme gibi disiplinlerde, belirsiz ve karmaşık verilerle başa çıkabilmek için bulanık mantık tabanlı modellerin uygulanması giderek yaygınlaşmaktadır. Bulanık mantık, insan benzeri esnek ve sezgisel karar verme mekanizmalarının modellenmesine de olanak sağlayarak, klasik yöntemlerin yetersiz kaldığı karmaşık sistemlerde anlamlı çözümler üretir.

Sonuç olarak, bulanık mantık modern veri bilimi ve yapay zeka uygulamalarının temel yapı taşlarından biri haline gelmiş, hem teorik hem pratik düzeyde belirsizlik yönetimi ve esnek modelleme alanlarında önemli ilerlemeler sağlamıştır. Belirsizliğin ve muğlaklığın etkin biçimde ele alınmasıyla, gerçek dünya problemlerinin daha doğru, güvenilir ve uygulanabilir çözümleri mümkün kılınmaktadır. Bu yönüyle, bulanık mantık ve bulanık kümeler, gelecekteki karmaşık sistemlerin tasarımında ve analitik yaklaşımların geliştirilmesinde vazgeçilmez araçlar olarak ön plana çıkmaktadır.

3. Klasik ve Bulanık Regresyon Yöntemleri

Klasik doğrusal regresyonda en yaygın kullanılan tahmin yöntemi en küçük kareler yöntemidir. Bu yöntem regresyon analizinin temel taşlarından biridir. En küçük kareler yöntemi, parametre tahmininde yanlılıktan uzak, minimum varyanslı sonuçlar sunması nedeniyle tercih edilmektedir. Ancak en küçük kareler yönteminin geçerli ve güvenilir sonuçlar verebilmesi için bazı temel varsayımların sağlanması zorunludur. Bu varsayımlar

arasında bağımlı ve bağımsız değişkenler arasındaki ilişkinin doğrusal olması, gözlemlerin birbirinden bağımsız ve aynı dağılıma sahip olması, hata terimlerinin ortalamasının sıfır olması, hata varyansının sabit olması, hata terimleri arasında otokorelasyon bulunmaması ve hata terimlerinin normal dağılım göstermesi yer alır. Ayrıca bağımsız değişkenler arasında ilişki olmaması da önemli bir varsayımdır. Bağımsız değişkenler arasındaki ilişki olması durumu çoklu bağlantı problemini ortaya çıkarır. Bu konu ise regresyon analizinde önemli bir araştırma ve analiz konusudur. Çoklu bağlantı, iki ya da daha fazla bağımsız değişkenin birbirleriyle ilişkili olması durumunda ortaya çıkar. Bu ilişkinin düzeyi arttıkça bu durum regresyon katsayılarının tahminindeki standart hataların büyümesine, tahminlerin güvenilirliğinin azalmasına ve sonuçların aşırı değişkenlik göstermesine yol açar. Sonuç olarak, modelin yorumlanması zorlaşır ve pratikte kullanılabilirliği azalır.

Çoklu bağlantının temel kaynakları arasında veri toplama sürecindeki eksiklikler, yetersiz örneklem büyüklüğü, doğası gereği birbirine bağlı değişkenlerin varlığı ve modelde yer alan değişken sayısının fazla olması sayılabilir. Çoklu bağlantı sorunu yalnızca modelin istatistiksel geçerliliğini değil, aynı zamanda tahmin başarısını ve genellenebilirliğini de olumsuz etkiler. Bu probleme bir çözüm yanlı tahmin yöntemlerinin kullanılmasıdır. Ridge regresyonu ve Liu regresyonu ise çoklu bağlantı sorununu çözmek için kullanılan yanlı tahmin yöntemlerindendir.

Ridge regresyon, klasik en küçük kareler yöntemine kıyasla regresyon katsayılarını küçültme işlemi uygular, böylece tahminlerin varyansı azalır, modelin aşırı öğrenmesi engellenir ve çoklu bağlantıdan kaynaklanan problemler hafifletilir. Ridge regresyonda küçültme işleminde kullanılan yanlılık parametresi k'nın doğru belirlenmesi tahmin başarısını doğrudan etkiler ve genellikle çapraz doğrulama gibi tekniklerle optimize edilir. Ridge regresyon yöntemi özellikle gözlem sayısının az, değişken sayısının çok ve değişkenler arasında orta ve orta üstü korelasyonun bulunduğu durumlarda avantajlıdır. Liu regresyonu ise Ridge regresyona benzer şekilde çoklu bağlantı sorununu azaltmayı hedefler ancak yanlılık parametresi d ile daha esnek bir denge kurar. Yanlılık parametresi d'nin seçimi, sapma-varyans dengesinin hassas bir şekilde ayarlanmasına olanak sağlar ve genellikle klasik yöntemlere kıyasla daha düşük hata oranları ve daha güvenilir tahminler elde edilmesini sağlar. Liu regresyonu hesaplama açısından genellikle daha basit bir yapıya sahiptir ve $d = 1$ durumunda klasik en küçük kareler yöntemine eşdeğer sonuçlar verir. Her iki yöntem de modeldeki tüm değişkenleri koruyarak ve tahminlerin kararlılığını artırarak çoklu bağlantının olumsuz etkilerini azaltır. Bu özelliklerinden dolayı Ridge ve Liu regresyonları, çoklu bağlantı probleminin sıkça görüldüğü istatistiksel modelleme uygulamalarında yaygın şekilde tercih edilmektedir.

Regresyon analizi genel olarak, bağımlı değişken ile bağımsız değişkenler arasındaki ilişkiyi nicel olarak modellemek için kullanılan temel istatistiksel yöntemlerden biridir. Gerçek dünyada karmaşık sistemlerde, insan kaynaklı süreçlerde veya ölçümlerde ortaya çıkan kesin olmayan, belirsiz ve bulanık bilgi yapılarının modellenmesi ihtiyacı doğmuştur. Bu bağlamda, klasik regresyona alternatif olarak geliştirilen bulanık regresyon teknikleri, belirsiz ve kesin olmayan veri yapılarında daha esnek, kapsamlı ve uyarlanabilir modelleme olanağı sağlar.

Klasik regresyonda model ile gözlenen değerler arasındaki farklar genellikle rastgele hata terimleri olarak kabul edilirken, bu farklar ölçüm hataları, sistemdeki yapısal belirsizlikler veya çevresel değişkenlerdeki netlikten uzak rastgele olmayan unsurlar nedeniyle de oluşabilir. Böyle durumlarda belirsizliğin rastgelelikten ziyade bulanıklık olarak ifade

edilmesi daha uygun olur. Bulanık küme teorisi, Lotfi Zadeh tarafından geliştirilmiş ve belirsiz değişkenlerin sayısal olarak temsil edilmesini, üyelik fonksiyonları yoluyla analizini mümkün kılmıştır. Bu teori temelinde geliştirilen bulanık regresyon modelleri, niteliksel değişkenleri ve belirsizlikleri üyelik dereceleri şeklinde modele entegre eder.

Bulanık regresyonda, merkez eğilimi klasik regresyona oldukça yakın sonuçlar verirken, yayılma eğilimi modelin esnekliğini artırır ve sistemdeki bireysel değişimleri, yapısal bulanıklığı yansıtır. Böylece klasik regresyonun hata terimleri için getirdiği sabit varyans varsayımı gevşetilir ve modelde farklı yayılma seviyeleri temsil edilebilir. Bu durum, klasik regresyonun kısıtlamalarından bağımsız olarak daha gerçekçi ve kapsamlı modellemeye imkân sağlar.

Bulanık regresyonun gelişimi, geleneksel yöntemlerin çok sayıda gözlem gerektirmesi, klasik ekonometrik varsayımların bulanık regresyonda esnetilmesi gibi ihtiyaçlara dayanmaktadır. Klasik yöntemlerde karmaşık dönüşümler ve veri ön işlemleri gerekebilirken, bulanık regresyon doğrudan ham veri seviyesinde çalışabilir. Sayısal olarak ifade edilemeyen sıralama, derecelendirme veya niteliksel değişkenlerde bulanık regresyon etkin bir alternatiftir. Ekonomik ilişkilerin kesin fonksiyonlarla ifade edilemediği karmaşık durumlarda, bulanık regresyon daha kapsayıcı ve esnek modellemelere imkân tanır.

Bulanık regresyon hem bulanık hem de kesin verileri aynı modelde işleyebilmekte ve veri yapısına göre uyarlanabilmektedir. Klasik regresyonda hata terimlerinin birbirinden bağımsız olması beklenirken, bulanık regresyonda bu zorunluluk kaldırılmıştır. Böylece belirsizlikler daha esnek biçimde modellenir. Klasik regresyon olasılık teorisine dayanırken, bulanık regresyon olasılık ve bulanık küme teorilerini bir arada kullanarak hem rastgele hem rastgele olmayan belirsizlikleri birlikte yönetme yeteneğine sahiptir.

Sonuç olarak, klasik regresyonda model ile veri arasındaki sapmalar rastgele hata olarak değerlendirilirken, bulanık regresyonda bu sapmalar modelin doğal ve sistematik bulanıklığı olarak yorumlanır. Bu farklı bakış açısı, veri-model ilişkisinde yeni bir perspektif sunarak, özellikle belirsizlik içeren karmaşık uygulamalarda daha gerçekçi ve kapsamlı analiz yapılmasını sağlar.

Bulanık regresyon, kullanılan tahmin yöntemleri ve veri türlerine göre çeşitli sınıflara ayrılır. Tahmin yöntemleri açısından, doğrusal programlama, amaç programlama ve aralık regresyonu gibi teknikleri kapsayan olasılıksal regresyon ile bulanık en küçük kareler yöntemi olmak üzere iki temel yaklaşım vardır. Bulanık en küçük kareler yöntemi, bulanık regresyon alanında yaygın kullanılan ve tamamlayıcı bir yöntemdir. Veri yapısına göre ise dört ana kategori tanımlanmıştır. Bunlar, kesin girdi-kesin çıktı, kesin girdi-bulanık çıktı, bulanık girdi-kesin çıktı ve hem bulanık girdi hem de bulanık çıktı içeren modellenebilir. Değişkenler arası ilişki yapısına göre ise bulanık regresyon modelleri doğrusal ve doğrusal olmayan biçimde sınıflandırılır. Uygulama amacına göre tanımlayıcı ve tahmine yönelik bulanık regresyon yöntemleri mevcuttur.

Bulanık en küçük kareler yönteminde amaç, gözlenen ve tahmin edilen değerleri arasındaki toplam kare hatayı minimize etmektir. Zaman içinde çeşitli revizyonlar ve geliştirmelerle bulanık regresyon analizinde merkezi bir konuma gelmiştir.

Bulanık sayıların oluşturduğu uzayın tam bir metrik uzay olduğu Öklid uzaklığı ile doğrulanmış olup, negatif olmama, simetriklik ve üçgen eşitsizliği gibi temel metrik özellikleri sağlanmaktadır.

Bulanık doğrusal regresyon modeli, klasik modellerde karşılaşılan çoklu doğrusal bağlantı sorunlarını da kapsar. Çoklu bağlantı, en küçük kareler tahmin edicisinin tanımsız veya tahminin gerçekten uzak olmasına yol açarken, Ridge regresyon etkisi az olan değişkenlerin katsayılarını küçülterek varyansı stabilize eder ve sorunu hafifletir. Benzer şekilde bulanık doğrusal regresyon modellerinde de çoklu bağlantı önemli bir problemdir. Bulanık verilerde bağımsız değişken sayısı genellikle sınırlı olduğundan değişken eleme yöntemleri pratik olmaz. Bu nedenle, bulanık sayılar arasındaki uzaklığa dayalı yöntemlerle bulanık Ridge tahmin edicileri geliştirilmiştir; ancak uygulama alanları sınırlıdır.

Daha önce geliştirilen α-kesim tabanlı tahmin yöntemi, klasik Ridge ve klasik en küçük kareler tahminlerini bulanık küme teorisinin temel kavramlarından biri olan α-kesim yaklaşımıyla genişletmektedir. Bu yöntem, bulanık kümelerin sadece alt ve üst sınırlarına değil, çoklu kesimlerine odaklanarak esnek tahmin olanağı sunar. Böylece, üçgen veya diğer şekillerde tanımlanan bulanık bağımlı değişkenin olası değer aralıkları daha kapsamlı analiz edilir.

Doğrusal regresyonda, α-kesim tabanlı tahmin süreci altı adımdan oluşur: İlk olarak, üçgen bulanık bağımlı ve bağımsız değişkenler için α-kesim kümeleri oluşturulur. α değerleri [0,1] aralığında seçilir ve değişkenler için alt ve üst sınırlar tanımlanır. İkinci adımda, bu α-kesimlerden elde edilen kesin değerler üzerinde en küçük kareler, Ridge veya Liu gibi klasik tahmin yöntemleri uygulanır. Üçüncü adımda, ara tahminlerden üçgen bulanık sayı üyelik fonksiyonuna uygun tahminler elde edilir ve bu modifikasyonlarla tahminlerin üçgen bulanık sayı özelliklerini sağlaması garanti edilir. Dördüncü adımda, belirli noktalar üzerinde doğrusal regresyon uygulanarak üçgen bulanık katsayı tahmin edilir ve üyelik fonksiyonunun üçgen bulanık sayı koşullarını koruması için kısıtlamalar getirilir. Beşinci adımda, simetrik bulanık girdi veya çıktı durumlarında tahmin edilen üyelik fonksiyonunun simetrik olması garanti edilmez; bu yüzden gerçek ve tahmin edilen değerler arasındaki farkı azaltmak için iki aday üyelik fonksiyonu değerlendirilir. Son adımda ise önerilen bulanık uzaklık ölçütüne dayalı kriter kullanılarak performans değerlendirilir ve en düşük değeri veren fonksiyon seçilir. Bu α-kesim tabanlı bulanık tahmin algoritması, kesin girdi ve bulanık çıktı varsayımıyla çalışır. Seçilen α-kesim değerlerine göre sol, sağ uç noktalar ile merkez için klasik tahminler uygulanır; elde edilen katsayıların üçgen bulanık sayı koşullarını sağlayıp sağlamadığı kontrol edilip gerekirse düzeltilir. Daha sonra katsayıların simetrikliği optimize edilir ve performans en yakın uzaklık kriterine göre değerlendirilir. Bu yöntem, bulanık katsayı tahmininde doğruluk ve hesaplama etkinliği açısından optimal sonuçlar sunar.

Bu çalışmada, α-kesim tabanlı bulanık regresyon yöntemi önceki çalışmalardan genişletilerek önemli sınırlamalar ele alınmakta ve çerçeveye yeni bileşenler eklenmektedir. Önceki araştırmalar genellikle tek bir örneklem büyüklüğüne odaklanırken, bu çalışma farklı örneklem büyüklükleri üzerinden kapsamlı değerlendirmeler yaparak algoritmanın veri hacmine duyarlılığını da incelemektedir. Ayrıca, bulanık Ridge regresyonda yanlılık parametresi k'nın seçimine sistematik ve kapsamlı bir yaklaşım getirilmekte, önceki çalışmalarda detaylandırılmamış bu konu farklı teknikler kullanılarak ele alınmaktadır. Böylece yanlılık parametresi k'nın hesaplanması için K-katlı çapraz doğrulama gibi uzun ve

yoğun yöntemlere ihtiyaç duyulmadan, daha etkin, güvenilir ve pratik sonuçlar elde edilmektedir.

Önemli bir yenilik olarak, bu çalışma α-kesim tabanlı bulanık regresyon çerçevesine bulanık Liu tahmin edicisini entegre eden ilk çalışmadır. Bulanık Liu regresyon, bulanık Ridge regresyondan farklı bir sapma yapısı getirir ve yanlılık parametresi d'nin seçimine odaklanır. Buradan yola çıkarak çalışmada, d parametresinin etkileri ve seçimi üzerine detaylı deneysel ve teorik analizler yapılmıştır. Böylece, hem Ridge hem de Liu tahmin edicilerinin bulanık versiyonları aynı α-kesim tabanlı platformda karşılaştırılarak, en uygun yanlılık parametre seçimleri elde edilmiştir. Ayrıca, α-kesim tabanlı bulanık en küçük kareler ile kıyaslamalar da yapılarak farklı yöntemlerin performansları, doğrulukları ve pratik kullanımları ortaya konulmuştur.

Son olarak, çalışma bulanık ve kesin regresyon sonuçlarının yorumlanması konusunda yenilikçi yaklaşımlar geliştirmektedir. Bu kapsamda, bulanık katsayılar ve tahminlerin grafiksel yorumlanması ve karşılaştırılması için açık, anlaşılır ve kullanıcı dostu yöntemler önerilmektedir. Böylece, alan dışı kullanıcılar ve uygulayıcılar da sonuçları daha kolay yorumlayabilmekte ve modellerin pratik işlevselliği artmaktadır.

4. Simülasyon Çalışması

Bu bölüm, α-kesim tabanlı bulanık Ridge regresyon yöntemine dayanan temel çalışmalardan esinlenerek geliştirilmiş, kapsamı genişletilmiş ve detaylandırılmış bir simülasyon çalışması sunulmaktadır. Çalışmanın bu bölümünde, bulanık regresyon alanında öne çıkan, bulanık en küçük kareler regresyonu, bulanık Ridge regresyonu ve bulanık Liu regresyonu gibi üç farklı yöntemin performanslarının karşılaştırmalı olarak incelemesi hedeflemektedir. Bu amaçla, farklı veri yapıları, belirsizlik koşulları ve parametrik değişkenlik altında bu yöntemlerin tahmin doğruluğu ve tutarlılığı objektif ve kapsamlı biçimde değerlendirilmektir.

Simülasyon tasarımı, gerçek dünya verilerinde sıkça rastlanan bulanık ve belirsiz veri özelliklerini yansıtacak şekilde dört bağımsız değişken kullanılarak yapılandırılmıştır. Örneklem büyüklükleri, istatistiksel analizlerde yaygın şekilde tercih edilen ve küçükten büyüğe doğru sırasıyla $n = 20,50,100,200$ olarak seçilmiştir. Bu örneklem büyüklükleri, yöntemlerin veri hacmine bağlı olarak nasıl performans gösterdiklerini ortaya koymak, özellikle küçük ve orta büyüklükteki veri setlerinde tahminlerin güvenilirliğini incelemek amacıyla belirlenmiştir.

Bağımsız değişkenler arasındaki bağımlılık yapısı, simülasyon çalışmasının bir diğer kritik parametrelerinden biri olarak ele alınmıştır. Bu nedenle, bağımlılık dereceleri orta, yüksek ve çok yüksek korelasyon olmak üzere üç farklı düzeyde modellenmiştir. Bu çeşitlendirme, çoklu doğrusal bağlantı probleminin etkilerini simüle etmek ve yöntemlerin bu tür zorluklara karşı dayanıklılığını ölçmek için önem taşımaktadır.

Bağımlı değişken olarak kullanılan bulanık değişken, önceden belirlenmiş regresyon katsayıları ve bağımsız değişkenlerin doğrusal kombinasyonu ile oluşturulmuştur. Bu modele, farklı varyans seviyelerinde hata terimleri eklenerek belirsizlik ve gürültü düzeyleri çeşitlendirilmiştir. Yüksek hata varyansı, gözlemler arasında daha yoğun bir belirsizliğe ve karmaşıklığa işaret ederken, düşük hata varyansı daha temiz ve güvenilir veri koşullarını simgeler. Böylece, yöntemlerin farklı hata düzeylerindeki dayanıklılıkları ve tahmin doğrulukları test edilmiş olur.

Bulanık değişkenlerin yayılım yapısı da simülasyonda ayrı bir parametre olarak ele alınmıştır. İki temel yayılım senaryosu incelenmiştir: simetrik yayılım ve asimetrik yayılım. Simetrik yayılımda, bulanık sayının sağ ve sol yayılım uzunlukları eşit kabul edilirken, asimetrik yayılımda bu iki yayılım birbirinden farklı büyüklükte belirlenmiştir. Bu yaklaşım, gerçek dünyada sıkça karşılaşılan çeşitli bulanık veri biçimlerinin modellenmesine olanak tanımakta ve yöntemlerin bu çeşitlilik karşısındaki performans farklılıklarını ortaya koymaktadır.

Simülasyon parametrelerinin belirlenmesinde, regresyon katsayıları sabit tutulmuştur. Bu sayede, yöntemlerin performansları üzerinde sadece bağımlılık düzeyi, hata varyansı, örneklem büyüklüğü ve yayılım tipi gibi deneysel olarak kontrol edilen faktörlerin etkisi gözlemlenmiştir. Böylece analiz, model parametrelerindeki değişkenlikten kaynaklanabilecek karmaşıklıklardan arındırılmış ve yöntemlerin karşılaştırmalı başarısı daha net ortaya konmuştur.

Bağımlılık düzeyi, hata varyansı ve örneklem büyüklüğü parametrelerinin etkilerini detaylı incelemek için α-kesim değerleri sabit tutulmuş ve farklı aralıklarda seçilmiştir. Her bir parametre kombinasyonu için 1000 tekrar sayısına sahip simülasyonlar gerçekleştirilmiştir. Her bir tekrar, bağımsız değişkenler ve katsayıların sabit kalması ancak hata terimlerinde rastgelelik olması nedeniyle, sonuçların genel geçerliliğini artırmakta ve yöntemlerin tutarlılığını güvenilir biçimde ortaya koymaktadır.

Simülasyon verileri, eğitim ve test olmak üzere iki ayrı veri setine bölünmüştür. Üç farklı bulanık regresyon yöntemi, eğitim setinde modellenmiş ve elde edilen modeller test setinde doğrulanmıştır. Model performansının objektif değerlendirilmesi için, tahmin hatalarının büyüklüğünü doğrudan ve kolay yorumlanabilir biçimde ölçen hata karelerinin ortalamasının karekökü kriteri kullanılmıştır. Hata karelerinin ortalamasının karekökü, istatistiksel model karşılaştırmalarında standart bir başarı ölçütü olup, tahminlerin ortalama sapmasını anlamak için uygundur.

Simülasyon çalışmasının toplamda 92 farklı parametre ve yöntem kombinasyonunu içermesi, sonuçların geniş kapsamlı ve farklı koşulları temsil eden bir yapıda değerlendirilmesine olanak sağlamıştır. Elde edilen sonuçlar, tablo biçiminde detaylı olarak sunulmuş ve her bir yöntemin başarı sıralaması, hata büyüklükleri ile farklı parametrelerin etkileri açık biçimde görselleştirilmiştir. Bu kapsamlı analiz, yöntemlerin hangi koşullarda daha başarılı ve güvenilir olduklarına dair kapsamlı bilgiler sağlamaktadır.

Bununla birlikte, simülasyon çalışması sadece yöntemlerin genel başarılarını karşılaştırmakla kalmayıp, α-kesim tabanlı bulanık regresyon modellerinin pratikteki uygulanabilirliği ve etkinliği hakkında da önemli çıkarımlar yapmaktadır. Literatürde α-kesim yöntemlerinin bulanık regresyondaki etkinliği giderek artan bir şekilde vurgulanmakta olup, bu çalışmanın sonuçları da bu doğrultuda değerli ampirik kanıtlar sunmaktadır.

Özet olarak simülasyon çalışmasının yapıldığı bölümde sunulan simülasyon sonuçları, farklı hata varyansları, örneklem büyüklükleri ve veri simetrisi koşullarında bulanık Ridge regresyon, bulanık Liu regresyon ve bulanık en küçük kareler regresyon yöntemlerinin karşılaştırılmalı performanslarını kapsamaktadır. α-kesim dizisi tüm analiz boyunca sabit tutulmuş ve böylece yalnızca diğer faktörlerin etkilerinin izole edilmesi amaçlanmıştır.

Simülasyon çalışmasıyla elde edilen bulgulara göre, özellikle korelasyon katsayısının 0.80 gibi orta düzeyde olduğu durumlarda ve bağımlı değişkenin asimetrik üçgensel bulanık sayılarla ifade edildiği senaryolarda, bulanık Ridge regresyonun yanlılık parametresi k'yı tahmin etmek için literatürde sunulan formüller genel olarak en düşük hata ölçütlerine ulaşmıştır. Bu yöntemler, özellikle küçük hata varyanslarında ve artan örneklem büyüklüklerinde oldukça tutarlı sonuçlar vermiş olup, bazı durumlarda K-katlı çapraz doğrulama yaklaşımıyla benzer düzeyde performans göstermişlerdir. Ancak formül tabanlı durumların dikkat çeken bir avantajı, çok daha kısa hesaplama süresine sahip olmasıdır. Buna karşın çapraz doğrulama yöntemi, her durumda en uygun yanlılık parametresini bulamayabilmekte ve özellikle düşük örneklem büyüklüğü veya yüksek varyans durumlarında karmaşık matris işlemleri gerektirerek hesaplama yükünü artırabilmektedir.

Simülasyon çalışması bulanık Liu regresyonu açısından değerlendirildiğinde, formül tabanlı yanlılık parametrelerinin özellikle düşük hata varyansı ve büyük örneklemlerde iyi performans sergilediği görülmüştür. Bununla birlikte, tahmin yöntemlerinin performansına bakıldığında, genel olarak bulanık Ridge regresyonun formül tabanlı yöntemlerinin gerisinde kaldığı görülmüştür. Bulanık Ridge ve bulanık Liu yöntemlerinin başarısı, yanlılık parametre seçiminin veri özelliklerine duyarlı şekilde yapılmasının model başarısını doğrudan etkilediğini göstermektedir.

Bulanık en küçük kareler yöntemi ise genel olarak en düşük performansa sahip yöntem olarak öne çıkmıştır. Bulanık Ridge regresyonu, yanlılık parametresi sıfır seçildiğinde bu yönteme indirgenebilir nitelikte olduğundan, en küçük kareler yönteminin performansı dolaylı olarak bulanık Ridge yöntemlerinin özel durumu olarak da değerlendirilmiştir.

Korelasyon düzeyinin 0.90 gibi yüksek olduğu senaryolarda, çapraz doğrulama yöntemi daha fazla senaryoda daha iyi sonuçlar sağlamış; fakat bazı formül tabanlı parametre seçimleriyle elde edilen sonuçlar da oldukça yakın düzeyde kalmıştır. Bu bağlamda, bulanık Ridge yöntemi için formüle dayalı yanlılık parametrelerinin yalnızca hızlı değil, aynı zamanda yeterince etkili bir alternatif sunduğu görülmektedir. Bulanık Liu regresyonu ise benzer senaryolarda belirli varyans düzeylerinde kabul edilebilir düzeyde hata oranlarına ulaşmış, ancak varyans arttıkça başarısı azalmıştır.

Bağımlı değişkenin simetrik üçgensel bulanık sayılarla ifade edilen durumlar gibi simetrik olduğu senaryolarda, bulanık Ridge regresyon yöntemi diğer yöntemlere göre daha güçlü bir performans sergilemiştir. Bununla birlikte, bulanık Ridge regresyon yönteminin performansı asimetrik veri koşullarına çapraz doğrulama yönteminin performansına göre daha zayıf kalmıştır. Bu durum, bazı formül tabanlı parametre seçimlerinin daha istikrarlı ve veri yapısından bağımsız çalıştığını göstermektedir. Bulanık Liu yöntemi ise, simetrik yapıda bazı senaryolarda çapraz doğrulama ile benzer sonuçlar üretse de, genel sıralamada daha geride kalmıştır.

Korelasyonun 0.99 gibi çok yüksek olduğu yani ciddi çoklu bağlantı içeren senaryolarda, çapraz doğrulama yöntemi düşük varyans değerlerinde diğer iki yönteme göre en iyi sonuçları vermiştir. Ancak, bu sonuçlara belirli formül tabanlı parametre seçimleriyle de ulaşmak mümkün olduğundan, hesaplama verimliliği açısından önemli bir avantaj elde edilmiştir. Yüksek varyans değerlerinde ise özellikle bulanık Ridge ve bulanık Liu yöntemlerinde hata oranları önemli ölçüde artarken, bazı formüller diğerlerine göre daha kararlı sonuçlar üretebilmiştir. Bulanık en küçük kareler yöntemi bu koşullarda da zayıf bir performans sergilemiştir.

Simülasyon çalışması sonuçlarından genel olarak, bağımlı değişkenin simetrik yapıda tanımlandığı senaryolarda elde edilen hata oranlarının, asimetrik yapıya göre daha yüksek olduğu gözlenmiştir. Bu durum, modelleme sürecinde asimetri özelliğinin dikkate alınmasının önemine işaret etmektedir. Ayrıca, beklenildiği gibi daha büyük örneklem büyüklüklerinin, sistematik olarak daha düşük hata paylarına sahip oldukları ve dolayısıyla bu durumun model kararlılığını artırdığını da göstermektedir.

Sonuç olarak simülasyon çalışması, bulanık regresyon alanında hem akademik literatüre hem de pratik uygulamalara önemli katkılar sunmakta; farklı örneklem büyüklükleri, bağımlılık yapıları, hata varyansları ve bulanık yayılım türleri gibi gerçekçi veri koşullarında bulanık Ridge, bulanık Liu ve bulanık en küçük kareler yöntemlerinin karşılaştırmalı performanslarını detaylı biçimde ortaya koymaktadır. Ayrıca, çalışma α-kesim tabanlı yöntemlerin performansını artıracak parametre seçimi ve model geliştirme süreçleri için yol gösterici niteliktedir.

5. Ampirik Uygulama

Bu bölümde, α-kesim tabanlı tahmin algoritması temel alınarak geliştirilmiş olan bulanık Ridge, bulanık Liu ve bulanık en küçük kareler regresyon yöntemlerinin performansları, üç farklı gerçek veri seti üzerinde kapsamlı ve karşılaştırmalı bir şekilde ele alınmaktadır. İncelenen veri setleri, hem bulanık veri yapıları hem de çoklu doğrusal bağlantı problemleri açısından birbirinden farklı özellikler taşıması sebebiyle, yöntemlerin değişken veri koşullarındaki etkinlik ve dayanıklılık düzeylerini ortaya koymak açısından büyük önem taşımaktadır. Böylece, bu yöntemlerin pratik uygulamalarda karşılaşabilecekleri çeşitli zorluklara karşı tahmin başarısı ve stabilitesi detaylı biçimde analiz edilmiştir.

İlk olarak değerlendirilen veri seti, Çin'in ekonomik açıdan önemli kentlerinden biri olan Shanghai'ye ait konut fiyatlarını kapsamaktadır. Toplamda 147 gözlem içeren bu veri setinde 6 adet kesin bağımsız değişken yer almakta, bağımlı değişken ise asimetrik özellikler gösteren bulanık konut fiyatı olarak tanımlanmaktadır. Veri seti üzerinde gerçekleştirilen ön analizler sonucunda, bağımsız değişkenler arasında anlamlı bir çoklu doğrusal bağlantı problemi gözlemlenmemiştir. Bu durum, klasik regresyon yöntemlerinin temel varsayımlarının büyük ölçüde sağlandığını ve dolayısıyla bulanık en küçük kareler tabanlı yöntemlerin geçerli tahminler üretebileceği uygun bir ortam oluşturduğunu göstermektedir.

Bu çalışmada kullanılan ilk veri seti üzerinde gerçekleştirilen modelleme ve tahmin analizleri, özellikle bulanık Ridge regresyon yönteminin diğer yöntemlere kıyasla daha başarılı sonuçlar verdiğini göstermektedir. Bulanık Ridge regresyonunun sunduğu düzenlileştirme mekanizması sayesinde, hata ölçütleri anlamlı düzeyde düşmüş ve bu yöntem veri seti üzerindeki genel başarısıyla öne çıkmıştır. Buna karşılık, bulanık Liu regresyonuna ve bulanık Ridge regresyonuna dayalı çapraz doğrulama yaklaşımı daha yüksek hata oranlarına ulaşarak performans açısından ikinci planda kalmıştır. Bulanık en küçük kareler yöntemi ise, bu veri ortamında en zayıf sonuçları üretmiş ve özellikle belirsizliği ifade etme açısından yetersiz kalmıştır. Ayrıca, α-kesim dizisinin doğru seçilmesinin model performansına önemli katkı sağladığı görülmüş; uygun aralıklarla seçilen kesim dizileri hata oranlarında belirgin iyileşmelere yol açmıştır.

İkinci veri seti olarak literatürde bilinen klasik bir örnek kullanılmış ve burada da bulanık Ridge regresyonunun genel olarak en iyi performansı sunduğu gözlemlenmiştir. Bu veri setinde yer alan orta düzeydeki çoklu doğrusal bağlantı problemi, bazı yöntemlerin

parametre kesiminde kararsızlığa neden olmuş; ancak bulanık Ridge yönteminin düzenlileştirme özelliği bu olumsuzluğu dengeleyerek düşük hata değerleriyle dikkat çekmiştir. Bulanık en küçük kareler yöntemi, yine en zayıf sonuçları üretmiş; özellikle verinin yayılım bileşenini doğru biçimde tahmin etmede başarısız olmuştur. Bulanık Liu ve çapraz doğrulama yöntemleri ise ortalama düzeyde bir performans sergilemiş, özellikle merkez değer tahminlerinde kabul edilebilir sonuçlar üretmiştir.

Üçüncü veri setinde ise yüksek düzeyde çoklu bağlantı problemi içeren bir yapı söz konusudur. Bu durum, model tahminlerinin kararlılığını ciddi şekilde etkilemiştir. Bu veri seti üzerinde yapılan analizler, bulanık Ridge regresyonunun diğer yöntemlere göre çok daha kararlı ve güvenilir sonuçlar verdiğini bir kez daha ortaya koymuştur. Özellikle düzenlileştirme etkisi, bu gibi karmaşık yapılı verilerde modelin istikrarını artırmış ve hata oranlarını düşük seviyede tutmuştur. Bulanık Liu ve bulanık en küçük kareler yöntemleri ise benzer ancak daha düşük düzeyde performans göstermiştir.

Genel olarak tüm veri setleri ve yöntemler üzerinde yapılan kapsamlı analizler, bulanık Ridge regresyonunun hem merkez hem de yayılım bileşenlerini başarılı şekilde tahmin etmede güçlü bir araç olduğunu ortaya koymaktadır. Bu yöntem, özellikle çoklu doğrusal bağlantının bulunduğu ortamlarda yüksek doğruluk ve düşük hata oranları sağlamış, veri yapısına duyarlı parametre seçimi sayesinde tutarlı bir genel performans sunmuştur.

Öte yandan, bulanık en küçük kareler yöntemi merkez tahminlerde zaman zaman iyi sonuçlar verse de, belirsizlik içeren yayılım bileşenini yeterince temsil edememiştir. Bu durum, modelin bulanık veri yapısının karmaşıklığını tam anlamıyla yansıtamamasından kaynaklanmaktadır. Bulanık Liu regresyonu ve çapraz doğrulama tabanlı bulanık Ridge regresyonu yöntemleri ise bazı senaryolarda istikrarlı ve kabul edilebilir sonuçlar üretmiş, ancak formül tabanlı bulanık Ridge yönteminin gerisinde kalmıştır. Bununla birlikte, bulanık Liu yönteminin esnek yapısı sayesinde bazı durumlarda alternatif bir seçenek olabileceği de gözlemlenmiştir.

Tüm bu bulgular, α-kesim dizilerinin model başarısı üzerindeki etkisini açıkça ortaya koymaktadır. Uygun biçimde seçilmiş kesim dizileri, yalnızca hata oranlarını düşürmekle kalmayıp aynı zamanda model çıktılarının yorumlanabilirliğini ve güvenirliğini de artırmaktadır.

6. Sonuçlar ve Öneriler

Bu çalışmada, α-kesim tabanlı tahmin algoritması temel alınarak geliştirilmiş üç farklı bulanık regresyon yöntemi olan bulanık Ridge, bulanık Liu ve bulanık en küçük kareler regresyon yöntemleri kapsamlı ve sistematik bir şekilde değerlendirilmiştir. Analizler, hem kontrollü simülasyon ortamında oluşturulan yapay veri setleri hem de üç farklı gerçek veri seti kullanılarak gerçekleştirilmiştir. Özellikle çoklu doğrusal bağlantı problemlerinin olduğu durumlarda bu yöntemlerin tahmin doğruluğu, hesaplama etkinliği ve belirsiz veri yapılarındaki performansları detaylı biçimde incelenmiş, karşılaştırılmış ve yöntemlerin avantajları ile kısıtları ortaya konmuştur.

Elde edilen bulgular, k parametresi formüllerinin kullanıldığı bulanık Ridge regresyon yönteminin genel olarak diğer yöntemlere göre en yüksek başarıyı ortaya koyduğunu göstermiştir. Bulanık Ridge yöntemi, özellikle çoklu bağlantı problemlerinin yoğun olarak bulunduğu veri setlerinde tahminlerin kararlılığını önemli ölçüde artırmakta ve hata

oranlarını anlamlı biçimde düşürmektedir. Bu başarının temelinde, bulanık Ridge regresyonunun kullandığı düzenlileştirme mekanizması yatmaktadır. Bu mekanizma, parametre tahminlerindeki aşırı varyansı kontrol altına alarak modelin genellenebilirliğini artırmaktadır. Öte yandan, K-katlı çapraz doğrulama yöntemi ise bazı durumlarda iyi sonuçlar elde edilmesini sağlamasına rağmen, hesaplama maliyetinin yüksekliği ve yineleme sayısının artması sebebiyle pratik uygulamalarda özellikle büyük veri setlerinde sınırlamalar getirmektedir. Bu durum, özellikle gerçek zamanlı analiz gerektiren veya büyük çaplı veri ile çalışılan senaryolarda önemli bir kısıtlayıcı faktör olarak değerlendirilmektedir.

Bulanık Liu regresyon yöntemi ise genellikle istikrarlı ve tutarlı sonuçlar sunmakla birlikte, bulanık Ridge regresyonun sağladığı performans seviyesine tam anlamıyla ulaşamamıştır. Bununla birlikte, bulanık Liu yönteminin esnek yanlılık parametresi, farklı veri yapıları altında bulanık Ridge yöntemine yakın veya bazı durumlarda daha iyi tahmin sonuçları elde edilmesine imkân vermektedir. Bu özellik, bulanık Liu regresyonunun alternatif ve tamamlayıcı bir yöntem olarak değerlendirilmesini mümkün kılmaktadır. Bulanık en küçük kareler yöntemi ise çoğunlukla diğer yöntemlere kıyasla daha düşük tahmin doğruluğu göstermiştir. Bu performans farkının temel nedeni, bulanık Ridge regresyonunun yanlılık parametresi k sıfırlandığında bulanık en küçük karelere indirgenebilmesi ve böylece bulanık Ridge regresyonunun en küçük kareler yönteminin daha güçlü, esnek ve genelleştirilmiş bir versiyonu olarak tanımlanabilmesidir. Bulanık en küçük kareler yönteminin belirsizlik boyutunu tam olarak modelleyememesi ve çoklu bağlantı durumlarındaki yetersizliği, onun tahmin başarısını sınırlayan önemli etkenlerdir.

Çalışmada ayrıca, farklı α-kesim dizilerinin kullanımının yöntemler arasındaki performans farklarının daha belirgin hale gelmesini sağladığı gözlemlenmiştir. α-kesim yaklaşımı, bulanık küme teorisinin temel bileşenlerinden biri olarak, verideki belirsizliklerin çoklu kesimler üzerinden daha hassas ve esnek bir şekilde modellenmesine olanak tanımaktadır. Bu sayede, tahminlerin duyarlılığı ve güvenilirliği anlamlı biçimde artırılmaktadır. Dolayısıyla, α-kesim dizilerinin seçimi, bulanık regresyon modellerinin etkinliğinde kritik bir rol oynamaktadır.

Bulanık Ridge ve bulanık Liu regresyon yöntemleri, özellikle çoklu bağlantı sorunu bulunan karmaşık veri yapılarında α-kesim tabanlı modelleme ile birleştiğinde oldukça etkili ve güçlü araçlar olarak öne çıkmaktadır. Bu çalışmada, Ridge yöntemi için 77 farklı k parametresi formülü, Liu yöntemi için ise 13 farklı d parametresi formülü kapsamlı bir şekilde test edilmiştir. Bu geniş parametre seti, doğruluk ve hesaplama verimliliği arasında dengeli ve optimal bir denge kurulmasına olanak tanımış, farklı senaryolarda en uygun parametre seçimlerinin nasıl yapılabileceği üzerine önemli bulgular sağlamıştır.

Geleceğe yönelik araştırmalar için öneriler arasında, Gauss tipi, trapez biçimli veya diğer üyelik fonksiyonu yapıları gibi farklı bulanık sayı türlerinin modellenmesi ve böylece bulanık regresyon modellerinin kapsamının genişletilmesi yer almaktadır. Özellikle, küçük örneklem büyüklüğüne sahip veri setleri veya yüksek boyutlu değişken sayısına sahip karmaşık veri yapılarında α-kesim tabanlı parametre optimizasyonunun daha derinlemesine incelenmesi ve bu alandaki algoritmaların geliştirilmesi, model performansının artırılması açısından kritik görülmektedir. Ayrıca, değişik disiplinlerden ve farklı yapısal özelliklere sahip gerçek veri setleri üzerinde modelin uygulanabilirliğinin ve genellenebilirliğinin test edilmesi, araştırmanın pratik değerini ve yaygın kullanım potansiyelini artıracaktır.

Sonuç olarak, bu çalışma bulanık regresyon alanında α-kesim tabanlı yöntemlerin özellikle çoklu bağlantı ve belirsizlik sorunlarına karşı güçlü çözümler sunduğunu ortaya koymakta, araştırmacılara ve uygulayıcılara yöntem seçimi ve parametre ayarı konusunda önemli rehberlik sağlamaktadır. Ayrıca, geliştirilen yaklaşımların pratik uygulamalarda ve farklı disiplinlerde geniş bir kullanım potansiyeline sahip olduğu görülmektedir. Gelecekteki çalışmaların, mevcut yöntemlerin kapsamını genişletmek, performansını artırmak ve yeni bulanık veri türleri için uyarlanabilirlik sağlamak yönünde yoğunlaşması önerilmektedir.

Bilim Kodu	:	20512-20513
Anahtar Kelimeler	:	α-kesim, bulanık mantık, çoklu bağlantı, Ridge regresyonu, Liu regresyonu, en küçük kareler.
Sayfa Adedi	:	106
Danışman	:	Prof. Dr. Meral EBEGİL

ACKNOWLEDGEMENTS

I would like to express my deepest and heartfelt gratitude to my supervisor, Prof. Dr. Meral EBEGIL, whose exceptional mentorship, continuous encouragement, and dedicated support have been central to the completion of this thesis. Her profound academic knowledge, thoughtful guidance, and patience have inspired and challenged me throughout my research journey, and I am truly fortunate to have worked under her supervision. I am also sincerely grateful to Assoc. Prof. Mustafa Hilmi PEKALP, who not only served on my Thesis Monitoring Committee but also acted as a second supervisor, offering critical insights, constructive feedback, and generous support that significantly enhanced the quality and direction of my work. My sincere thanks go as well to Prof. Dr. Fikri GÖKPINAR, a valued member of the Thesis Monitoring Committee, for his thoughtful input and support during this process. I would also like to thank the jury members of my thesis defense for their careful evaluation and helpful recommendations. I extend my gratitude to Gazi University, the Department of Statistics, and all the academic staff who guided and supported me throughout my education. As a foreign student, I am deeply thankful to Türkiye for its warmth, hospitality, and for becoming a second home during my academic journey. I owe my deepest appreciation to my mother and father, whose sacrifices, unconditional love, and unwavering belief in me have been the foundation of everything I have achieved—their strength, values, and constant support have carried me through every step of this path. I am also profoundly grateful to my family for their continued encouragement, and most especially to my wife, whose endless patience, love, and faith in me have been my greatest motivation—her presence has given me strength in difficult times and joy in moments of success. I want to express my thanks to all the friends and people I have met from around the world, whose kindness and companionship have enriched this journey in countless ways. Finally, I would like to express my sincere appreciation to all of my teachers from school to university in my home country, Syria, who laid the foundation for my academic path and whose dedication and efforts continue to inspire me.

CONTENTS

Page

ABSTRACT iv

ÖZET v

ACKNOWLEDGEMENTS xix

CONTENTS xx

LIST OF TABLES xxii

LIST OF FIGURES xxiii

SYMBOLS AND ABBREVIATIONS xxiv

1. INTRODUCTION 1

2. FUZZY LOGIC 13

2.1. Fuzzy Set and Fuzzy Number 15

2.2. Fuzzification and Defuzzification 23

2.2.1. Center of sums method 25

2.2.2. Maximum membership principle 27

2.2.3. Principle of the maximum of the maximums and the minimum of the maximums 28

2.2.4. Mean of maxima principle 29

2.3. Operations on Fuzzy Sets and Fuzzy Numbers 29

2.3.1. Arithmetic operations on fuzzy numbers using the extension principle 32

2.3.2. Arithmetic operations on fuzzy numbers using α-cut method 34

3. CLASSICAL REGRESSION AND FUZZY REGRESSION 39

3.1. Ordinary Least Squares 39

3.2. Multicollinearity 41

3.3. Ridge Regression 44

3.4. Liu Regression 52

Page

3.5. Fuzzy Regression 54

3.6. Fuzzy Least Squares 57

3.7. α-Cut Based Fuzzy Regression 61

3.8. Hybrid α-Cut-Based Estimation Framework 68

4. SIMULATION STUDY 71

5. EMPIRICAL STUDY 81

6. CONCLUSION AND RECOMMENDATIONS 91

REFERENCES 95

RESUME 104

LIST OF TABLES

Table **Page**

Table 2.1. Subtraction of two fuzzy numbers 33

Table 2.2. Multiplication of two fuzzy numbers..... 33

Table 2.3. Division of two fuzzy numbers..... 34

Table 3.1. Some of the proposed formulas for finding the optimal formulas for finding the optimal bias parameters for Ridge regression..... 48

Table 3.2. Some of the proposed formulas for finding the optimal bias parameters for Liu regression 54

Table 4.1. The performance measures when $\rho = 0.80$ and asymmetric data. 73

Table 4.2. The performance measures when $\rho = 0.80$ and symmetric data. 75

Table 4.3. The performance measures when $\rho = 0.90$ and asymmetric data. 76

Table 4.4. The performance measures when $\rho = 0.90$ and symmetric data. 77

Table 4.5. The performance measures when $\rho = 0.99$ and asymmetric data. 78

Table 4.6. The performance measures when $\rho = 0.99$ and symmetric data. 79

Table 5.1. Suggested candidates for α-cut sequences. 82

Table 5.2. Results of applying different methods to the 3 datasets. 83

Table 5.3. Relative errors of cores and spreads 89

LIST OF FIGURES

Figure **Page**

Figure 2.1. Normal and non-normal fuzzy sets..... 17

Figure 2.2. Convex and non-convex fuzzy sets. 19

Figure 2.3. Triangular fuzzy membership function. 22

Figure 2.4. Trapezoidal fuzzy membership function. 22

Figure 2.5. Gaussian fuzzy membership function. 23

Figure 2.6. Center of sums method of $\tilde{A}$ and $\tilde{B}$. 26

Figure 2.7. Maximum membership principle defuzzification method 27

Figure 2.8. Minimum of the maximums and maximum of the maximums method 28

Figure 2.9. The union of two fuzzy sets..... 30

Figure 2.10. The intersection of two fuzzy sets 30

Figure 2.11. The complement of a fuzzy set..... 31

Figure 3.1. Fuzzy output data 63

Figure 3.2. Intermediate estimators for the α-cut sequence (0, 0.25, 0.5, 0.75, 1)..... 64

Figure 3.3. Modified estimators for the α-cut sequence (0, 0.25, 0.5, 0.75, 1) 64

Figure 3.4. Estimated fuzzy coefficient and its membership function 65

Figure 3.5. α-cut-based fuzzy estimation algorithm..... 67

Figure 5.1. $RMSE_F$ values across α-cut sequences for various approaches 86

Figure 5.2. Expected and observed values for Tanaka's dataset 88

SYMBOLS AND ABBREVIATIONS

The symbols and abbreviations used in this study are presented below, along with their explanations.

Symbols	Explanations
$\widetilde{A}, \widetilde{B}, \ldots$	Fuzzy sets
$\mu_{\widetilde{A}}$	Membership function of fuzzy set $\tilde{A}$
σ	Standard deviation
β	Vector of coefficients in the regression model
λ	Eigenvalue
k	Bias parameter for Ridge regression
I	Identity matrix
d	Bias parameter for Liu regression
ρ	Correlation Coefficient
ρ^2	Coefficient of Determination

Abbreviations	Explanations
LASSO	Least Absolute Shrinkage and Selection Operator
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MSE	Mean Squared Error
$MSPE$	Mean Squared Prediction Error
OLS	Ordinary Least Squares
$RMSE_F$	Root Mean Square Error for Fuzzy Numbers
SST_E	Sum of Squared Errors
VIF	Variance Inflation Factor

1. INTRODUCTION

In today's data-driven world, the importance of information and statistics has increased significantly, influencing virtually every aspect of society. From ancient civilizations, which employed basic counting methods to monitor resources and populations, to modern societies utilizing vast datasets for advanced analytics, the evolution of statistical thought reflects humanity's continuous pursuit of knowledge and informed decision-making. Statistics not only serves as a fundamental tool for scientific inquiry and innovation but also plays a pivotal role in shaping public policies, guiding business strategies, and enhancing our understanding of complex phenomena. As we navigate an era characterized by unprecedented access to information, the ability to analyze and interpret statistical data has become essential for making well-informed decisions that affect social, economic, and environmental landscapes. This underscores the enduring relevance of statistics, highlighting its transformative impact in both historical and contemporary contexts.

Building upon the historical evolution of statistical thought and its critical role in decision-making, the development of logical systems also reflects a parallel journey toward more sophisticated frameworks for understanding complex phenomena. As statistics advanced, so did the exploration of logical structures, driven by the need to address the growing intricacies of human reasoning and the limitations of earlier systems. The exploration of logic throughout history reveals a rich intellectual evolution, adapting to the growing complexities of human reasoning. Classical logic, grounded in the work of ancient philosophers such as Aristotle, established a binary framework that dominated philosophical thought for centuries, paralleling the early use of statistics in ancient societies. However, as scholars delved deeper into the intricacies of human thought and natural language, the limitations of classical logic became increasingly evident. This recognition led to the development of alternative logical systems, including fuzzy logic, which emerged in the mid-20th century. Fuzzy logic challenges the rigid dichotomy of classical logic by introducing degrees of truth, thereby enabling more nuanced reasoning that better reflects real-world phenomena. In the broader context of the evolution of logical systems, the development of logical frameworks has progressed from classical foundations to more sophisticated models. Initially, classical logic relied on binary truth values, but advancements such as the exploration of three-valued logic by Lukasiewicz in 1930 [1] and later by Knuth in 1981 [2] marked significant strides

in this field. Simultaneously, German physicist Heisenberg's research on uncertainty, particularly his work between 1924 and 1927, culminated in the formulation of the Heisenberg Uncertainty Principle in 1927 [3], which provided foundational insights for multi-valued logic. Fuzzy logic, introduced by Zadeh in 1965 [4], extends the boundaries of classical logic by addressing uncertainty through graded truth values. Unlike the binary decision-making structure of classical logic, fuzzy logic allows propositions to exist on a continuum between "true" and "false," offering a more flexible and realistic model for reasoning.

Mathematically, fuzzy logic assigns values within the real number interval [0,1], enabling the examination of vague and imprecise concepts inherent in human thought and behavior [5]. While classical set theory defines membership in a set in a strict manner, fuzzy set theory emphasizes degrees of membership, allowing elements to belong to a set to varying degrees. This flexibility facilitates more nuanced interpretations and greater adaptability across diverse contexts, illustrating the continuous evolution of logic in response to the increasing complexity of reasoning and uncertainty. Together, these logical frameworks and statistical methods reflect humanity's enduring quest to understand and model the complexities of knowledge and decision-making, demonstrating a dynamic interplay between philosophy, logic, and mathematics.

Zadeh's foundational work in fuzzy logic, introduced in [4, 6-9], established key principles that underpin the field. In fuzzy logic, fuzzification is a fundamental process that transforms crisp input values into fuzzy values using membership functions. A membership function assigns a value between 0 and 1 to each element in the universal set, where 1 indicates full membership in the fuzzy set. This process tailors the membership values of each element to the specific characteristics of the dataset, enabling effective representation of uncertainty and imprecision. This process captures the uncertainty and imprecision inherent in real-world data, often using membership functions with shapes such as triangular and trapezoidal to define the degree of membership of input values. However, fuzzy numbers, due to their inherent nature, cannot be directly applied in practical situations since most existing tools are designed for standard numerical values. Consequently, transforming fuzzy numbers into conventional numerical forms is necessary, a process known as defuzzification. Defuzzification converts fuzzified outputs into a single crisp value derived from a fuzzy set, and it is a crucial step in fuzzy rule-based systems, such as fuzzy logic controllers. Although

these systems generate fuzzy outputs, defuzzification is required to produce precise values that can be interpreted and used in practical applications. Several methods are available for defuzzification, including the centroid method, mean of maximum, largest of maximum, and smallest of maximum, each offering a distinct approach to deriving crisp values from fuzzy outputs, ensuring that fuzzy systems can be effectively employed in decision-making and control tasks [4, 10, 11].

Returning to the discussion on statistical methods, they serve as fundamental tools in data analysis, allowing researchers to extract meaningful insights that play a crucial role across various disciplines, including the social sciences, natural sciences, and business. Among these methods, regression analysis is one of the most fundamental and widely employed techniques. Regression analysis is a statistical process used to estimate the relationships among variables, specifically by modeling the relationship between a dependent variable and one or more independent variables. This allows researchers to examine how changes in the predictors influence the outcome of interest. By quantifying these relationships, regression analysis aids in making predictions, testing hypotheses, and supporting informed decision-making processes. The versatility of regression analysis is evident in its various forms, each designed to address specific types of data and research questions. Common forms include linear regression, which models the relationship between variables as a straight line, multiple regression, which extends linear regression to include multiple predictors, and logistic regression, used for modeling binary outcomes. Further classifications, such as polynomial regression and non-linear regression, accommodate more complex relationships between variables.

Additionally, as the complexity of data continues to increase in an era of vast information, regression analysis adapts to incorporate advanced techniques, such as machine learning algorithms, to enhance its predictive power and broaden its applicability in real-world scenarios. This evolution highlights the fundamental importance of statistical methods, particularly regression analysis, in navigating the challenges of data interpretation, decision-making, and policy development in contemporary society. As data becomes increasingly multifaceted and dynamic, the continuous development of regression methods is essential for improving the accuracy and relevance of insights drawn from complex datasets.

Regression analysis is subject to several challenges that can compromise its efficacy. One common issue is the presence of outliers, which are data points that significantly deviate from the overall trend and can disproportionately influence the regression results, leading to biased estimates and skewed interpretations. Another critical concern is the assumption of linearity, which assumes that the relationship between independent and dependent variables is linear. Deviations from this assumption can result in poor model fit and inaccurate predictions. The independence of observations is also a key assumption, and violations—such as those found in clustered data or time series—can lead to unreliable estimates and affect the validity of statistical tests. Heteroscedasticity, or non-constant variance of the error terms across different levels of the independent variables, is another issue that violates the assumptions of ordinary least squares (OLS) regression, leading to inefficient estimates and misleading hypothesis tests. Additionally, autocorrelation, which is particularly problematic in time series data, occurs when residuals from one observation are correlated with those from another, resulting in underestimated standard errors and overly optimistic significance levels. Among these issues, multicollinearity is particularly troublesome, arising when two or more independent variables exhibit high correlations. This high correlation complicates the isolation of individual effects of each predictor on the dependent variable, inflates the standard errors of regression coefficients, and undermines the reliability of statistical inferences. These challenges highlight the importance of rigorous model assessment and validation in regression analysis to ensure robust, reliable, and accurate results.

Among these challenges, multicollinearity is particularly problematic, as it renders the OLS estimator unreliable and can lead to misleading conclusions. Multicollinearity arises when predictor variables are highly correlated, causing instability in the estimation of regression coefficients. Contributing factors include the inclusion of correlated predictors, data collection methods that result in redundant variables, and small sample sizes that do not capture sufficient variation within the data. While it may not always be possible to precisely identify the root cause of multicollinearity, understanding the underlying structure of this issue is crucial for researchers. A comprehensive understanding enables the development of strategies to mitigate the problem effectively. Common methods for detecting multicollinearity include the calculation of the variance inflation factor (VIF), which measures the extent to which the variance of a regression coefficient is inflated due to collinearity. A VIF value exceeding 10 typically signals a significant multicollinearity issue. Additionally, examining the correlation matrix of independent variables can help identify

high correlations that may indicate the presence of multicollinearity, guiding researchers in refining their model specifications.

To address multicollinearity, several strategies can be implemented to stabilize the regression estimates and improve the model's reliability. One common approach is to remove one of the highly correlated independent variables, thereby simplifying the model without sacrificing significant explanatory power. Alternatively, correlated independent variables can be combined using techniques such as principal component analysis, which transforms the correlated variables into a set of uncorrelated components.

Another widely used method is Ridge regression, introduced by Hoerl and Kennard [12] in 1970. This technique applies a penalty to the size of the coefficients, effectively regularizing the regression estimates and reducing the influence of multicollinearity. Liu regression, proposed by Liu [13] in 1993, also offers a similar solution by introducing a biasing parameter, which shrinks the regression coefficients towards zero, thus balancing bias and variance. Additionally, the least absolute shrinkage and selection operator (LASSO) regression, developed by Tibshirani [14] in 1996, uses a penalty to shrink some coefficients to zero, performing both variable selection and regularization to mitigate the effects of multicollinearity. Finally, researchers may also consider increasing the sample size to introduce more variability, which can help reduce the impact of multicollinearity. These approaches offer effective solutions to multicollinearity, ensuring the robustness and accuracy of regression models.

Fuzzy logic, much like classical logic, plays a significant role in enhancing regression analysis by addressing the inherent uncertainties and imprecisions present in real-world data. Traditional regression methods typically rely on crisp values and deterministic models, which may struggle to accurately capture the complexities of human reasoning and the vagueness inherent in natural phenomena. In contrast, fuzzy regression integrates fuzzy numbers and linguistic variables, allowing for a more nuanced representation of relationships between dependent and independent variables. In fuzzy regression, input values are expressed as fuzzy sets, offering a range of possible values rather than a single fixed number. This approach enables the model to account for the uncertainty and ambiguity present in the data. For example, instead of specifying an exact temperature, a fuzzy regression model might categorize it as "warm," encompassing a range of values that

represent varying degrees of warmth. Additionally, fuzzy regression models incorporate fuzzy rules, which further enhance their ability to model relationships in a way that is more intuitive and human understandable. These rules enable the model to learn from imprecise data, making fuzzy regression particularly valuable in fields such as economics, environmental science, and social sciences, where human behavior and perception often play a significant role in data interpretation. Ultimately, fuzzy regression strengthens the robustness of regression analysis by providing a framework that accommodates uncertainty and vagueness, leading to more reliable and interpretable results, particularly in domains where precision is difficult to achieve, and where capturing the subtleties of real-world phenomena is essential for informed decision-making.

Following Zadeh's seminal work on fuzzy set theory in 1965 [4], researchers began integrating qualitative variables into mathematical models using fuzzy membership degrees to address uncertainties inherent in real-world data. This innovation paved the way for the application of fuzzy logic in regression analysis. In 1982, Tanaka et al. [15] introduced a groundbreaking model that incorporated fuzzy set theory into linear regression by fuzzifying model coefficients while retaining crisp dependent and independent variables. This approach allowed for the representation of uncertainties by associating variables with fuzzy membership degrees and aimed to minimize the spread of predicted values, offering a novel method for handling imprecision in data interpretation. However, Tanaka's model exhibited limitations, particularly its tendency to produce overly crisp solutions that constrained model flexibility. To address these shortcomings, Celmins [16] developed a refined methodology that enabled the fitting of differentiable fuzzy model functions to crisp data. By employing conical membership functions for vector data and optimizing coefficients using a least-squares objective, Celmins introduced a more adaptable and efficient approach to fuzzy regression, accommodating both scalar and multidimensional data. These advancements have significantly shaped the evolution of fuzzy regression, establishing it as a robust framework for managing uncertainties and imprecisions in regression analysis while enhancing flexibility and interpretability [17].

In the late 1980s, Diamond [18] advanced the field with the fuzzy least squares method, which utilized triangular fuzzy numbers to enhance the estimation process, thereby reducing the dependency on exact values. This approach facilitated broader applicability in real-world scenarios where data precision is often variable. Subsequently, in 1992, Watada [19]

introduced an intersection-based method for fuzzy regression, illustrating its adaptability to dynamic data environments. That same year, Sakawa and Yano [20] proposed an iterative coefficient estimation algorithm, though they noted its sensitivity to outliers, an issue that continues to influence fuzzy regression models.

Building upon these foundational advancements, the early 1990s witnessed a surge of research aimed at enhancing the interpretability, robustness, and practical applicability of fuzzy regression models. Moskowitz and Kim [21] explored variations in coefficients' spreads and membership function shapes, enabling a deeper and more flexible interpretation of fuzzy relationships. Chang and Lee [22] introduced a least absolute deviation method, which significantly improved the robustness of fuzzy regression against outliers, while Peters [23] proposed interval regression as an alternative to Tanaka's approach, offering a distinct methodology for managing uncertainty. By 1996, Chang [24] advanced the field further with a generalized fuzzy weighted least squares regression method that separated the estimation of modal values and spreads. This innovation not only facilitated the effective treatment of outliers but also allowed for the incorporation of decision-makers' confidence in the data. Furthermore, this method was versatile, accommodating diverse types of fuzzy data, and enhanced computational efficiency, solidifying its impact on the development of robust and adaptable fuzzy regression models.

Advancements in fuzzy regression continued into the late 1990s, marked by the contributions of Redden and Woodall [25], who employed orthogonal least squares combined with linear programming to effectively mitigate the influence of outliers. During the same period, Kim et al. [26], followed by Kim and Chen [27], demonstrated the superiority of fuzzy regression methods over classical approaches, particularly in scenarios with limited data availability. Building on these developments, D'Urso and Gastaldi [28] proposed a doubly adaptive fuzzy regression model that independently estimated the centers and spreads of fuzzy observations. This innovation provided greater flexibility and depth in capturing and interpreting the complexities inherent in fuzzy data.

In the early 2000s, research on fuzzy regression saw the emergence of asymmetric and generalized methodologies aimed at enhancing the accuracy and flexibility of models in reflecting real-world data structures. Ishibuchi and Nii [29] introduced asymmetrical triangular and trapezoidal fuzzy coefficients, which allowed for more realistic

representations of data variability. Building on these advancements, Chang [30] developed a hybrid method combining fuzzy weighted arithmetic with fuzzy least squares, enabling the model to handle diverse data structures effectively. Around the same time, Lee and Chen [31] proposed a generalized nonlinear fuzzy regression model, opening new possibilities for addressing complex data relationships. Continuing this trajectory, Nasrabadi et al. [32] advanced the field with a multi-objective fuzzy regression model that enhanced predictive accuracy, further refining the nonlinear framework established by Lee and Chen [31].

Research in the mid-2000s emphasized confidence-based estimations, with Buckley [33] introducing a methodology that utilized confidence intervals to estimate fuzzy coefficients through α-cuts of fuzzy numbers. This approach brought statistical rigor to fuzzy regression analysis by enhancing the precision of triangular fuzzy number estimation. Concurrently, Kim et al. [34] advanced the fuzzy least absolute deviation approach by incorporating fuzzy inputs, further strengthening the robustness of fuzzy regression methods in handling outliers and improving the reliability of the results.

Subsequent research concentrated on refining the technical aspects of fuzzy regression models, focusing on enhancing both their theoretical foundations and practical applicability. Zeng and Zhang [35] explored distance-based estimations, employing various measures such as Euclidean, Y-K, and D_k distances to calculate coefficients, demonstrating consistency and reliability across different methods. In 2020, Chen and Nien [36] introduced an intuitionistic fuzzy regression model grounded in least absolute deviations, leveraging intuitionistic fuzzy numbers to integrate richer fuzzy information and improve decision-making accuracy. This innovative model offered greater flexibility in handling coefficient signs and symmetry, resulting in superior prediction outcomes compared to traditional fuzzy regression approaches.

Building on previous advancements, Belhadj [37] proposed a fuzzy regression model that incorporated Gaussian membership functions, replacing the traditional triangular forms. This novel approach was applied to data from Middle East and North Africa countries, where it successfully estimated the employment-poverty relationship. The use of Gaussian membership functions offered a more accurate representation of real-world conditions, where data often exhibit natural variability that cannot be captured by linear models alone. These developments mark a significant progression in fuzzy regression, shifting the

methodology from its early simple fuzzy parameterization to more sophisticated and flexible models capable of addressing a wider range of uncertainties and data complexities. As a result, fuzzy regression techniques have evolved to handle diverse forms of data irregularities and improve decision-making in complex environments.

As we mentioned, in classical regression models, multicollinearity poses significant challenges, and similar issues are encountered in fuzzy regression analysis, often with heightened complexity due to the distinct properties of fuzzy data. Fuzzy regression, like its classical counterpart, is affected by problems such as multicollinearity, outliers, non-linearity, and heteroscedasticity, all of which can substantially impair model performance. For instance, outliers can distort fuzzy membership functions, undermining the reliability of the model. Non-linearity presents an even greater challenge in fuzzy regression, as these models aim to capture relationships across ranges of values rather than fixed points. Heteroscedasticity, characterized by unequal variances across observations, adds further complexity, hindering the ability to establish consistent relationships within fuzzy datasets. Among these challenges, multicollinearity is particularly critical. In fuzzy regression, high correlations among independent variables not only destabilize coefficient estimates, as seen in classical regression, but also amplify the variance of fuzzy estimates, thereby reducing the model's interpretability. Addressing these issues is pivotal for advancing fuzzy regression methodologies capable of managing the complexities of real-world data while enhancing model robustness and interpretability.

The preceding discussion highlights the critical importance of addressing multicollinearity in fuzzy regression models. To tackle this issue, several studies have focused on adapting Ridge regression techniques to the fuzzy context, aiming to improve predictive stability under conditions of uncertainty. These extensions of Ridge regression to fuzzy datasets have played a significant role in enhancing the robustness and interpretability of fuzzy regression models. The most notable contributions in this area are summarized as follows:

Hong et al. [38] extended Ridge regression to the fuzzy domain by incorporating crisp input variables and Gaussian fuzzy outputs. Their work built on Xu's earlier models [39, 40] and integrated distance measures from Diamond [18] as well as Yang and Ko [41]. This method demonstrated improved effectiveness in modeling fuzzy data relationships, as evidenced by enhanced goodness-of-fit metrics. Building on this foundation, Hong and Hwang [42]

proposed a novel Ridge regression method tailored for estimating multivariable fuzzy linear and nonlinear regression models. Their approach utilized triangular fuzzy numbers within a dual Ridge regression framework, enabling the representation of fuzzy nonlinear relationships by constructing high-dimensional feature spaces for datasets with crisp inputs and fuzzy outputs. Experimental results further highlighted the performance improvements achieved through this method.

Building on advancements in fuzzy regression, Donoso et al. [43] in 2007 tackled specific limitations of Tanaka's linear programming model by incorporating asymmetric triangular membership functions and quadratic programming. Their approach, applied to house price data, utilized non-possibilistic methods with positive coefficients, enhancing both model interpretability and the scope of applicable data types. Similarly, Bas et al. [44] introduced a novel method by proposing a type-1 fuzzy function, which is characterized by fuzzy sets with crisp membership values ranging between 0 and 1. Unlike type-2 fuzzy sets, which extend this concept by incorporating a fuzzy membership function that allows for uncertainty in membership values, type-1 fuzzy functions are simpler and computationally less intensive. Bas et al. [44] employed Ridge regression instead of traditional multiple linear regression within this framework. Their approach, tested on various time series datasets, demonstrated improved forecasting accuracy, particularly in addressing the complexities of real-world data.

Building upon the foundational work in fuzzy Ridge regression, Rabiei et al. [45] made further advancements by introducing models in which both the independent and dependent variables are fuzzy. To address multicollinearity, they applied a generalized VIF and developed a new objective function to stabilize predictions. Their model evaluated using mean squared prediction error ($MSPE$) and fuzzy distance measures, demonstrated improved handling of multicollinearity in various simulations and cross-validation tests. Extending this work, Karbasi et al. [46] incorporated recurrent neural networks to estimate coefficients in Ridge fuzzy regression models. By validating the method's accuracy through $MSPE$ and Monte Carlo simulations, they showed that these Ridge regression adaptations remain effective, even when variables are linear combinations. Similarly, Baş [47] applied a type-1 fuzzy regression model based on Ridge regression to forecast monthly house sales to foreigners, marking a novel application of the method. Evaluations revealed that this model

outperformed several others in the literature, further emphasizing the utility of Ridge regression-based fuzzy models for accurate forecasting across diverse datasets.

A substantial body of research in fuzzy regression focuses on estimating the bounds of model coefficients by jointly analyzing the upper and lower limits of independent and dependent variables. In this context, Choi et al. [48] and Kim and Jung [49] introduced an algorithm grounded in a fundamental concept of fuzzy set theory known as the α-cut, or α-level. The α-cut divides or refines the range of crisp numbers based on an α value, where $\alpha \in [0,1]$, reflecting membership values. By setting α to 0, the entire range of values is encompassed, enabling coefficient estimation from the outer bounds of the data. Adjusting α between 0 and 1 provides flexibility to exclude low membership values from the estimation or to employ multiple α cuts concurrently, allowing for more nuanced coefficient estimates drawn from specific segments of the fuzzy independent or dependent variables, or both. This approach yields results comparable to those from fuzzy OLS or fuzzy Ridge regression through different sets of α-cuts, which allow researchers to generate symmetric or asymmetric results, depending on the selected error metrics. By comparing fuzzy OLS and fuzzy Ridge methods with these α-cut configurations, Ridge k values for Ridge regression were determined using K-fold cross-validation or by defining a range close to 0 and adjusting incrementally. Performance was then evaluated using error metrics. Both the dataset from Tanaka et al. [15] and simulated datasets with varying correlation structures were utilized for testing, where the proposed algorithm consistently outperformed alternative approaches in all scenarios when applying the α-cuts-based fuzzy Ridge method, or simply the fuzzy Ridge method. This work, centered on the α-cuts-based fuzzy Ridge method, is the focus of the current study, which builds upon and extends these advancements to further improve the handling of fuzzy data in regression analysis.

Our study aims to evaluate an extensive set of 77 formulas for determining the k parameter in fuzzy Ridge regression within the framework of fuzzy data. Specifically, the evaluation employs the α-cut-based estimation method, which has not been comprehensively assessed in this field. The investigation considers a wide range of scenarios to ensure robustness and thorough analysis. Building on Hoerl and Kennard's [12] foundational framework and the recent work of Kocasoy [50] in crisp data analysis, this research uniquely applies these formulas to fuzzy data, advancing the estimation process in the field of fuzzy regression.

Additionally, this study pioneers the application of Liu regression with α-cut estimation methods, marking the first systematic comparison of various optimization approaches for determining the d parameter in Liu regression within a fuzzy context. Extensive simulations were conducted across numerous scenarios, including variations in the number of dependent variables, sample size, multicollinearity, and triangular fuzzy numbers, allowing for a robust assessment of both methods. Furthermore, the study validated its findings using well-established real fuzzy datasets, providing a practical comparison between fuzzy Ridge and fuzzy Liu regression approaches. The results demonstrate the efficiency and precision of these formulas in identifying optimal bias parameters, offering a significant advancement over traditional computational methods, such as K-fold cross-validation. This work contributes a systematic, data-driven approach to enhance fuzzy regression analysis, marking a substantial step forward in the field and offering new avenues for future research and application.

The remainder of this thesis is organized as follows:

Section 2 introduces the foundational concepts of fuzzy logic, including fuzzy sets and fuzzy numbers, along with their operations, supported by examples. It also explores various methods of fuzzification and defuzzification. Section 3 provides an introduction to regression analysis, discussing OLS regression in both classical and fuzzy logic contexts. It examines multicollinearity, methods for its detection, and approaches to address it, followed by an overview of Ridge and Liu estimation methods within classical logic. The section also details the primary estimation algorithms, particularly the α-cut-based fuzzy estimation method, highlighting the unique contributions of this study. Section 4 presents a comprehensive simulation study conducted across numerous scenarios, while Section 5 applies the proposed methods to real-world datasets. Finally, Section 6 summarizes the results and includes a discussion of the findings.

2. FUZZY LOGIC

In daily life, many terms are often used interchangeably, and the responses to various situations frequently contain a considerable degree of vagueness. Whether offering commentary, providing critique, or making decisions, individuals commonly encounter mental ambiguity, a characteristic inherent in human cognition. Humans, by nature, are beings who engage in debate, analysis, and conclusion-drawing; they are creatures driven by ideas. People express their thoughts and perceptions of the world around them through both verbal and numerical terms. Examples of such expressions include linguistic terms such as "partly cloudy skies," "lukewarm tea," "a tall child," or "a dash of sugar." Human decision-making processes are largely influenced by terms that lack absolute certainty, and this inherent ambiguity in human reasoning has led to the development and refinement of various logical systems. Over time, logic has evolved from classical binary systems to more sophisticated structures, such as multivalued and fuzzy logic, which offer a more accurate representation of the uncertainty and imprecision that naturally characterize human thought and decision-making processes [51].

Building on the inherent ambiguity of human reasoning and the need to better represent imprecise concepts, the evolution of scientific and logical systems has transitioned classical logic from 3-valued to multi-valued frameworks, ultimately culminating in infinite-valued logic. This progression gave rise to fuzzy logic, introduced by Zadeh in 1965 [4], which addresses uncertainty by extending beyond the binary decision-making structure of classical logic to incorporate graded degrees of truth and falsehood. In fuzzy logic, propositions are no longer restricted to the rigid "true" or "false" dichotomy but can instead take on values within an infinite range. Mathematically, this framework assigns propositions values from the real number interval [0,1], providing a robust system for analyzing and mathematically representing the vague concepts and subjective nuances inherent in human thought and behavior [5].

By allowing for degrees of membership rather than binary states, fuzzy logic reflects human experiences and linguistic imprecision, offering a more flexible approach to decision-making and analysis. Zadeh's innovation enables computations based on qualitative information, such as the varying tastes represented by the term "ice cream," which cannot be easily

quantified by traditional binary logic [9]. This paradigm shift not only overcomes the limitations of classical logic but also paves the way for diverse applications across mathematical science, engineering, and beyond. Although initially met with skepticism and viewed as a philosophical debate due to its limited early applications, fuzzy logic has since gained widespread acceptance. Notably, its practical impact was demonstrated when Mamdani and Assilian [52] applied it to steam engine control systems in 1975. Today, fuzzy set theory continues to thrive, influencing a wide array of fields, including engineering, business, and economics [53].

At first glance, the concept of fuzzy sets may appear similar to probability theory, as both deal with uncertainty. However, they are fundamentally distinct tools grounded in different logics. Probability measures the likelihood of an event's occurrence, answering questions such as "How often does it happen?" In contrast, fuzzy sets measure the degree of certainty, addressing questions like "How confident are you that it happens?" [54]. Zadeh [8] explored these distinctions and similarities, emphasizing that while probability theory and fuzzy logic may intersect in certain aspects, their foundations and objectives differ significantly. The relationship between fuzzy sets and random sets does not equate fuzziness with randomness, nor does it suggest that fuzzy logic and probability theory share the same theoretical framework. For instance, probability theory cannot adequately describe inherently fuzzy events such as "tomorrow will be a warm day" or "there will be a strong earthquake in the near future." It also lacks mechanisms to handle fuzzy quantifiers like "many," "most," "several," or "few" and does not provide a framework for computing fuzzy probabilities expressed in terms such as "likely," "unlikely," or "not very likely." While subjective probability theory attempts to address elicitation, it falls short in estimating fuzzy probabilities, leaving questions like "What is the probability that my car may be stolen?" inadequately resolved. In many cases, subjective probabilities are rooted in fuzzy perceptions of frequency-based probabilities, further complicating assessments. Moreover, probability theory's limitations as a language for representing meaning make it difficult to express statements like "It is not likely that there will be a sharp increase in the price of oil in the near future." These constraints highlight that classical probability theory often struggles with real-world problems where fuzziness is central rather than peripheral, whereas fuzzy logic serves as a complementary framework that effectively addresses imprecision and vagueness. The complementarity lies in their ability to address different types of uncertainty: probability theory excels in handling stochastic or randomness-based uncertainty, while

fuzzy logic is better suited for modeling ambiguity and vagueness, offering a more comprehensive understanding of complex systems.

Further emphasizing this distinction, Hajek et al. [55] highlight the fundamental differences in how fuzzy logic and probability theory treat truth values. In truth-functional contexts, probabilities do not operate in a truth-functional manner, adhering instead to a binary framework of true or false, though this limitation does not apply in non-truth-functional contexts. For instance, Gerla [56] presents an abstract, non-truth-functional fuzzy logical system where interpretations encompass all probabilities concerning the set of all formulas, supported by a complete deductive system. The primary divergence lies in their approach to truth values: fuzzy logic allows for partial degrees of truth, enabling a continuum of truth values, whereas probability theory operates on crisp notions, assigning probabilities to propositions as a degree of belief in their truthfulness.

Formally, fuzzy logic is a many-valued logic, whereas probability theory aligns more closely with a two-valued modal logic, underscoring the unique strengths of both frameworks and highlighting their complementary roles in addressing diverse forms of uncertainty and reasoning.

2.1. Fuzzy Set and Fuzzy Number

Consider X as a space of objects and x as an arbitrary element of X. In classical set theory, an element either belongs to or does not belong to a set $A \subseteq X$. The value of 1 can be assigned to x as the pair $(x, 1)$ if $x \in A$, indicating that x belongs to set A. Similarly, the pair $(x, 0)$ represents the case where $x \notin A$, meaning x is not an element of set A .

Classical logic does not accommodate any situations beyond strict membership or non-membership. Zadeh [4] introduced an extension of classical logic known as fuzzy logic. In this framework, rather than limiting each element to the binary states of either belonging or not belonging to a set, fuzzy logic allows for a spectrum of membership. An element x may fully belong, not belong, or possess a degree of membership, represented by a value between 0 and 1. A value of 0 indicates non-membership, 1 indicates full membership, and any value between 0 and 1 reflects partial membership. This degree of membership is quantified by a

function known as the membership function, which can take either discrete or continuous values [57].

2.1.1 Definition: A fuzzy set $\tilde{A}$ in universal set X is defined as a set of ordered pairs $\tilde{A} = \{(x, \mu_{\tilde{A}}(x) | x \in X)\}$, where $\mu_{\tilde{A}}(x)$ is the membership function of the fuzzy set $\tilde{A}$. This function assigns a membership value to each element $x \in X$, with values ranging from 0 to 1. When the membership function assigns values of either 0 or 1 to all elements, the fuzzy set reduces to a classical set, adhering to the principles of classical logic. For classical logic, the membership degree is represented by the function μ, as shown in Eq. 2.1.

$$\mu_{\tilde{A}}(x) = \begin{cases} 1 & ; \quad x \in A \\ 0 & ; \quad x \notin A \end{cases} \tag{2.1}$$

Zadeh defined the representation of fuzzy sets with both finite and infinite elements using Eq. 2.2 and Eq. 2.3:

$$\tilde{A} = \frac{\mu_{\tilde{A}}(x_1)}{x_1} + \frac{\mu_{\tilde{A}}(x_2)}{x_2} + \cdots + \frac{\mu_{\tilde{A}}(x_n)}{x_n} = \sum_{i=1}^{n} \frac{\mu_{\tilde{A}}(x_i)}{x_i} \tag{2.2}$$

$$\tilde{A} = \int_X \frac{\mu_{\tilde{A}}(x)}{x} \tag{2.3}$$

In this context, the symbols Σ and $\int$ are not used in the conventional sense of summation and integration, but rather to indicate the aggregation of membership functions. Additionally, the division symbol / in the equations does not carry its usual mathematical meaning. Instead, the denominator represents the set element, while the numerator indicates the degree of membership associated with that element in the set [58].

An example of a fuzzy set can be seen in the context of selecting a city for an international film festival. Let $X = \{New\ York,\ Berlin,\ Sydney\}$ represent the cities under consideration. The fuzzy set $\tilde{A}$, which reflects the suitability of a city for hosting the festival, can be defined as $\tilde{A} = \{(New\ York, 0.85), (Berlin, 0.7), (Sydney, 0.6)\}$. In this scenario, the membership grades indicate each city's perceived suitability, with New York rated the highest at 0.85, followed by Berlin at 0.7, and Sydney at 0.6. These membership values are inherently subjective; different organizers might assign different scores based on their experiences or the cities' cultural appeal. For instance, an organizer with a preference for

European cinema might rate Berlin higher than New York. This example highlights how fuzzy sets effectively encapsulate diverse evaluations, thereby aiding in the decision-making process for selecting the most appropriate city for the film festival.

As previously mentioned, a fuzzy set is fully defined by its membership function. The following section will outline the key definitions related to fuzzy sets.

Given that X is the universal set, the following definitions are provided [4, 6, 57, 59-63]:

2.1.2 Definition: The support of a fuzzy set $\tilde{A}$ is defined as the set of all points $x \in X$ such that $\mu_{\tilde{A}}(x) > 0$, and is denoted by $S(\tilde{A}) = \{x|\mu_{\tilde{A}}(x) > 0\}$.

2.1.3 Definition: The height of a fuzzy set $\tilde{A}$ is defined as the maximum membership value achieved by any element in the set and is defined as $h(\tilde{A}) = \sup_{x \in X} \mu_{\tilde{A}}(x)$.

2.1.4 Definition: The core of a fuzzy set $\tilde{A}$ is defined as the set of all points $x \in X$ such that $\mu_{\tilde{A}}(x) = 1$, and is denoted by $C(\tilde{A}) = \{x|\mu_{\tilde{A}}(x) = 1\}$, provided that the height of $\tilde{A}$ is 1.

2.1.5 Definition: A fuzzy set $\tilde{A}$ is considered normal when its core set is not empty, meaning there is at least one element with a membership function value of 1, as illustrated in Figure 2.1.

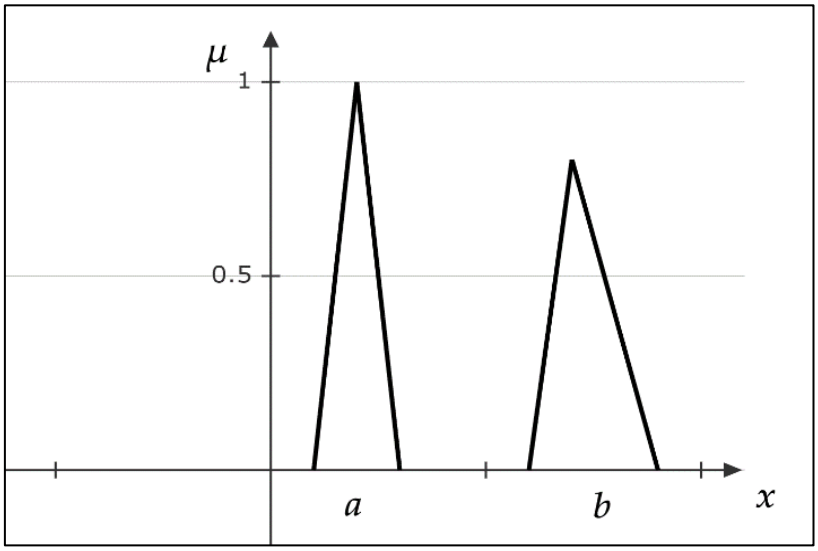


Figure 2.1. Normal (a) and non-normal (b) fuzzy sets.

2.1.6 Definition: The boundary of a fuzzy set $\tilde{A}$ is defined as the set of all points $x \in X$ such that $0 < \mu_{\tilde{A}}(x) < 1$, and it is denoted by $B(\tilde{A}) = \{x|0 < \mu_{\tilde{A}}(x) < 1\}$.

2.1.7 Definition: A crossover point of a fuzzy set $\tilde{A}$ is a point $x \in X$ such that $\mu_{\tilde{A}}(x) = 0.5$, and is denoted by $CO(\tilde{A}) = \{x|\mu_{\tilde{A}}(x) = 0.5\}$.

2.1.8 Definition: A fuzzy singleton is a fuzzy set whose support is a single point.

2.1.9 Definition: A nonnegative fuzzy set is a fuzzy set whose membership function are $\mu_{\tilde{A}}(x) = 0$ for all $x < 0$.

2.1.10 Definition: The α-cut or α-level set of a fuzzy set $\tilde{A}$ is a crisp set defined by $\tilde{A}_\alpha = \{x|\mu_{\tilde{A}}(x) \geq \alpha\}$, where α can obviously take a value from the range $[0,1]$. Similarly, the strong α-cut or strong α-level set is defined as $\tilde{A}'_\alpha = \{x|\mu_{\tilde{A}}(x) > \alpha\}$.

Using the notation for α-cut set, the support and core of a fuzzy set $\tilde{A}$ can be expressed as $S(\tilde{A}) = \tilde{A}'_0$ and $C(\tilde{A}) = \tilde{A}_1$.

Let us explain the definitions presented up to this point through an example.

Example: Given the universal set $X = \{1,2,\dots,12\}$ and the fuzzy set $\tilde{A} = \{(3,0.4), (5,0.3), (7,0.8), (9,0.7), (11,1)\}$:

The support set is: $S(\tilde{A}) = \{x|\mu_{\tilde{A}}(x) > 0\} = \{3,5,7,9,11\}$. Clearly, since the support set consists of more than one element, $\tilde{A}$ is not a singleton set.

The height of the set is: $h(\tilde{A}) = \sup_{x \in X} \mu_{\tilde{A}}(x) = 1$.

The core set is: $C(\tilde{A}) = \{x|\mu_{\tilde{A}}(x) = 1\} = \{11\}$. Clearly, since the core set is not empty and the height is 1, $\tilde{A}$ is a normal fuzzy set.

The boundary set is: $B(\tilde{A}) = \{x|0 < \mu_{\tilde{A}}(x) < 1\} = \{3,5,7,9\}$.

The set of crossover points is: $CO(\tilde{A}) = \{x|\mu_{\tilde{A}}(x) = 0.5\} = \phi$.

The α-cuts for $\alpha \in \{0.3,0.4,0.7,0.8,1\}$ are: $\tilde{A}_{0.3} = \{3,5,7,9,11\}$, $\tilde{A}_{0.4} = \{3,7,9,11\}$, $\tilde{A}_{0.7} = \{7,9,11\}$, $\tilde{A}_{0.8} = \{7,11\}$, $\tilde{A}_1 = \{11\}$.

The strong α-cuts for $\alpha \in \{0.3,0.4,0.7,0.8,1\}$ are: $\tilde{A}'_{0.3} = \{3,7,9,11\}$, $\tilde{A}'_{0.4} = \{7,9,11\}$, $\tilde{A}'_{0.7} = \{7,11\}$, $\tilde{A}'_{0.8} = \{11\}$, $\tilde{A}'_1 = \phi$.

2.1.11 Definition: A fuzzy set can either be convex or non-convex. Convexity is a key attribute of fuzzy membership functions, as it ensures smooth, gradual transitions between membership degrees, preserving logical consistency and simplifying optimization and decision-making in fuzzy systems. It also makes the representation of partial membership more intuitive and easier to interpret. A fuzzy set $\tilde{A}$ is convex if $\mu_{\tilde{A}}(\gamma x_1 + (1-\gamma)x_2) \geq \min\{\mu_{\tilde{A}}(x_1), \mu_{\tilde{A}}(x_2)\}, x_1, x_2 \in X, \gamma \in [0,1]$. Figure 2.2 giving an example of convex and non-convex sets [4].

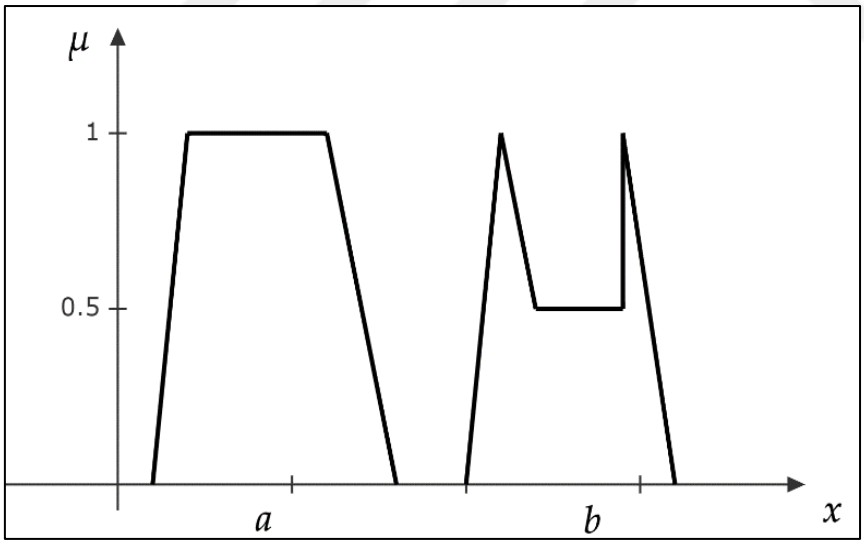


Figure 2.2. Convex (a) and non-convex (b) fuzzy sets.

Before discussing the membership function shapes, it is important to provide a summary of the extension principle, a key concept introduced by Zadeh [6], also known as Zadeh's extension principle [53, 62].

Let X be the Cartesian product of universes, denoted as $X = X_1 \times ... \times X_r$, and let $\tilde{A}_1, ..., \tilde{A}_r$ represent r fuzzy sets in $X_1, ..., X_r$, respectively. The Cartesian product of $\tilde{A}_1, ..., \tilde{A}_r$ is defined in Eq. 2.4:

$$\tilde{A}_1 \times ... \times \tilde{A}_r = \int_{X_1 \times ... \times X_r} \frac{min\left(\mu_{\tilde{A}_1}(x_1), ..., \mu_{\tilde{A}_r}(x_r)\right)}{f(x_1, ..., x_r)} \tag{2.4}$$

Let f be a mapping from $X_1 \times ... \times X_r$ to a universe Y, such that $y = f(x_1, ..., x_r)$. According to Zadeh's extension principle [6], a fuzzy set $\tilde{B}$ on Y can be induced from the fuzzy sets $\tilde{A}_1, ..., \tilde{A}_r$ via f as shown in Eq. 2.5.

$$\mu_{\tilde{B}}(y) = \sup_{x_1 \times ... \times x_r} min\left(\mu_{\tilde{A}_1}(x_1), ..., \mu_{\tilde{A}_r}(x_r)\right) \quad ; \quad y = f(x_1, ..., x_r) \tag{2.5}$$

Additionally, $\mu_{\tilde{B}}(y) = 0$ if $f^{-1}(y) = \phi$, where $f^{-1}(y)$ denotes the preimage of y. The membership value $\mu_{\tilde{B}}(y)$ is the maximum of the membership values $\mu_{\tilde{A}_1 \times ... \times \tilde{A}_r}(x_1, ..., x_r)$ corresponding to all realizations of y using the r-tuples $(x_1, ..., x_r)$.

For the special case when $r = 1$, this was addressed by Zadeh [4]. If f is one-to-one, Eq. 2.5 simplifies to $\mu_{\tilde{B}}(y) = \mu_{\tilde{A}}\left(f^{-1}(y)\right)$ whenever $f^{-1}(y) \neq \phi$.

Example: Let $X = \{-9, -8, ..., 8, 9\}$ be the universal set and let $\tilde{A} = \{(-2, 0.4), (-1, 0.8), (0, 1), (1, 0.3), (2, 0.1)\}$ be a fuzzy set. Using the extension principle for the function $f(x^2 + 1)$:

$$\tilde{B} = \{f(-2, 0.4), f(-1, 0.8), f(0, 1), f(1, 0.3), f(2, 0.1)\}$$
$$= \{(5, 0.4), (2, 0.8), (1, 1), (2, 0.3), (5, 0.1)\}$$

Since $\tilde{B}$ has elements with multiple membership values, it becomes:

$$\tilde{B} = \{(1, 1), (2, \max(0.3, 0.8)), (5, \max(0.1, 0.4))\} = \{(1, 1), (2, 0.8), (5, 0.4)\}$$

2.1.12 Definition: A fuzzy number is represented by a set of real numbers and an associated membership function. A fuzzy number is a type of fuzzy set that represents a quantity with a degree of uncertainty. It is characterized by a membership function that defines the degree of membership of each element in the set. A fuzzy number is often used to model imprecise or uncertain information in various applications, such as decision-making, control systems, and economic forecasting. A fuzzy number can be formally defined as follows:

Let $\tilde{A}$ be a fuzzy number on the real line $\mathbb{R}$. Then, $\tilde{A}$ is defined as a fuzzy set characterized by a membership function $\mu \rightarrow [0,1]$, where [4]:

- $\mu_{\tilde{A}}(x) = 1$ indicates full membership,
- $\mu_{\tilde{A}}(x) = 0$ indicates non-membership, and
- $0 < \mu_{\tilde{A}}(x) < 1$ indicates partial membership.

Another definition of fuzzy numbers is that for a fuzzy set to qualify as a fuzzy number, it must satisfy the following criteria [64]:

- Normality: The fuzzy set should be a normal fuzzy set, meaning there is at least one element with a membership degree of 1.
- Convexity: The fuzzy set must be convex, indicating that if an element x has a degree of membership greater than 0, then all elements between the left and right bounds of x must also have degrees of membership bigger than 0.
- Limited support: The support set of the fuzzy set must be bounded, ensuring that the fuzzy number is defined over a finite interval.
- Closed range for α-set: Each α-set of the fuzzy set must fall within a closed range on the real number line, indicating that all elements with membership grades equal to or above a certain threshold α are contained in a closed interval.

The shape of a membership function varies depending on the context and the nature of the data. Commonly used membership functions in the literature include triangular, trapezoidal, Gaussian, and sigmoidal functions [53, 54, 65].

The triangular membership function is widely regarded as the most used shape due to its simplicity. It is defined by three parameters, a, b and c, where a and c represent the boundary values, while b denotes the core of the fuzzy set. As shown in Figure 2.3, the triangular membership function is mathematically expressed in Eq. 2.6:

$$\mu_{\tilde{A}}(x; a, b, c) = \begin{cases} \frac{x-a}{b-a} & ; \quad a \leq x \leq b \\ \frac{c-x}{c-b} & ; \quad b \leq x \leq c \\ 0 & ; \quad otherwise \end{cases} = max\left[min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right] \tag{2.6}$$

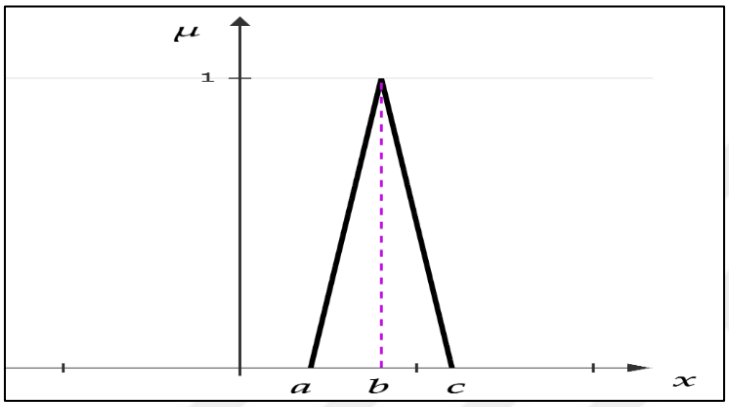


Figure 2.3. Triangular fuzzy membership function.

The second type is the trapezoidal membership function as illustrated in Eq. 2.7, defined by four parameters a, b, c and d. This is illustrated in Figure 2.4.

$$\mu_{\tilde{A}}(x; a, b, c, d) = \begin{cases} \frac{x-a}{b-a} & ; \quad a \leq x \leq b \\ 1 & ; \quad b \leq x \leq c \\ \frac{d-x}{d-c} & ; \quad c \leq x \leq d \\ 0 & ; \quad otherwise \end{cases} = max\left[min\left(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}\right), 0\right] \tag{2.7}$$

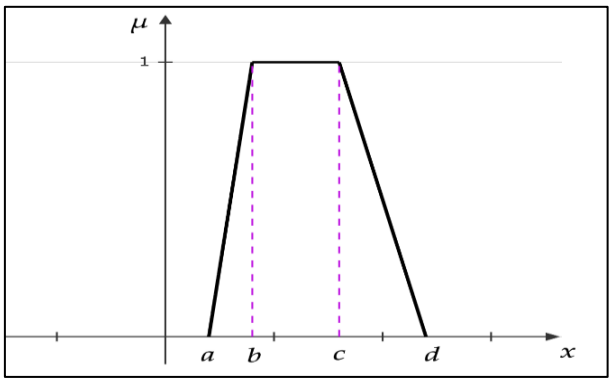


Figure 2.4. Trapezoidal fuzzy membership function.

As observed in Figure 2.4, when the values of b and c are equal, the trapezoidal membership function simplifies to a triangular fuzzy set.

In addition to the triangular and trapezoidal membership functions, the Gaussian membership function—expressed in Eq. 2.8—is another commonly used shape, ideal for capturing uncertainty and imprecision. It is defined by two parameters: c, representing the center, and σ, specifying the width of the fuzzy set $\tilde{A}$ as we can see in Figure 2.5 [59].

$$\mu_{\tilde{A}}(x; c, \sigma) = e^{-\frac{(c-x)^2}{2\sigma^2}} \tag{2.8}$$

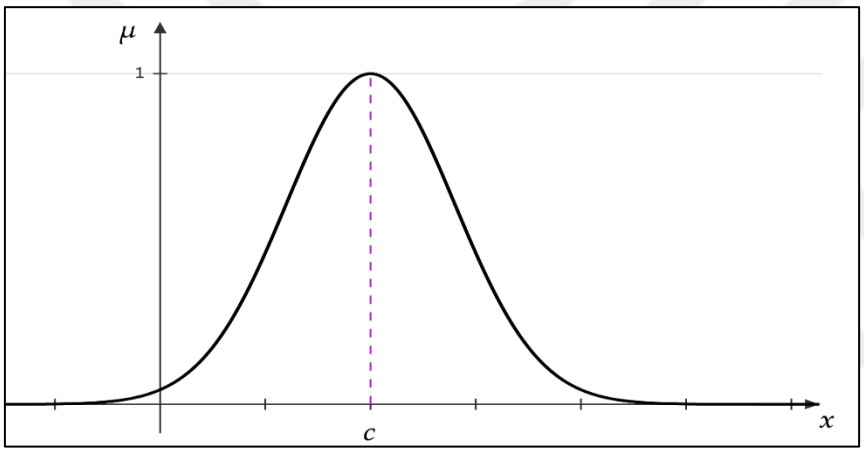


Figure 2.5. Gaussian fuzzy membership function.

The center parameter c can take any real number, depending on where the peak of the Gaussian curve is positioned. The standard deviation σ, on the other hand, must be a positive real number, as it controls the width of the curve.

2.2. Fuzzification and Defuzzification

Fuzzification refers to the process of transforming crisp values into fuzzy values through the application of membership functions specifically designed for the characteristics of a given dataset. This transformation is essential for effectively capturing the uncertainty and imprecision inherent in real-world data. The selection of an appropriate membership function is a critical step, as it profoundly influences the accuracy and performance of fuzzy systems [4, 10].

For instance, a fuzzy control system regulating room temperature can classify the temperature into fuzzy sets such as “Cold,” “Comfortable,” and “Hot,” utilizing triangular membership functions. These fuzzy sets are defined by specific parameters; for example, the “Cold” set is represented as a triangular fuzzy set with a lower limit of 0 degree, a peak value of 10 degree, and an upper limit of 20 degree. This representation can be expressed as $\tilde{A} = (0,10,20)_T$ where T stands for triangular. The corresponding membership function for this fuzzy number is defined by Eq. 2.9.

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{x}{10} & ; \quad 0 \le x \le 10 \\ \frac{20-x}{10} & ; \quad 10 \le x \le 20 \\ 0 & ; \quad otherwise \end{cases} \tag{2.9}$$

Similarly, the fuzzy set for “Comfortable” can be represented as $\tilde{B} = (15,20,25)_T$, while the fuzzy set for “Hot” can be defined as $\tilde{C} = (20,30,35)_T$. These fuzzy sets enable the classification of crisp temperature values into corresponding categories, facilitating the representation of temperature within the fuzzy control system.

For illustration, consider the crisp temperature values 17, 22, and 25 degrees. The membership degrees for these temperatures in the three fuzzy sets (“Cold,” “Comfortable,” and “Hot”) are computed as follows:

- Cold:
  - $\mu_{\tilde{A}}(17) = 0.3$
  - $\mu_{\tilde{A}}(22) = 0$
  - $\mu_{\tilde{A}}(25) = 0$
- Comfortable:
  - $\mu_{\tilde{A}}(17) = 0.4$
  - $\mu_{\tilde{A}}(22) = 0.6$
  - $\mu_{\tilde{A}}(25) = 0$
- Hot:
  - $\mu_{\tilde{A}}(17) = 0$
  - $\mu_{\tilde{A}}(22) = 0.2$
  - $\mu_{\tilde{A}}(25) = 0.5$

It can be observed that a single crisp value can simultaneously belong to multiple fuzzy sets, each with varying degrees of membership. This flexibility enables a more nuanced and comprehensive representation of data, in contrast to classical logic, which rigidly classifies elements as either belonging to a set or not.

Fuzzy numbers are often challenging to apply directly in real-world scenarios due to the limitations of current technology, as most tools and devices are designed to operate with standard numerical values. Consequently, it becomes necessary to transform fuzzy numbers into conventional numerical forms to ensure compatibility with existing systems. This transformation process, known as defuzzification, involves converting fuzzy numbers into crisp values and plays a critical role, as the resulting crisp value serves as the final output integrated into the system. For instance, a fuzzy number representing temperature, such as a range between 20 degrees and 30 degrees with varying degrees of membership, can be converted into a single crisp value, such as 25 degrees, using a defuzzification method. The choice of defuzzification method significantly influences the output, as different techniques may yield different crisp values from the same fuzzy number, underscoring the importance of selecting an appropriate method. The defuzzification methods presented in the literature [51, 66-68], along with the specific technique employed in this study, will be thoroughly examined in the following section, as detailed in [11].

2.2.1. Center of sums method

The centroid method, also known as the center of gravity or center of area method, is the most commonly utilized defuzzification technique. This approach determines the defuzzified value by calculating the center of the area under the aggregated fuzzy output distribution, taking into account the contributions of all relevant fuzzy sets. Notably, overlapping regions between fuzzy sets are accounted for multiple times, ensuring that each set's influence is proportionately represented in the calculation. The defuzzified value, x_D (D stands for defuzzified), is mathematically expressed as shown in Eq. 2.10.

$$x_D = \frac{\sum_{i=1}^{N} x_i \sum_{j=1}^{n} \mu_{\tilde{A}_j}(x_i)}{\sum_{i=1}^{N} \sum_{j=1}^{n} \mu_{\tilde{A}_j}(x_i)} \tag{2.10}$$

where n represents the number of fuzzy sets, N denotes the total number of discrete points for the fuzzy variable, and $\mu_{\tilde{A}_j}(x_i)$ is the membership function value for the j-th fuzzy set at the i-th data point x_i. The numerator represents the weighted sum of these values, while the denominator aggregates the membership values, providing a balanced crisp output. This method is widely preferred due to its ability to offer a smooth and logical interpretation of the fuzzy set's geometric properties. Alternatively, the defuzzified value can also be determined using Eq. 2.11, as follows:

$$x_D = \frac{\sum_{i=1}^{t} A_i \bar{x}_i}{\sum_{i=1}^{t} A_i} \tag{2.11}$$

Here, A_i represents the firing area of the i-th rule, t is the total number of rules fired, and $\bar{x}_i$ represents the center of area.

Example: In Figure 2.6 below, the aggregated fuzzy set of two fuzzy sets $\tilde{A}$ and $\tilde{B}$ is illustrated.

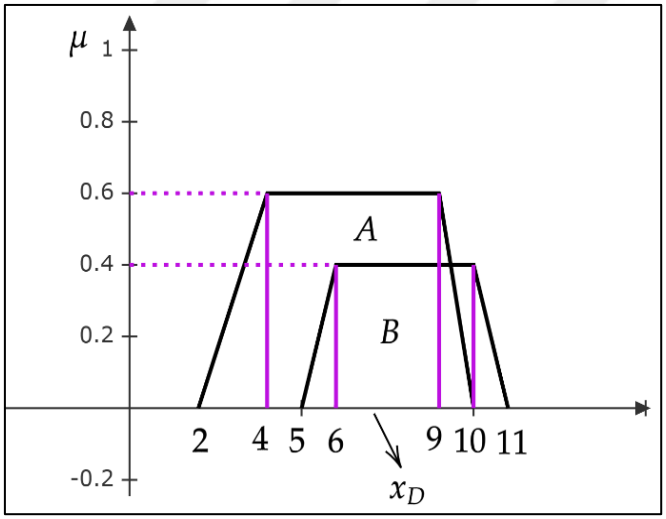


Figure 2.6. Center of sums method of $\tilde{A}$ and $\tilde{B}$.

Let the areas of these fuzzy sets be denoted by A and B, respectively. The calculations are as follows:

$$A = \frac{1}{2} \times [(10-2) + (9-4)] \times 0.6 = \frac{1}{2} \times 13 \times 0.6 = 3.9$$

$$B = \frac{1}{2} \times [(11-5) + (10-6)] \times 0.4 = \frac{1}{2} \times 10 \times 0.4 = 2$$

The center of area for the fuzzy set $\tilde{A}$, denoted as $\bar{x}_1$, is calculated as $a = \frac{9+4}{2} = 6.5$, and similarly, the center of area for the fuzzy set $\tilde{B}$, denoted as b, is $b = \frac{10+6}{2} = 8$. The defuzzified value x_D is then calculated using the weighted average formula:

$$x_D = \frac{Aa+Bb}{A+B} = \frac{(3.9\times6.5)+(2\times8)}{3.9+2} \approx 7.01$$

Thus, the defuzzified value x_D is approximately 7.01 as we see in Figure 2.6.

2.2.2. Maximum membership principle

This method, commonly referred to in the literature as the height method, is applicable when the fuzzy set in question possesses a well-defined peak. In other words, this defuzzification technique is specifically used for fuzzy sets that exhibit a distinct and clear peak. Mathematically, it can be expressed as $\mu(x_D) \geq \mu(x), \forall x \in X$, where each element x of the fuzzy set $\tilde{A}$ has a corresponding membership degree denoted by $\mu_{\tilde{A}}(x)$. The crisp value x_D, obtained through defuzzification represents the point at which the membership function $\mu_{\tilde{A}}(x)$ reaches its maximum. In other words, the maximum membership principle selects x_D such that it satisfies $\mu(x_D) = \max \mu(x)$ for all $x \in \tilde{A}$. Figure 2.7 provides an illustration of how the crisp value is determined using this principle.

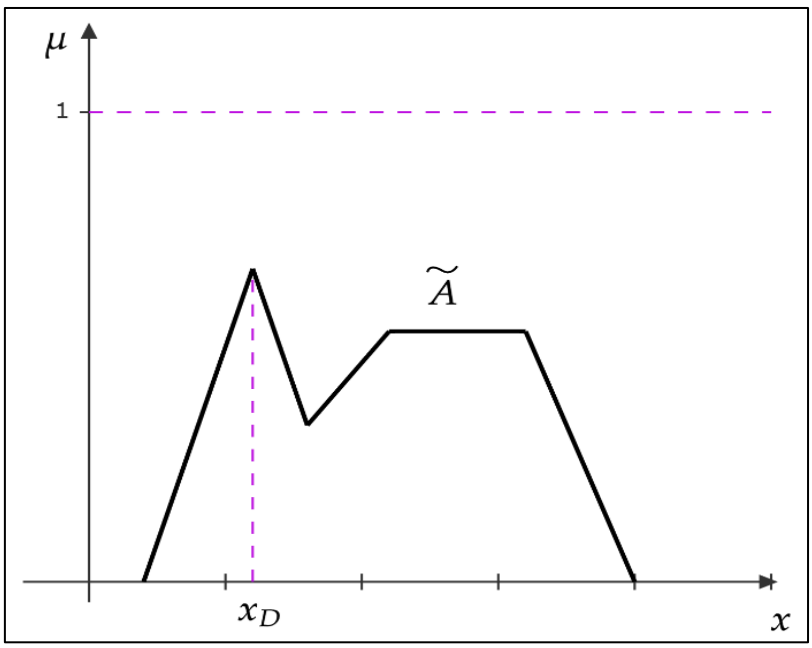


Figure 2.7. Maximum membership principle defuzzification method.

2.2.3. Principle of the maximum of the maximums and the minimum of the maximums

The principles of the maximum of the maximums and minimum of the maximums are frequently discussed in the literature due to their computational simplicity and ease of implementation. For these methods to be applicable, the fuzzy set under consideration must exhibit one or more peak points. In the defuzzification method known as the minimum of the maximums, the crisp value is determined by selecting the smallest element among those with the highest membership degrees. Let $\tilde{A}$ represent the fuzzy set, with each element denoted as x and its membership degree expressed as $\mu_{\tilde{A}}(x)$. The element x_l, which has the smallest value among those with the highest membership degree, is selected, and the corresponding crisp value obtained from the fuzzy set is denoted as x_D. The crisp value x_D determined using the minimum of the maximums method is depicted in Figure 2.8. This approach effectively captures the most significant elements of the fuzzy set while providing a clear and straightforward output, making it valuable in various practical applications.

Similarly, the defuzzification method that selects the largest element x_u among those with the highest membership degrees is known as the maximum of the maximums method. For the same fuzzy set $\tilde{A}$, this method is also illustrated in Figure 2.8.

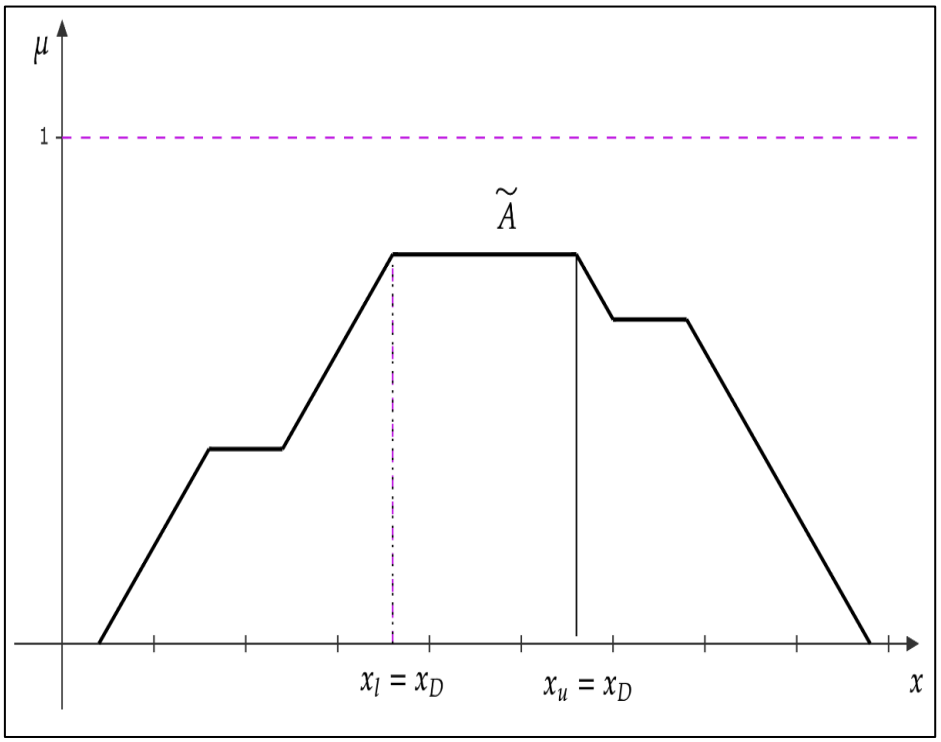


Figure 2.8. Minimum of the maximums and maximum of the maximums method.

2.2.4. Mean of maxima principle

The mean of maxima method is a widely employed defuzzification technique, particularly useful when a fuzzy set contains multiple elements with the highest membership degrees. This method determines the crisp value of the fuzzy set by averaging the values of the elements that exhibit the maximum membership degrees as expressed in Eq. 2.12.

$$x_D = \frac{x_l + x_u}{2} \tag{2.12}$$

where x_l and x_u can be founded as in previous section. This approach effectively captures the central tendency of the elements with the maximum membership values, providing a clear and balanced output that reflects the key characteristics of the fuzzy set.

2.3. Operations on Fuzzy Sets and Fuzzy Numbers

In addition to the standard set operations of union, intersection, and complement, fuzzy sets possess analogous operations, as originally defined in Zadeh's foundational work [4]. Prior to examining these fundamental fuzzy set operations, it is crucial to introduce the concept of containment, which plays a central role in both classical and fuzzy set theory. The notion of containment is an extension of its classical counterpart, adapted to accommodate the inherent flexibility of fuzzy sets [57].

Let X represent the universal set and let $\mu_{\tilde{A}}$ and $\mu_{\tilde{B}}$ denote the membership functions of the fuzzy sets $\tilde{A}$ and $\tilde{B}$, respectively. The set operations on fuzzy sets can be defined as follows [4, 6, 69]:

2.3.1 Definition: A fuzzy set $\tilde{A}$ is considered contained within another fuzzy set $\tilde{B}$ or equivalently, $\tilde{A}$ is a subset of $\tilde{B}$ if and only if $\mu_{\tilde{A}}(x)$ is less than or equal to $\mu_{\tilde{B}}(x)$ for every x. This relationship can be mathematically expressed as $\tilde{A} \subseteq \tilde{B}$ if and only if $\mu_{\tilde{A}}(x) \leq \mu_{\tilde{B}}(x)$.

2.3.2 Definition: For all $x \in X$, if $\mu_{\tilde{A}}(x) = \mu_{\tilde{B}}(x)$, then the fuzzy sets $\tilde{A}$ and $\tilde{B}$ are said to be equal, i.e., $\tilde{A} = \tilde{B} \Leftrightarrow \mu_{\tilde{A}}(x) = \mu_{\tilde{B}}(x)$.

2.3.3 Definition: The union of two fuzzy sets $\tilde{A}$ and $\tilde{B}$ is a fuzzy set $\tilde{C}$, denoted as $\tilde{C} = \tilde{A} \cup \tilde{B}$, whose membership function is given by $\mu_{\tilde{C}}(x) = \max\left(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\right)$. (Figure 2.9)

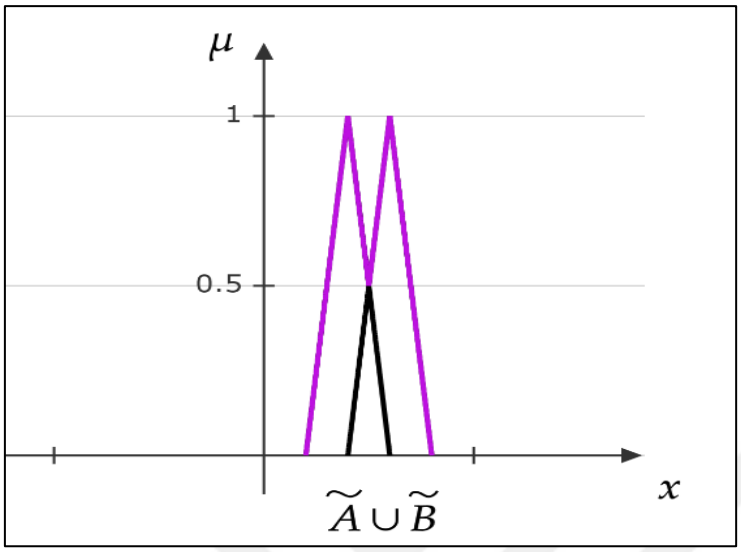


Figure 2.9. The union of two fuzzy sets $\tilde{A}$ and $\tilde{B}$.

2.3.4 Definition: The intersection of two fuzzy sets $\tilde{A}$ and $\tilde{B}$ is a fuzzy set $\tilde{C}$, denoted as $\tilde{C} = \tilde{A} \cap \tilde{B}$, whose membership function is given by $\mu_{\tilde{C}}(x) = \min\left(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\right)$. (Figure 2.10)

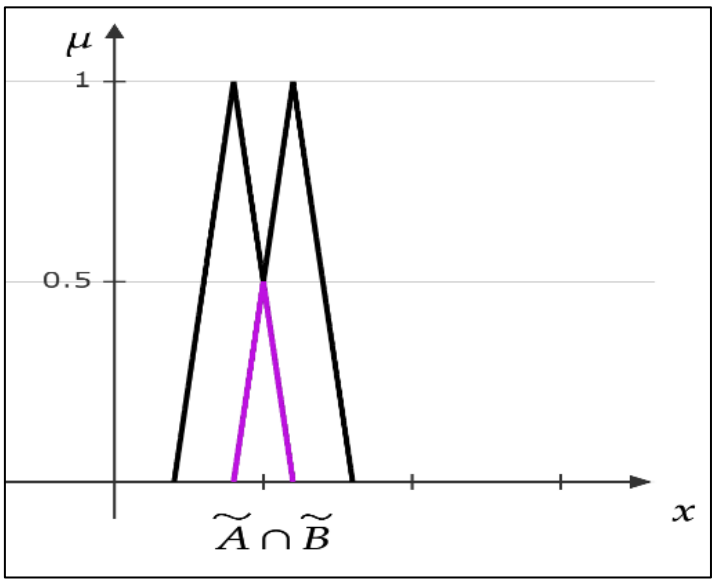


Figure 2.10. The intersection of two fuzzy sets $\tilde{A}$ and $\tilde{B}$.

2.3.5 Definition: Let $\tilde{C} = \tilde{A} \cdot \tilde{B}$, where $\mu_{\tilde{C}}(x) = \mu_{\tilde{A}}(x) \cdot \mu_{\tilde{B}}(x)$. The set formed by all values of x satisfying this condition is called the algebraic product of the fuzzy sets $\tilde{A}$ and $\tilde{B}$, i.e., $\tilde{C} = \left\{\left(x, \mu_{\tilde{C}}(x)\right) | \mu_{\tilde{C}}(x) = \mu_{\tilde{A}}(x) \cdot \mu_{\tilde{B}}(x)\right\}$.

2.3.6 Definition: Let $\tilde{C} = \tilde{A} \oplus \tilde{B}$, where $\mu_{\tilde{C}}(x) = \mu_{\tilde{A}}(x) + \mu_{\tilde{B}}(x) - \mu_{\tilde{A}}(x) \cdot \mu_{\tilde{B}}(x)$. The set formed by all values of x satisfying this condition is called the algebraic sum of the fuzzy sets $\tilde{A}$ and $\tilde{B}$, i.e., $\tilde{C} = \{(x, \mu_{\tilde{C}}(x)) | \mu_{\tilde{C}}(x) = \mu_{\tilde{A}}(x) + \mu_{\tilde{B}}(x) - \mu_{\tilde{A}}(x) \cdot \mu_{\tilde{B}}(x)\}$.

2.3.7 Definition: The complement of a fuzzy set $\tilde{A}$, denoted by $\neg\tilde{A}$, is defined by the membership function $\mu_{\neg\tilde{A}}(x) = 1 - \mu_{\tilde{A}}(x)$. (Figure 2.11)

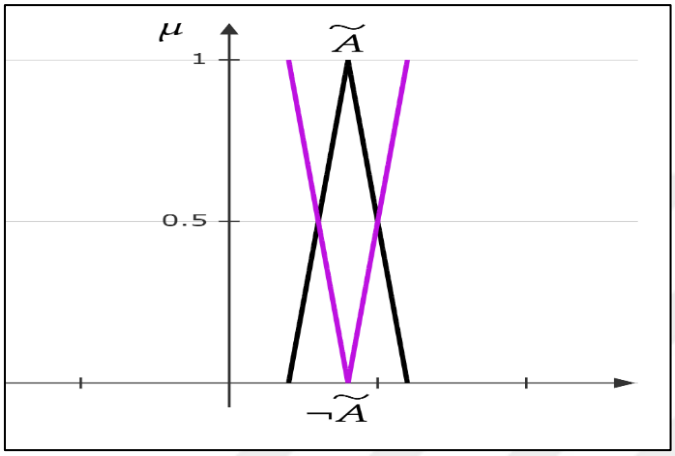


Figure 2.11. The complement of a fuzzy set $\tilde{A}$.

2.3.8 Definition: Let $\tilde{C} = \tilde{A} - \tilde{B}$, where $\mu_{\tilde{C}}(x) = \min\{\mu_{\tilde{A}}(x), \mu_{\neg\tilde{B}}(x)\}$. The set formed by all values of x satisfying this condition is called the difference of the fuzzy sets $\tilde{A}$ and $\tilde{B}$, i.e., $\tilde{C} = \{(x, \mu_{\tilde{C}}(x)) | \mu_{\tilde{C}}(x) = \min\{\mu_{\tilde{A}}(x), \mu_{\neg\tilde{B}}(x)\}\}$.

The most recently presented definitions are elucidated below through an illustrative example.

Example: Consider the universal set $X = \{1,2,\ldots,10\}$, and let $\tilde{A} = \{(1,0.5), (5,1), (9,0.3)\}$ and $\tilde{B} = \{(1,0.8), (4,1), (9,0.7)\}$ represent two fuzzy sets.

Union: $\tilde{A} \cup \tilde{B} = \{(1,0.8), (4,1), (5,1)(9,0.7)\}$.
Intersection: $\tilde{A} \cap \tilde{B} = \{(1,0.5), (9,0.3)\}$.
Algebraic product: $\tilde{A} \cdot \tilde{B} = \{(1,0.4), (9,0.21)\}$.
Algebraic sum: $\tilde{A} \oplus \tilde{B} = \{(1,0.9), (4,1), (5,1), (9,0.79)\}$.
Complement of $\tilde{B}$: $\neg\tilde{B} = \{(1,0.2), (2,1), (3,1), (5,1), (6,1), (7,1), (8,1), (9,0.3), (10,1)\}$.
Difference: $\tilde{A} - \tilde{B} = \{(1,0.2), (5,1), (9,0.3)\}$.

In fuzzy logic, arithmetic operations differ based on the representation of fuzzy numbers. Unlike classical logic, where exact numerical operations follow a single standard approach, arithmetic operations on fuzzy numbers are designed to accommodate their unique structure, which parallels but does not replicate classical logic. These operations are derived from classical arithmetic, with addition and subtraction maintaining commutative and associative properties, while multiplication and division do not inherently exhibit these properties. The implementation of arithmetic operations typically employs the extension principle or the α-cut method. The following two sections introduce these methodologies and illustrate their application through examples [2, 57, 70-74].

2.3.1. Arithmetic operations on fuzzy numbers using the extension principle

Arithmetic operations for discrete or countable fuzzy numbers can be conducted using the extension principle. Let # denote any arithmetic operation. To execute the fuzzy # operation via the extension principle, consider $x\#y = z$, where $x \in \tilde{A}$ and $y \in \tilde{B}$. The membership value of $z \in (\tilde{A}\#\tilde{B})$ is determined using Eq. 2.13 as follows [4, 6, 63]:

$$\mu_{\widetilde{A\#B}}(x) = max\{min\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(y)\}\} \tag{2.13}$$

Example: Let $\tilde{A} = \{(6, 0.5), (9, 1), (10, 0.7)\}$ and $\tilde{B} = \{(3, 0.4), (5, 0.8), (7, 1), (9,0.6), (12, 0.4)\}$ be two fuzzy numbers on the universal set X. Using the extension principle, the basic operations can be applied as follows:

Addition

$$\begin{aligned}\tilde{A} + \tilde{B} = \{&(6+3, min\{0.5,0.4\}), (6+5, min\{0.5,0.8\}), (6+7, min\{0.5,1\}), (6+9, min\{0.5,0.6\}, (6+12, min\{0.5,0.4\}), (9+3, min\{1,0.4\}), (9+5, min\{1,0.8\}), (9+7, min\{1,1\}), (9+9, min\{1,0.6\}), (9+12, min\{1,0.4\}), (10+3, min\{0.7,0.4\}, (10+5, min\{0.7,0.8\}), (10+7, min\{0.7,1\}), (10+9, min\{0.7,0.6\}), (10+12, min\{0.7,0.4\}\}\\ = \{&(9,0.4), (11,0.5), (13,0.5), (15,0.5), (18,0.4), (12,0.4), (14,0.8), (16,1), (18, 0.6), (21,0.4), (13,0.4), (15,0.7), (17,0.7), (19,0.6), (22,0.4)\}\end{aligned}$$

Using the extension principle, the maximum membership function value is selected for each corresponding value. It is observed that certain values may have multiple membership values.

$\tilde{A} + \tilde{B} = \{(9,0.4), (11,0.5), (12,0.4), (13,0.5), (14,0.8), (15,0.7), (16,1), (17,0.7), (18,0.6), (19,0.6), (21,0.4), (22,0.4)\}$

Subtraction: Let $\tilde{A} = \{(5, 0.5), (6, 1), (8, 0.6)\}$ and $\tilde{B} = \{(2,0.4), (3,0.7), (4,1), (5,0.6), (6,0.3)\}$. Subtraction using the extension principle can be represented as follows:

Table 2.1. Subtraction of two fuzzy numbers using the extension principle.

$\tilde{A} - \tilde{B}$	$(2, 0.4)$	$(3, 0.7)$	$(4, 1)$	$(5, 0.6)$	$(6, 0.3)$
$(5, 0.5)$	3 0.4	2 0.5	1 0.5	0 0.5	-1 0.3
$(6, 1)$	4 0.4	3 0.7	2 1	1 0.6	0 0.3
$(8, 0.6)$	6 0.4	5 0.6	4 0.6	3 0.6	2 0.3

Thus, $\tilde{A} - \tilde{B} = \{(-1,0.3), (0,0.5), (1,0.6), (2,1), (3,0.7), (4,0.6), (5,0.6), (6,0.4)\}$.

Multiplication: Let $\tilde{A} = \{(4, 0.3), (6, 1), (8, 0.6)\}$ and $\tilde{B} = \{(3,0.4), (4,0.7), (5,1), (6,0.6) (7,0.3)\}$. Multiplication using the extension principle can be represented as follows:

Table 2.2. Multiplication of two fuzzy numbers using the extension principle.

$\tilde{A} \cdot \tilde{B}$	$(3, 0.4)$	$(4, 0.7)$	$(5, 1)$	$(6, 0.6)$	$(7, 0.3)$
$(4, 0.3)$	12 0.3	16 0.3	20 0.3	24 0.3	28 0.3
$(6, 1)$	18 0.4	24 0.7	30 1	36 0.6	42 0.3
$(8, 0.6)$	24 0.4	32 0.6	40 0.6	48 0.6	56 0.3

$\tilde{A} \cdot \tilde{B} = \{(12,0.3), (16,0.3), (18,0.4), (20,0.3), (24,0.7), (28,0.3), (30,1), (32,0.6), (36, 0.6), (40,0.6), (42,0.3), (48,0.6), (56,0.3)\}$

Division: Let $\tilde{A} = \{(9, 0.5), (12, 1), (15, 0.4)\}$ and $\tilde{B} = \{(3, 0.6), (4, 1), (5, 0.5)\}$.

Table 2.3. Division of two fuzzy numbers using the extension principle.

$\frac{\tilde{A}}{\tilde{B}}$	$(3, 0.6)$	$(4, 1)$	$(5, 0.5)$
$(9, 0.5)$	~~3~~ ~~0.5~~	2.25 0.5	1.8 0.5
$(12, 1)$	4 0.6	3 1	2.4 0.5
$(15, 0.4)$	5 0.4	3.75 0.4	~~3~~ ~~0.4~~

Thus, $\frac{\tilde{A}}{\tilde{B}} = \{(1.8,0.5), (2.25,0.5), (2.4,0.5), (3,1), (3.75,0.4), (4,0.6), (5,0.4)\}$.

2.3.2. Arithmetic operations on fuzzy numbers using α-cut method

Suppose that $\tilde{A}$ and $\tilde{B}$ are two fuzzy numbers and let the # symbol represent basic arithmetic operations. In this case, the operation $\tilde{A}\#\tilde{B}$ is defined by Eq. 2.14, as follows [63, 75]:

$$\tilde{C} = \tilde{A}\#\tilde{B} = \bigcup_{\alpha\in[0,1]} \tilde{A}_\alpha \#\tilde{B}_\alpha = \left\{\left(x, \mu_{\tilde{C}}(x)\right): \exists y, z; x = y\#z, y \in \tilde{A}, z \in \tilde{B}\right\} \quad (2.14)$$

For the division operation denoted by /, it is assumed that $0 \notin \tilde{B}_\alpha$.

To explain arithmetic operations on fuzzy numbers more clearly, triangular fuzzy numbers will be considered. Let $\tilde{A} = (a, b, c)_T$ be a triangular fuzzy number. The membership function of $\tilde{A}$ is defined by Eq. 2.15, as follows:

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{x-a}{b-a} & ; \quad a \le x \le b \\ \frac{c-x}{c-b} & ; \quad b \le x \le c \\ 0 & ; \quad otherwise \end{cases} \quad (2.15)$$

According to the definition of the α-cut method, the α-cut set of the fuzzy number $\tilde{A}$ defines an interval. For the triangular fuzzy number $\tilde{A}$, the left and right reference functions according to α are expressed respectively as $\alpha = \frac{x-a}{b-a}$ and $\alpha = \frac{c-x}{c-b}$, where $\alpha \in [0,1]$.

The α-cut of the triangular fuzzy number $\tilde{A}_\alpha$ is given by $\tilde{A}_\alpha = [(b-a)\alpha + a\,, c - (c-b)\alpha]$, where the corresponding x values based on α can be calculated as $x = (b-a)\alpha + a$ and $x = c - (c-b)\alpha$. When $\alpha = 0$, it is clear that $\tilde{A}_0 = [a, c]$, and when $\alpha = 1$, $\tilde{A}_1 = [\mathrm{b}\,, \mathrm{b}] = b$. If the left and right lengths are equal, the fuzzy number is symmetric.

Fuzzy numbers can be positive, negative, or partially negative. For example, for a triangular fuzzy number $\tilde{A} = (a, b, c)$ [76]:

- When $\alpha > 0$, $\tilde{A}$ is positive.
- When $c < 0$, $\tilde{A}$ is negative.
- When $a < 0$ and at least one value is positive, $\tilde{A}$ is partially negative.

Arithmetic operations using the α-cut method is performed by following these steps [4, 6, 71, 77]:

1. Determine the α-cuts of the fuzzy numbers.
2. Perform the operation according to interval arithmetic rules.
3. Calculate the left and right bounds, or x values, of the result based on the α value.
4. Calculate the membership function values over the intervals for each α value in the result.
5. Determine the membership function of the performed operation.

Let $\tilde{A} = (a, b, c)_T$ and $\tilde{B} = (d, e, f)_T$ be two triangular fuzzy numbers, with their membership functions given as follows:

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{x-a}{b-a} & ; \quad a \le x \le b \\ \frac{c-x}{c-b} & ; \quad b \le x \le c \\ 0 & ; \quad otherwise \end{cases} \qquad \mu_{\tilde{B}}(x) = \begin{cases} \frac{x-d}{e-d} & ; \quad d \le x \le e \\ \frac{f-x}{f-e} & ; \quad e \le x \le f \\ 0 & ; \quad otherwise \end{cases}$$

According to the proposed steps, the α-cuts of the fuzzy numbers $\tilde{A}$ and $\tilde{B}$ are given as $\tilde{A}_\alpha = [(b-a)\alpha + a\,, c-(c-b)\alpha]$ and $\tilde{B}_\alpha = [(e-d)\alpha + d\,, f-(f-e)\alpha]$, respectively. Now, using these α-cuts, let's explain how to apply the above steps to perform some arithmetic operations.

Addition:

2. Step: $\tilde{A}_\alpha + \tilde{B}_\alpha = [(b-a)\alpha + a\,, c-(c-b)\alpha] + [(e-d)\alpha + d\,, f-(f-e)\alpha]$

$= [(b-a+e-d)\alpha + a + d, c+f-(c-b+f-e)\alpha]$

3. Step: $x = (b-a+e-d)\alpha + a + d$ and $x = c+f-(c-b+f-e)\alpha$

4. Step: $\alpha = \frac{x-(a+d)}{(b+e)-(a+d)};\ (a+d) \le x \le (b+e)$ and

$\alpha = \frac{(c+f)-x}{(c+f)-(b+e)};\ (b+e) \le x \le (c+f)$

5. Step: $\mu_{\tilde{A}+\tilde{B}}(x) = \begin{cases} \frac{x-(a+d)}{(b+e)-(a+d)} & ;\quad (a+d) \le x \le (b+e) \\ \frac{(c+f)-x}{(c+f)-(b+e)} & ;\quad (b+e) \le x \le (c+f) \\ 0 & ;\quad otherwise \end{cases}$

Subtraction:

$$\mu_{\tilde{A}-\tilde{B}}(x) = \begin{cases} \frac{x-(a-f)}{(b-e)-(a-f)} & ;\quad (a-f) \le x \le (b-e) \\ \frac{(c-d)-x}{(c-d)-(b-e)} & ;\quad (b-e) \le x \le (c-d) \\ 0 & ;\quad otherwise \end{cases}$$

Multiplication: The multiplication operation can yield different results for positive and non-positive numbers. Let $\tilde{A}$ and $\tilde{B}$ be two positive fuzzy numbers, so that for $\tilde{A}$, $a > 0$, and for $\tilde{B}$, $d > 0$.

2. Step: $\tilde{A}_\alpha \cdot \tilde{B}_\alpha = [(b-a)\alpha + a\,, c-(c-b)\alpha] \cdot [(e-d)\alpha + d\,, f-(f-e)\alpha]$

$= \big[\big((b-a)\alpha + a\big)\big((e-d)\alpha + d\big), (c-(c-b)\alpha)(f-(f-e)\alpha)\big]$

3. Step: $x = (b-a)(e-d)\alpha^2 + \big((b-a)d + (e-d)a\big)\alpha + ad$

$x = (c-b)(f-e)\alpha^2 - \big((f-e)c + (c-b)f\big)\alpha + cf$

4. Step: $\alpha = \frac{-((b-a)l+(e-d)a)+\sqrt{((b-a)d+(e-d)a)^2-4(b-a)(e-d)(ad-x)}}{2(b-a)(e-d)}$; $ad \le x \le be$

$$\alpha = \frac{((f-e)c+(c-b)f)-\sqrt{((f-e)c+(c-b)f)^2-4(c-b)(f-e)(cf-x)}}{2(c-b)(f-e)}; \quad be \le x \le cf$$

5. Step: $\mu_{\tilde{A}\cdot\tilde{B}}(x) = \begin{cases} \frac{-((b-a)l+(e-d)a)+\sqrt{((b-a)d+(e-d)a)^2-4(b-a)(e-d)(ad-x)}}{2(b-a)(e-d)} & ; \quad ad \le x \le be \\ \frac{((f-e)c+(c-b)f)-\sqrt{((f-e)c+(c-b)f)^2-4(c-b)(f-e)(cf-x)}}{2(c-b)(f-e)} & ; \quad be \le x \le cf \\ 0 & ; \quad otherwise \end{cases}$

<u>*Division*</u>: Similar to multiplication, division can yield different results depending on whether the numbers are positive or non-positive. Let be two positive fuzzy numbers, so that for $\tilde{A}$, $a > 0$, and for $\tilde{B}$, $d > 0$.

2. Step: $\frac{\tilde{A}_\alpha}{\tilde{B}_\alpha} = \frac{[(b-a)\alpha+a\ ,c-(c-b)\alpha]}{[(e-d)\alpha+d\ ,f-(f-e)\alpha]} = \left[\frac{(b-a)\alpha+a}{f-(f-e)\alpha}, \frac{c-(c-b)\alpha}{(e-d)\alpha+d}\right]$

3. Step: $x = \frac{(b-a)\alpha+a}{f-(f-e)\alpha}$ and $x = \frac{c-(c-b)\alpha}{(e-d)\alpha+d}$

4. Step: $\alpha = \frac{xf-a}{(b-a)+(f-e)x}$; $\frac{a}{f} \le x \le \frac{b}{e}$ and $\alpha = \frac{c-dx}{(c-b)+(e-d)x}$; $\frac{b}{e} \le x \le \frac{c}{d}$

5. Step: $\mu_{\frac{\tilde{A}}{\tilde{B}}}(x) = \begin{cases} \frac{xf-a}{(b-a)+(f-e)x} & ; \quad \frac{a}{f} \le x \le \frac{b}{e} \\ \frac{c-dx}{(c-b)+(e-d)x} & ; \quad \frac{b}{e} \le x \le \frac{c}{d} \\ 0 & ; \quad otherwise \end{cases}$

Example: Let $\tilde{A} = (4,5,9)_T$ and $\tilde{B} = (3,7,9)_T$ be two triangular fuzzy numbers. The α-cuts of $\tilde{A}$ and $\tilde{B}$ are found as $\tilde{A}_\alpha = [\alpha + 4, 9 - 4\alpha]$ and $\tilde{B}_\alpha = [4\alpha + 3, 9 - 2\alpha]$.

The result of the addition: $\mu_{\tilde{A}+\tilde{B}}(x) = \begin{cases} \frac{x-7}{5} & ; \quad 7 \le x \le 12 \\ \frac{18-x}{6} & ; \quad 12 \le x \le 18 \\ 0 & ; \quad otherwise \end{cases}$

The result of the subtraction: $\mu_{\tilde{A}-\tilde{B}}(x) = \begin{cases} \frac{x+5}{3} & ; \quad -5 \le x \le -2 \\ \frac{6-x}{8} & ; \quad -2 \le x \le 6 \\ 0 & ; \quad otherwise \end{cases}$

The result of the multiplication: $\mu_{\tilde{A}\cdot\tilde{B}}(x) = \begin{cases} \frac{-19+\sqrt{361-16(12-x)}}{8} & ; \quad 12 \leq x \leq 35 \\ \frac{54-\sqrt{2916-32(81-x)}}{16} & ; \quad 35 \leq x \leq 81 \\ 0 & ; \quad otherwise \end{cases}$

The result of the division: $\mu_{\frac{\tilde{A}}{\tilde{B}}}(x) = \begin{cases} \frac{9x-4}{1+2x} & ; \quad \frac{4}{9} \leq x \leq \frac{5}{7} \\ \frac{9-3x}{4+4x} & ; \quad \frac{5}{7} \leq x \leq 3 \\ 0 & ; \quad otherwise \end{cases}$

For additional operations, along with step-by-step examples and detailed explanations, reference [63] provides the necessary information.

3. CLASSICAL REGRESSION AND FUZZY REGRESSION

3.1. Ordinary Least Squares

Regression analysis is a foundational technique for examining and quantifying the relationships between variables. By assessing how one or more independent variables influence or predict the value of a dependent variable, this method offers critical insights into complex systems. Statistical tools are employed to explore and model these relationships, enabling the identification of patterns and the formulation of predictions based on empirical data. In regression models, independent variables, also referred to as predictors, are those that impact the outcome, while the dependent variable represents the outcome being measured. This analytical framework provides researchers and practitioners with a robust basis for making informed decisions and drawing meaningful conclusions from data.

Regression analysis in the literature is classified based on several key criteria, including the number of independent variables, the form of the regression function, the type of dependent variable, and the structural properties of the model. Models are categorized as simple regression, involving a single independent variable, or multiple regression, incorporating two or more independent variables. The regression function can take various forms, such as linear for straight-line relationships, non-linear for complex interactions, or polynomial for higher-order terms. The type of dependent variable further refines the classification into continuous, binary, count data, ordinal, or multinomial. Additionally, regression models are classified as parametric, which assume a predefined functional form, or non-parametric, which are more flexible and do not impose such assumptions. Features such as interaction terms, outlier handling, and regularization techniques can further enhance and customize regression models, enabling their application to diverse datasets and research contexts [53, 78].

In simple linear regression, the relationship between a dependent variable and a single independent variable is analyzed. In contrast, multiple regression models examine the effect of multiple independent variables on a single dependent variable. The conditional expected value $E(\boldsymbol{Y}|X_1 = x_1, X_2 = x_2, \dots, X_p = x_p)$ takes the form $f(x_1, x_2, \dots, x_p)$, where the

function f determines the type of regression model. If f is a linear function of $x_1, x_2, \dots, x_p$, the model is referred to as a multiple linear regression model, or simply multiple regression. The multiple regression equation with p independent variables is expressed as in Eq. 3.1:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} + \varepsilon_i \qquad i = 1,2,\dots,n \tag{3.1}$$

The matrix representation of the regression model in Eq. 3.1 is given by Eq. 3.2, as follows:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{p1} \\ 1 & x_{12} & x_{22} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{pn} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{3.2}$$

where $\boldsymbol{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$, $\boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{p1} \\ 1 & x_{12} & x_{22} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{pn} \end{bmatrix}$, $\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}$, and $\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$.

OLS method is a widely recognized and efficient method for estimating unknown coefficients in linear regression models, appreciated for its simplicity and effectiveness in identifying relationships between variables. The OLS estimator is consistent when the regressors are exogenous, meaning that the regressors are uncorrelated with the error term. According to the Gauss-Markov theorem, it is also optimal within the class of linear unbiased estimators when the errors are homoscedastic and serially uncorrelated. Under these conditions, OLS provides minimum-variance, mean-unbiased estimates when the errors exhibit finite variances. Moreover, when the errors follow a normal distribution, OLS serves as the maximum likelihood estimator. However, it is crucial for the errors to conform to a normal distribution after model fitting to ensure valid conclusions through hypothesis testing. The primary objective of the OLS method is to minimize the errors present in the regression model. Since the errors may be both positive and negative, the approach focuses on minimizing the squared errors rather than the errors themselves.

Equation 3.2 can be compactly written as illustrated in Equation 3.3:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{3.3}$$

By multiplying both sides of the equation by the transpose of matrix $\boldsymbol{X}$, the normal equations are given by Eq. 3.4:

$$\boldsymbol{X'X\beta} = \boldsymbol{X'Y} \tag{3.4}$$

The solution to the normal equation in Equation 3.4 is obtained using Equation 3.5, as follows:

$$\widehat{\boldsymbol{\beta}}_{OLS} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'Y} \tag{3.5}$$

which are the estimators of the model coefficients using the OLS estimation method.

For the OLS estimators of the coefficients, the expected value and variance are derived in Eq. 3.6 and Eq. 3.7, as follows:

$$E(\widehat{\boldsymbol{\beta}}_{OLS}) = \boldsymbol{\beta} \tag{3.6}$$

$$Var(\widehat{\boldsymbol{\beta}}_{OLS}) = \sigma^2(\boldsymbol{X'X})^{-1} \tag{3.7}$$

3.2. Multicollinearity

In multiple linear regression models, the existence of correlations among the independent variables results in the issue of multicollinearity. This condition occurs when two or more independent variables are highly correlated, which can cause difficulties in estimating the regression coefficients accurately. Multicollinearity leads to inflated standard errors, making it challenging to determine the individual effect of each independent variable on the dependent variable. Consequently, the estimates of the regression coefficients may become unstable, resulting in unreliable and misleading conclusions from the model [50].

Independent variables that are not linearly related to each other are considered orthogonal, and when this condition holds, the inferences drawn from the estimated coefficients are more likely to reflect the true relationships in the data. However, in practical regression applications, independent variables are seldom orthogonal. In many instances, these variables may exhibit significant multicollinearity, which poses a challenge to accurate

coefficient estimation. One of the key assumptions for ensuring unbiased estimates in the classical linear regression model using the OLS method is the absence of a linear relationship among the independent variables. When this assumption is violated, multicollinearity arises, which can severely affect the reliability of the regression model. Specifically, the presence of multicollinearity leads to inflated standard errors of the coefficient estimates, making it difficult to ascertain the individual contributions of each independent variable and resulting in potentially misleading inferences and predictions. Therefore, it is critical to understand the conditions under which multicollinearity emerges and the potential consequences it may have on the validity of the model. Addressing this issue necessitates the application of alternative methods, such as Ridge regression [12] or Liu regression [13], which have been developed to mitigate the effects of multicollinearity and improve the accuracy of coefficient estimates.

Several factors contribute to the occurrence of multicollinearity, including the following [50]:

- Data collection methodology: The way in which data is collected can induce multicollinearity. For example, survey methods that elicit similar responses for correlated variables can artificially inflate the correlations among these variables.
- Sample representation: When the sample does not accurately represent the broader population, it can lead to spurious correlations. For instance, a sample that is biased toward a specific demographic may reveal relationships between variables that would not be observed in a more representative sample.
- Actual relationships among variables: In some cases, independent variables are inherently correlated due to their measurement of related concepts. For example, height and weight are often strongly correlated, which may lead to multicollinearity when both are included as explanatory variables.
- Observation count vs. variable count: Multicollinearity can also occur when the number of observations n is less than the number of independent variables $p + 1$, resulting in an unstable model. In such situations, the model fails to accurately estimate the relationships among the variables, leading to large variances in the coefficient estimates.
- Researcher-induced errors: Multicollinearity can also arise from errors in the model specification, such as incorrectly coding variables or including redundant or unnecessary

variables. These mistakes can artificially increase the correlations among the independent variables and distort the regression results.

Identifying the precise source of multicollinearity is often challenging; however, understanding its structural characteristics is crucial for researchers. A comprehensive grasp of the nature of multicollinearity enables the development and application of appropriate strategies to mitigate its effects.

The severity of multicollinearity directly correlates with its potential to produce problematic outcomes. Several diagnostic methods are available to detect multicollinearity by analyzing the structural properties of the data [79]. These methods are summarized below [50]:

- Examination of the $X'X$ matrix: A straightforward approach to detecting multicollinearity involves inspecting the off-diagonal elements of the $\boldsymbol{X'X}$ matrix. High correlations between explanatory variables indicate multicollinearity. However, this method is limited to pairwise relationships and cannot detect more complex forms of multicollinearity.
- Determinant of the $X'X$ matrix: The determinant range of the $\boldsymbol{X'X}$ matrix ranges from $0 \leq |\boldsymbol{X'X}| \leq 1$. A determinant value close to 0 signifies strong multicollinearity, though this measure does not provide insights into its underlying causes.
- Analysis of R^2 and tolerance: When regressing an independent variable on other independent variables, a high R^2 value nearing one indicates multicollinearity. Tolerance, defined as $1 - R^2$, provides a complementary measure; lower tolerance values approaching 0 reflect stronger multicollinearity.
- Variance inflation factor (VIF): Introduced by Marquardt [80], the VIF quantifies the extent to which the variance of a regression coefficient is inflated due to multicollinearity. A VIF greater than 10 suggests substantial multicollinearity.
- Eigenvalues of the $X'X$ matrix: The eigenvalues λ of the $\boldsymbol{X'X}$ matrix offer insights into multicollinearity. Eigenvalues close to 0 indicate potential multicollinearity.
- Condition number: Defined as the ratio of the largest eigenvalue λ_{max} to the smallest eigenvalue λ_{min}, the condition number provides a numerical assessment of multicollinearity. A condition number below 100 signifies negligible multicollinearity, while values exceeding 1000 indicate severe multicollinearity. The square root of the

condition number above 10 implies a problem, and values surpassing 20 necessitate corrective actions.

Additionally, other diagnostic tools, such as the F-statistic and t-statistic, can provide indications of multicollinearity. For example, a significant F-statistic paired with insignificant t-statistics often signals the presence of multicollinearity. Furthermore, substantial variations in coefficient estimates upon the addition or removal of a variable may also highlight multicollinearity [78].

3.3. Ridge Regression

In regression models affected by multicollinearity, Hoerl and Kennard [12] introduced the Ridge regression estimation method to enhance the accuracy of coefficient estimates. This approach incorporates a biasing parameter k into the diagonal elements of the $\boldsymbol{X'X}$ matrix, effectively mitigating the impact of multicollinearity [50].

The Ridge regression estimator is defined in Eq. 3.8, where k denotes the bias parameter, and, and $\boldsymbol{I}$ represents the identity matrix:

$$\widehat{\boldsymbol{\beta}}_R = (\boldsymbol{X'X} + k\boldsymbol{I})^{-1}\boldsymbol{X'Y}, \quad k \geq 0 \tag{3.8}$$

When $k = 0$, the Ridge regression estimator simplifies to the OLS estimator. The relationship between the Ridge and OLS estimators is expressed in Eq. 3.9 as a linear transformation of the OLS estimator:

$$\widehat{\boldsymbol{\beta}}_R = (\boldsymbol{X'X} + k\boldsymbol{I})^{-1}\boldsymbol{X'Y} = (\boldsymbol{X'X} + k\boldsymbol{I})^{-1}(\boldsymbol{X'X})\widehat{\boldsymbol{\beta}}_{OLS} = \boldsymbol{Z}_k\widehat{\boldsymbol{\beta}}_{OLS} \tag{3.9}$$

Here, $\boldsymbol{Z}_k$ is defined as $(\boldsymbol{X'X} + k\boldsymbol{I})^{-1}(\boldsymbol{X'X})$. The expected value of the Ridge regression estimator is given in Eq. 3.10.

$$E(\widehat{\boldsymbol{\beta}}_R) = (\boldsymbol{X'X} + k\boldsymbol{I})^{-1}\boldsymbol{X'X}\boldsymbol{\beta} \tag{3.10}$$

As shown, $E(\widehat{\boldsymbol{\beta}}_R) \neq \boldsymbol{\beta}$, indicating that the Ridge regression estimator is biased. The variance-covariance matrix of the Ridge estimator is given in Eq. 3.11.

$$Var(\widehat{\boldsymbol{\beta}}_R) = \sigma^2(\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-1}(\boldsymbol{X}'\boldsymbol{X})(\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-1} \quad (3.11)$$

The MSE of an estimator quantifies the deviation of the estimated coefficient from the true coefficient. Ridge regression often yields estimates closer to the true value compared to OLS. The MSE of the Ridge regression estimator is expressed in Eq. 3.12.

$$MSE(\widehat{\boldsymbol{\beta}}_R) = \sigma^2 \sum_{j=1}^{p} \frac{\lambda_j}{(\lambda_j+k)^2} + k^2\boldsymbol{\beta}(\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-2}\boldsymbol{\beta} \quad (3.12)$$

In Eq. 3.12, $\lambda_1, \lambda_2, \dots, \lambda_p$ denote the eigenvalues of the $\boldsymbol{X}'\boldsymbol{X}$ matrix, which represent the variance components of the coefficients and the squared bias term. As the bias parameter k increases, the variance decreases, while the bias term increases. The primary objective of Ridge regression is to determine a value of k that minimizes the MSE, where the reduction in variance compensates for the increase in bias. Hoerl and Kennard [12] established that such a k value exists, ensuring that $MSE(\widehat{\boldsymbol{\beta}}_R) < MSE(\widehat{\boldsymbol{\beta}}_{OLS})$ [50, 78].

The VIF for Ridge regression is presented in Eq. 3.13.

$$[VIF_1, VIF_2, \dots, VIF_p] = diag[(\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-1}(\boldsymbol{X}'\boldsymbol{X})(\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-1}] \quad (3.13)$$

The values λ_i corresponding to the i-th eigenvalue of the matrix $\boldsymbol{C} = \boldsymbol{X}'\boldsymbol{X}$ are arranged in descending order as $\lambda_1 > \lambda_2 > \cdots > \lambda_p$. The orthogonal matrix $\boldsymbol{D}$, which satisfies the equation $\boldsymbol{D}'\boldsymbol{C}\boldsymbol{D} = \boldsymbol{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$, is formed by considering the eigenvectors corresponding to these eigenvalues as columns. The canonical form of the classical multiple linear regression model is expressed as shown in Eq 3.14:

$$\boldsymbol{Y} = \boldsymbol{Z}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \quad (3.14)$$

Here, $\boldsymbol{D}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{D} = \boldsymbol{Z}'\boldsymbol{Z} = \boldsymbol{\Lambda}$, $\boldsymbol{\theta} = \boldsymbol{D}'\boldsymbol{B}$, and $\boldsymbol{Z} = \boldsymbol{X}\boldsymbol{D}$. Consequently, the OLS estimator for $\boldsymbol{\theta}$, is given by Eq 3.15 [78].

$$\widehat{\boldsymbol{\theta}}_{OLS} = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{Y} = \boldsymbol{\Lambda}^{-1}\boldsymbol{Z}'\boldsymbol{Y} \quad (3.15)$$

Now, with $\boldsymbol{A} = \boldsymbol{\Lambda} + k\boldsymbol{I}$, the canonical form of the Ridge regression estimator $\widehat{\boldsymbol{\theta}}_R$ is defined in Eq 3.16, as follows:

$$\widehat{\boldsymbol{\theta}}_R = \boldsymbol{A}^{-1}\boldsymbol{Z}'\boldsymbol{Y} \tag{3.16}$$

The parameter k, as previously mentioned, is a fundamental determinant in the Ridge regression method. Its value can be determined through various approaches and methodologies, as outlined below:

One prominent method is the Ridge trace method, introduced by Hoerl and Kennard [12]. This diagnostic tool examines the behavior of regression coefficients as the regularization parameter k varies. The Ridge trace method entails fitting Ridge regression models over a range of k values. Typically, the range begins with a large k and gradually decreases to a smaller value, approaching the OLS solution. The resulting Ridge trace plot illustrates the trajectory of each coefficient relative to k, providing insights into the stability of the model. A well-behaved Ridge trace plot exhibits smooth coefficient paths as k changes, whereas abrupt changes may indicate instability or suboptimal k values.

Another approach involves generating a set of k values within a predefined range using either fixed or randomized steps. These values are used to construct Ridge regression models, which are then assessed using a specific error measure. While straightforward, this method can be computationally intensive and may not consistently yield optimal results. The selection of k values and the evaluation metric significantly influence its effectiveness.

Among the most widely adopted techniques is K-fold cross-validation, which is employed in this study for determining the optimal k. This method is highly regarded for balancing bias and variance effectively. K-fold cross-validation is a robust procedure for evaluating model performance while mitigating overfitting. It ensures that each data point serves as part of the validation set, thereby enhancing the reliability of performance estimates. Historically rooted in statistical model validation, K-fold cross-validation does not have a single universally acknowledged originator.

The process involves partitioning the dataset into K subsets, commonly five or ten folds. For each iteration, the model is trained on $K - 1$ folds and validated on the remaining fold. This

process is repeated K times, with each fold serving as the validation set once. The average validation results across all folds are then used to determine the optimal k value. This approach involves evaluating a range of candidate k values to identify the one that minimizes the average validation error [81].

A critical consideration in these methods is the selection of the largest k value to test, which may approach infinity. One procedure for this selection, as described in [82], involves the following steps:

1. Identify the minimum k_{min} and maximum k_{max} values for k. The minimum value is a small positive number (e.g., 0.001, 0.0001), while the maximum value is chosen to sufficiently constrain the model coefficients toward 0.
2. Compute a sequence of k values using Eq. 3.17.

$$k_i = k_{min}\left(\frac{k_{max}}{k_{min}}\right)^{\frac{i-1}{v-1}} \qquad i = 1,2,\dots,v \tag{3.17}$$

Here, k_i denotes the i-th value in the sequence, and v represents the total number of candidate values. The logarithmically spaced values between k_{min} and k_{max} are computed based on Eq 3.18:

$$\lambda_{max} = \frac{\|\boldsymbol{X}^t\boldsymbol{Y}\|_\infty}{\|\boldsymbol{X}\|_2^2} \tag{3.18}$$

where $\|\boldsymbol{X}^t\boldsymbol{Y}\|_\infty$ is the infinity norm of $\boldsymbol{X}^t\boldsymbol{Y}$, defined as $\max_j \sum_i \left|(\boldsymbol{X}^t\boldsymbol{Y})_{ij}\right|$, and $\|\boldsymbol{X}\|_2^2$ is the squared Euclidean norm of $\boldsymbol{X}$, calculated as $\sqrt{\sum_{i,j}\boldsymbol{X}_{ij}^2}$.

In addition to, K-fold cross-validation, which relies on specified k_{min} and k_{max} values or randomly chosen ranges with steps, other researchers have developed formulas to determine the optimal k value. These formulas often surpass K-fold cross-validation in efficiency and accuracy, especially given the latter's computational demands and occasional inability to identify the optimal k. Since Hoerl and Kennard [12], numerous researchers have proposed

new methods for determining the bias parameter k in Ridge regression. Let $m_i = \sqrt{\frac{\hat{\sigma}^2}{\hat{\theta}_i^2}}$ and $q_i = \frac{\lambda_{max}}{(n-p)\hat{\sigma}^2+\lambda_{max}\hat{\theta}_i^2}$, where $\theta_i = \boldsymbol{D}'\boldsymbol{\beta}$ for $i = 1,2,\dots,p$. Table 3.1 presents 77 estimators, which are frequently compared in the literature [50].

Table 3.1. Some of the proposed formulas for finding the optimal bias parameters for Ridge regression.

Proposed Bias Parameter	Year	Authors
$k_1 = \frac{\hat{\sigma}^2}{\hat{\theta}_{max}^2},\ k_2 = \frac{\hat{\sigma}^2}{\hat{\theta}_{max}},\ k_3 = \frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\theta}_i^2}$	1970	Hoerl and Kennard [12]
$k_4 = \frac{p\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}}$	1975	Hoerl et al. [83]
$k_5 = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{p}\lambda_i\hat{\theta}_i^2}$	1976	Lawless and Wang [84]
$k_6 = \hat{\sigma}^2\frac{\sum_{i=1}^{p}(\lambda_i\hat{\theta}_i)^2}{\sum_{i=1}^{p}(\lambda_i\hat{\theta}_i^2)^2}$	1976	Hocking et al. [85]
$k_7 = \frac{1}{\hat{\theta}_{max}^2}$	1984	Schaefer et al. [86]
$k_8 = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{p}\left\{\frac{\hat{\theta}_i^2}{1+\left(1+\lambda_i\left(\frac{\hat{\theta}_i^2}{\hat{\sigma}^2}\right)^{1/2}\right)}\right\}}$	1988	Nomura [87]
$k_9 = \frac{\hat{\sigma}^2}{(\prod_{i=1}^{p}\hat{\theta}_i^2)^{1/p}}, k_{10} = median\left(\frac{\hat{\sigma}^2}{\hat{\theta}_i^2}\right), k_{11} = \frac{1}{p}\sum_{i=1}^{p}\frac{\hat{\sigma}^2}{\hat{\theta}_i^2}$	2003	Kibria [88]
$k_{12} = \frac{\lambda_{max}\hat{\sigma}^2}{(n-p-1)\hat{\sigma}^2+\lambda_{max}\hat{\beta}_{max}^2}$	2005	Khalaf and Shukur [89]
$k_{13} = \max\left(\frac{\lambda_i\hat{\sigma}^2}{(n-p)\hat{\sigma}^2+\lambda_i\hat{\beta}_i^2}\right), k_{14} = \frac{1}{p}\sum_{i=1}^{p}\left(\frac{\lambda_i\hat{\sigma}^2}{(n-p)\hat{\sigma}^2+\lambda_i\hat{\beta}_i^2}\right),$ $k_{15} = \text{median}\left(\frac{\lambda_i\hat{\sigma}^2}{(n-p)\hat{\sigma}^2+\lambda_i\hat{\beta}_i^2}\right), k_{16} = \max\left(\frac{s^2}{\hat{\beta}_i^2}\right)$	2006	Alkhamisi et al. [90]

Table 3.1 (Continue). Some of the proposed formulas for finding the optimal bias parameters for Ridge regression.

Proposed Bias Parameter	Year	Authors
$k_{17}=\frac{p\hat{\sigma}^2}{\hat{\theta}'\hat{\theta}}+\frac{1}{\lambda_{max}}$, $k_{18}=max\left(\frac{\hat{\sigma}^2}{\hat{\beta}_i^2}+\frac{1}{\lambda_i}\right)$, $k_{19}=\frac{1}{p}\sum_{i=1}^{p}\left(\frac{\hat{\sigma}^2}{\hat{\beta}_i^2}+\frac{1}{\lambda_i}\right)$, $k_{20}=median\left(\frac{\hat{\sigma}^2}{\hat{\beta}_i^2}+\frac{1}{\lambda_i}\right)$, $k_{21}=\frac{p\hat{\sigma}^2}{\sum_{i=1}^{p}\lambda_i\hat{\beta}_i^2}+\frac{1}{\lambda_{max}}$, $k_{22}=\frac{S^2}{\hat{\beta}_{max}^2}+\frac{1}{\lambda_{max}}$	2007	Alkhamisi and Shukur [91]
$k_{23}=\left(\prod_{i=1}^{p}\frac{\lambda_i\hat{\sigma}^2}{(n-p)\hat{\sigma}^2+\lambda_i\hat{\theta}_i^2}\right)^{1/p}$, $k_{24}=max\left(\frac{1}{m_i}\right)$, $k_{25}=max(m_i)$, $k_{26}=\left(\prod_{i=1}^{p}\frac{1}{m_i}\right)^{1/p}$, $k_{27}=\left(\prod_{i=1}^{p}m_i\right)^{1/p}$, $k_{28}=median\left(\frac{1}{m_i}\right)$, $k_{29}=median(m_i)$	2009	Muniz and Kibria [92]
$k_{30}=\hat{\sigma}^2\frac{\sum_{i=1}^{p}(\lambda_i\theta_i)^2}{\sum_{i=1}^{p}(\lambda_i\theta_i^2)^2}+\frac{1}{\lambda_{max}}$	2010	Al-Hassan [93]
$k_{31}=max\left(\frac{1}{q_i}\right)$, $k_{32}=max(q_i)$, $k_{33}=\prod_{i=1}^{p}\left(\frac{1}{q_i}\right)^{1/p}$, $k_{34}=\prod_{i=1}^{p}(q_i)^{1/p}$, $k_{35}=median\left(\frac{1}{q_i}\right)$, $k_{36}=median(q_i)$	2012	Kibria et al. [94]
$k_{37}=\frac{p\hat{\sigma}^2}{\hat{\alpha}'\hat{\alpha}}$	2012	Khalaf [95]
$k_{38}=max\left(\frac{1}{\sqrt{\frac{\lambda_{max}\hat{\sigma}^2}{(n-p)\hat{\sigma}^2+(\lambda_{max})\hat{\theta}_i^2}}}\right)$, $k_{39}=max\left(\sqrt{\frac{\lambda_{max}\hat{\sigma}^2}{(n-p)\hat{\sigma}^2+(\lambda_{max})\hat{\theta}_i^2}}\right)$, $k_{40}=\left(\prod_{i=1}^{p}\frac{1}{\sqrt{\frac{\lambda_{max}\hat{\sigma}^2}{(n-p)\hat{\sigma}^2+(\lambda_{max})\hat{\theta}_i^2}}}\right)^{1/p}$, $k_{41}=\left(\prod_{i=1}^{p}\sqrt{\frac{\lambda_{max}\hat{\sigma}^2}{(n-p)\hat{\sigma}^2+(\lambda_{max})\hat{\theta}_i^2}}\right)^{1/p}$, $k_{42}=median\left(\frac{1}{\sqrt{\frac{\lambda_{max}\hat{\sigma}^2}{(n-p)\hat{\sigma}^2+(\lambda_{max})\hat{\theta}_i^2}}}\right)$	2012	Muniz et al. [96]
$k_{43}=\frac{(\lambda_{max}+\lambda_{min})}{2\sum_{i=1}^{n}\lvert\hat{\beta}_i\rvert}\frac{p\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}}$, $k_{44}=\frac{(\lambda_{max}+\lambda_{min})}{2\sum_{i=1}^{n}\lvert\hat{\beta}_i\rvert}\frac{\lambda_{max}\hat{\sigma}^2}{\left((n-p-1)\hat{\sigma}^2+\lambda_{max}\hat{\beta}_{max}^2\right)}$	2013	Khalaf [97]
$k_{45}=\frac{2}{p\lambda_{max}}\sum_{i=1}^{p}\frac{\hat{\sigma}^2}{\hat{\theta}_i^2}$, $k_{46}=median\left(\frac{2\hat{\sigma}^2}{\lambda_{max}\hat{\theta}_i^2}\right)$, $k_{47}=\frac{2\hat{\sigma}^2}{\lambda_{max}(\prod_{i=1}^{p}\hat{\theta}_i^2)^{1/p}}$, $k_{48}=\frac{2p}{\lambda_{max}}\sum_{i=1}^{p}\frac{\hat{\sigma}^2}{\hat{\theta}_i^2}$	2014	Dorugade [98]

Table 3.1 (Continue). Some of the proposed formulas for finding the optimal bias parameters for Ridge regression.

$k_{49}=\frac{p^2}{\lambda_{max}^2}\frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\theta}_i^2}$, $k_{50}=\frac{p^3}{\lambda_{max}^3}\frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\theta}_i^2}$, $k_{51}=\frac{p}{\lambda_{max}^{1/3}}\frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\theta}_i^2}$, $k_{52}=\frac{p}{\sqrt[3]{(\prod_{i=1}^{p}\sqrt{\lambda_i})}}\frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\theta}_i^2}$, $k_{53}=\frac{2p}{\sqrt{\lambda_{max}}}\frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\theta}_i^2}$	2014	Asar et al. [99]
$k_{54}=\frac{\hat{\sigma}^2}{min\left(\frac{\hat{\sigma}^2}{\hat{\theta}_i^2}+\frac{1}{\lambda_i}\right)}$	2014	Dorugade [100]
$k_{55}=\frac{\hat{\sigma}^2}{2}\left(\frac{1}{\hat{\beta}_{max}^2}+\frac{p}{\hat{\beta}'\hat{\beta}}\right)$, $k_{56}=\frac{2\hat{\sigma}^2}{\hat{\beta}_{max}^2+\frac{\hat{\beta}'\hat{\beta}}{p}}$, $k_{57}=\hat{\sigma}^2\sqrt{\frac{p}{\hat{\beta}_{max}^2\hat{\beta}'\hat{\beta}}}$	2014	Khalaf and Iguernane [101]
$k_{58}=\sqrt{median(m_i)}$, $k_{59}=\frac{\hat{\sigma}^2}{\left(median(\hat{\theta}_i)\right)^2}$	2016	Göktaş and Sevinç [102]
$k_{60}=\frac{p\hat{\sigma}^2}{\hat{\theta}'\hat{\theta}}+\frac{1}{\lambda_{max}\hat{\theta}'\hat{\theta}}$, $k_{61}=\frac{p\hat{\sigma}^2}{\hat{\theta}'\hat{\theta}}+\frac{1}{2\left(\frac{\lambda_{max}}{\lambda_{min}}\right)^2}$	2016	Bhat and Vidya [103]
$k_{62}=\hat{\sigma}$	2016	Dorugade [104]
$k_{63}=\sqrt{\frac{\lambda_{max}\hat{\sigma}^2}{(n-p-1)\hat{\sigma}^2+\lambda_{max}\hat{\beta}_{max}^2}}$	2016	Khalaf and Iguernane [105]
$k_{64}=\frac{\sqrt{5}p}{\lambda_{max}}\frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\theta}_i^2}$, $k_{65}=\frac{p}{\sqrt{\lambda_{max}}}\frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\theta}_i^2}$, $k_{66}=\frac{2p}{\sum_{i=1}^{p}\left(\lambda_i^{1/4}\right)}\frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\theta}_i^2}$, $k_{67}=\frac{2p}{\sqrt{\sum_{i=1}^{p}\lambda_i}}\frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\theta}_i^2}$	2016	Karaibrahimoğlu et al. [106]
$k_{68}=p\sigma$	2017	Lukman and Olatunji [107]
$k_{69}=\max\left(\sqrt{\frac{2\hat{\sigma}^2}{\lambda_{max}\hat{\theta}_i^2}}\right)$, $k_{70}=\sqrt{\frac{2\hat{\sigma}^2}{\lambda_{max}p}}\sum_{i=1}^{p}\frac{1}{\hat{\theta}_i^2}$, $k_{71}=\frac{2\hat{\sigma}^2}{\sqrt{\sum_{i=1}^{p}\lambda_i}\max(\hat{\theta}_i^2)}$, $k_{72}=\sqrt{\frac{2\hat{\sigma}^2}{\sqrt{\sum_{i=1}^{p}\lambda_i}\max(\hat{\theta}_i^2)}}$, $k_{73}=\frac{2p\hat{\sigma}^2}{\sqrt{\sum_{i=1}^{p}\lambda_i}\sum_{i=1}^{p}\hat{\theta}_i^2}$, $k_{74}=\sqrt{\frac{2p\hat{\sigma}^2}{\sqrt{\sum_{i=1}^{p}\lambda_i}\sum_{i=1}^{p}\hat{\theta}_i^2}}$, $k_{75}=\frac{2\hat{\sigma}^2}{\sqrt{\sum_{i=1}^{p}\lambda_i}(\prod_{i=1}^{p}\hat{\theta}_i^2)^{1/p}}$, $k_{76}=\sqrt{\frac{2\hat{\sigma}^2}{\sqrt{\sum_{i=1}^{p}\lambda_i}(\prod_{i=1}^{p}\hat{\theta}_i^2)^{1/p}}}$, $k_{77}=median\left(\frac{2\hat{\sigma}^2}{\sqrt{\sum_{i=1}^{p}\lambda_i}\hat{\theta}_i^2}\right)$	2017	Lukman et al. [108]

The Ridge method, like any statistical technique, has its own set of advantages and disadvantages, which are detailed below [109].

Advantages of the Ridge method:

- Reduces coefficient size, particularly effective for mitigating multicollinearity.
- Simplifies model complexity through coefficient shrinkage.

- Offers flexibility with loss functions and regularization techniques.
- Addresses empirical indeterminacy in multidimensional datasets.
- Includes a penalty term to help prevent overfitting.
- Guarantees a solution by incorporating the penalty term.
- Reduces the likelihood of overfitting to training data.
- Does not require estimators to be unbiased.
- Introduces an appropriate bias to the data, improving the reliability of predictions.
- Performs well with multivariate data, especially when the number of predictors exceeds the number of observations.
- Is the optimal choice when seeking to improve OLS estimates under multicollinearity.
- Uses biased estimates to reduce variance in the presence of multicollinearity.
- Starts with a slightly worse fit on the training set but improves long-term predictive performance.
- Delivers the best results when most variables contribute meaningfully to the model.
- Excels when handling multiple medium-to-large impact variables.

Disadvantages of the Ridge method:

- It can be computationally intensive.
- Retains all predictors in the final model.
- Does not support feature selection.
- Shrinks coefficients towards zero.
- Alters variance by introducing bias.
- Increases bias in the model.
- Tends to reduce model interpretability.

3.4. Liu Regression

Liu [13] proposed an alternative methodology to Ridge regression for addressing the issue of multicollinearity. This method demonstrated superior performance compared to the OLS estimator in terms of MSE. A notable feature of the Liu method is the incorporation of a biasing parameter, d, which enhances its flexibility. This characteristic enables the Liu estimator to effectively manage multicollinearity by achieving an improved trade-off between bias and variance, thereby yielding more accurate coefficient estimates in various scenarios [50].

The Liu estimator is defined in Eq. 3.19:

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_L &= (\boldsymbol{X}'\boldsymbol{X}+\boldsymbol{I})^{-1}\left(\boldsymbol{X}'\boldsymbol{Y}+d\widehat{\boldsymbol{\beta}}_{OLS}\right)\\ &= (\boldsymbol{X}'\boldsymbol{X}+\boldsymbol{I})^{-1}(\boldsymbol{X}'\boldsymbol{X}+d\boldsymbol{I})\widehat{\boldsymbol{\beta}}_{OLS}, \quad d\in[0,1]\end{aligned} \tag{3.19}$$

Here, d represents the bias parameter of the Liu estimator. Notably, when $d=1$, the Liu estimator $\widehat{\boldsymbol{\beta}}_L$ becomes equivalent to the OLS estimator.

As shown in Eq. 3.20, the Liu estimator can be expressed as a linear transformation of the OLS estimator:

$$\widehat{\boldsymbol{\beta}}_L = (\boldsymbol{X}'\boldsymbol{X}+\boldsymbol{I})^{-1}(\boldsymbol{X}'\boldsymbol{X}+d\boldsymbol{I})\widehat{\boldsymbol{\beta}}_{OLS} = \boldsymbol{Z}_d\widehat{\boldsymbol{\beta}}_{OLS} \tag{3.20}$$

where $\boldsymbol{Z}_d = (\boldsymbol{X}'\boldsymbol{X}+\boldsymbol{I})^{-1}(\boldsymbol{X}'\boldsymbol{X}+d\boldsymbol{I})$. The expected value of the Liu estimator is given by Eq. 3.21:

$$E\left(\widehat{\boldsymbol{\beta}}_L\right) = (\boldsymbol{X}'\boldsymbol{X}+d\boldsymbol{I})\boldsymbol{\beta} \tag{3.21}$$

Since $E\left(\widehat{\boldsymbol{\beta}}_L\right) = E\left(\boldsymbol{Z}_d\widehat{\boldsymbol{\beta}}_{OLS}\right) \neq \boldsymbol{\beta}$, it follows that the Liu regression estimator $\widehat{\boldsymbol{\beta}}_L$ is a biased estimator of the coefficient $\boldsymbol{\beta}$. Additionally, the variance-covariance matrix of the Liu estimator is expressed in Eq. 3.22:

$$Var\left(\widehat{\boldsymbol{\beta}}_L\right) = \sigma^2(\boldsymbol{X}'\boldsymbol{X}+\boldsymbol{I})^{-1}(\boldsymbol{X}'\boldsymbol{X}+d\boldsymbol{I})(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{X}'\boldsymbol{X}+d\boldsymbol{I})(\boldsymbol{X}'\boldsymbol{X}+\boldsymbol{I})^{-1} \tag{3.22}$$

It has been demonstrated that the Liu estimator outperforms the OLS estimator based on the MSE criterion [13]. Accordingly, for $\boldsymbol{\alpha} = \boldsymbol{D}'\boldsymbol{\beta}$, the MSE value of the Liu estimator is given in Eq. 3.23:

$$MSE(\widehat{\boldsymbol{\beta}}_L) = \sigma^2 \sum_{i=1}^{p} \frac{(\lambda_i+d)^2}{\lambda_i(\lambda_i+1)^2} + (d-1)^2 \sum_{i=1}^{p} \frac{\theta_i^2}{(\lambda_i+1)^2} \tag{3.23}$$

Where $\lambda_1, \lambda_2, \dots, \lambda_p$ denote the eigenvalues of the $\boldsymbol{X}'\boldsymbol{X}$ matrix. The primary goal in Liu estimation is to select a bias parameter d such that the reduction in variance outweighs the corresponding increase in bias. Liu [13] demonstrated the existence of a d value for which $(\widehat{\boldsymbol{\beta}}_L) < MSE(\widehat{\boldsymbol{\beta}}_{OLS})$.

The VIF for the Liu estimator is presented in Eq. 3.24:

$$VIF_p = diag\left[(\boldsymbol{X}'\boldsymbol{X}+\boldsymbol{I})^{-1}(\boldsymbol{X}'\boldsymbol{X}+d\boldsymbol{I})(\boldsymbol{X}'\boldsymbol{X})^{-1}\left((\boldsymbol{X}'\boldsymbol{X}+d\boldsymbol{I})(\boldsymbol{X}'\boldsymbol{X}+\boldsymbol{I})^{-1}\right)'\right] \tag{3.24}$$

Since the introduction of the Liu estimator [13], various researchers have proposed methods for determining the bias parameter d in Liu regression. Table 3.2 lists 13 commonly used estimators, with $q_j = \frac{\theta_j^2 - 1}{\max\left(\frac{1}{\lambda_j}\right) + \theta_j^2}$ and $\boldsymbol{\theta} = \boldsymbol{D}'\boldsymbol{\beta}$, $i = 1,2, \dots, p$.

Table 3.2. Some of the proposed formulas for finding the optimal bias parameters for Liu regression.

Proposed Bias Parameter	Year	Authors
$d_1 = \max\left(0, \frac{\hat{\theta}^2_{max} - \hat{\sigma}^2}{\frac{1}{\lambda_{max}} + \hat{\theta}^2_{max}}\right), d_2 = \max\left(0, median\left(\frac{\hat{\theta}^2_j - \hat{\sigma}^2}{\frac{1}{\lambda_j} + \hat{\theta}^2_j}\right)\right)$ $d_3 = \max\left(0, mean\left(\frac{\hat{\theta}^2_j - \hat{\sigma}^2}{\frac{1}{\lambda_j} + \hat{\theta}^2_j}\right)\right), d_4 = \max\left(0, max\left(\frac{\hat{\theta}^2_j - \hat{\sigma}^2}{\frac{1}{\lambda_j} + \hat{\theta}^2_j}\right)\right)$ $d_5 = \max\left(0, min\left(\frac{\hat{\theta}^2_j - \hat{\sigma}^2}{\frac{1}{\lambda_j} + \hat{\theta}^2_j}\right)\right), d_6 = \max\left(0, median(q_i)\right)$ $d_7 = \max\left(0, mean(q_i)\right), d_8 = \max\left(0, max(q_i)\right), \ d_9 = \max\left(0, min(q_i)\right)$	2015	Shukur et al. [110]
$d_{10} = \max\left(0, median\left(\frac{\hat{\theta}^2_j - \hat{\sigma}^2}{\max\left(\frac{\hat{\sigma}^2}{\hat{\theta}^2_j}\right) + \hat{\theta}^2_{max}}\right)\right)$ $d_{11} = \max\left(0, mean\left(\frac{\hat{\theta}^2_j - \hat{\sigma}^2}{\max\left(\frac{\hat{\sigma}^2}{\hat{\theta}^2_j}\right) + \hat{\theta}^2_{max}}\right)\right)$ $d_{12} = \max\left(0, max\left(\frac{\hat{\theta}^2_j - \hat{\sigma}^2}{\max\left(\frac{\hat{\sigma}^2}{\hat{\theta}^2_j}\right) + \hat{\theta}^2_{max}}\right)\right), d_{13} = \max\left(0, min\left(\frac{\hat{\theta}^2_j - \hat{\sigma}^2}{\max\left(\frac{\hat{\sigma}^2}{\hat{\theta}^2_j}\right) + \hat{\theta}^2_{max}}\right)\right)$	2020	Qasim et al. [111]

3.5. Fuzzy Regression

Regression analysis serves as a cornerstone in data modeling, facilitating the examination of relationships between a dependent variable and one or more independent variables by expressing the former as a function of the latter. While conventional regression methods are designed to manage randomness in data, fuzzy regression extends this capability to address ambiguity and imprecision inherent in complex systems or resulting from human estimation errors. This approach proves particularly effective in capturing uncertainty in scenarios characterized by vague data or imprecisely defined systems, thereby enabling more flexible modeling and enhancing interpretability.

In traditional regression analysis, discrepancies between observed values and their corresponding estimates are typically attributed to random errors, with statistical techniques

employed for coefficient estimation and inference. Nonetheless, such deviations may also stem from the indeterminate structure of the system or the presence of imprecise measurements. In these instances, the uncertainty within the regression model is better described by fuzziness rather than randomness [112].

The development of fuzzy set theory by Zadeh [4] provided a foundation for representing vague variables as numerical entities, enabling their quantitative evaluation. Building on this framework, Tanaka et al. [15] introduced a linear regression model in 1982 that integrated fuzzy set theory. In this model, qualitative variables or observations are represented alongside their corresponding fuzzy membership degrees, facilitating the incorporation of imprecision into the analysis [53].

Fuzzy regression analysis evaluates two primary trends: the trend of the center and the trend of the spread of the response. The center trend aligns closely with outcomes from standard regression, representing the overall behavior of the system. In contrast, the spread trend captures individual variations and reflects the inherent fuzziness in the system's structure. By analyzing the spread trend, valuable insights can be obtained regarding the consistency of responses across different input conditions, providing additional information beyond the scope of traditional regression methods. This approach effectively relaxes the standard regression assumption of constant variance, allowing for varying spreads in error terms [53, 113].

The development of fuzzy regression analysis has been driven by several critical factors [114, 115]:

First, traditional econometric analyses require a sufficiently large number of observations to ensure reliable results, a condition that may not always be achievable. In contrast, fuzzy regression methods can deliver dependable results even with limited sample sizes. Second, classical econometric approaches rely on specific assumptions, such as the normality of error terms and the absence of autocorrelation. These assumptions may not always be held in practice, but fuzzy regression techniques mitigate the need to strictly adhere to these constraints. Furthermore, classical econometric analyses often necessitate data transformations (e.g., logarithmic transformations) to uncover relationships among

variables. Fuzzy regression, however, enables the use of raw-level data, retaining more information within the analysis.

Additionally, traditional econometric methods face challenges when variables cannot be expressed numerically, such as in cases involving rankings, ratings, or relative measures. Under such conditions, classical regression techniques become impractical, whereas fuzzy regression provides an effective alternative. Finally, articulating relationships in economics through precise functions or coefficients is not always feasible. Fuzzy regression techniques offer a means to explore these relationships, allowing for a broader and more nuanced understanding.

Moreover, several distinctions exist between fuzzy regression and classical regression methods. Fuzzy regression integrates both fuzzy and crisp data within a single model, whereas ordinary regression is confined to crisp data. While classical regression relies on assumptions about data, fuzzy regression is more flexible and can function without adhering to these assumptions, depending on the dataset structure.

Another key difference lies in the treatment of error terms. In classical regression, the unobserved error term must be independently distributed, and deviations from this requirement may compromise the effectiveness of the method. Fuzzy regression, in contrast, accommodates such deviations, offering a more adaptable approach. Unlike classical regression, which is grounded solely in probability theory, fuzzy regression leverages both probability theory and fuzzy set theory. This dual foundation allows fuzzy regression to handle data characterized by various types of uncertainty, including random and non-random sources, which classical regression cannot.

Finally, in classical regression, deviations between the regression model and observed data are treated as random observation errors, often regarded as a random variable. Fuzzy regression, however, interprets these deviations as the inherent fuzziness of the model itself, avoiding classification as errors and offering a distinct perspective on the relationship between the data and the model [115-117].

3.6. Fuzzy Least Squares

Fuzzy regression can be classified based on both estimation approaches and the types of data involved. Regarding estimation methods, it encompasses possibilistic regression, which includes techniques such as linear programming, goal programming, and interval regression, as well as the fuzzy least squares method, which, although distinct, complements the fuzzy regression framework. Additionally, fuzzy regression can be categorized according to the nature of the data, leading to four main classes: crisp input and crisp output, crisp input and fuzzy output, fuzzy input crisp output and fuzzy input and fuzzy output.

Fuzzy regression can also be distinguished based on the nature of the relationships between variables, with two primary categories: linear fuzzy regression and nonlinear fuzzy regression. Another important classification considers the application context, differentiating between descriptive fuzzy regression, which aims to interpret relationships, and predictive fuzzy regression, which focuses on forecasting outcomes. For more comprehensive discussions on these classifications, see Zadeh [6] on linguistic variables, Tanaka et al. [15] on fuzzy models in regression, and Bardossy and Fodor [118] on fuzzy regression analysis.

Prior to exploring the fuzzy least squares method, the following definition and properties will be presented [35]:

3.6.1 Definition: Let L and R be decreasing functions from $\mathbb{R}^+$ to [0,1], where $L(0) = R(0) = 1$. For any real number m, the membership function of the LR-type fuzzy number $\tilde{A}$ is defined by Eq. 3.25, as follows:

$$\mu_{\tilde{A}}(x) = \begin{cases} L\left(\frac{m-x}{a}\right) & ; \quad x \leq m, a > 0 \\ R\left(\frac{x-m}{b}\right) & ; \quad x \geq m, b > 0 \\ 0 & ; \quad otherwise \end{cases} \tag{3.25}$$

Here, $\tilde{A}$ is called an LR-type fuzzy number, where m represents the center of $\tilde{A}$, and a and b represent the left and right spreads, respectively. This is denoted as $\tilde{A} = (m, a, b)_{LR}$. When $a = b = 0$, the LR-type fuzzy number reduces to a conventional real number, represented as $(m, 0,0)$.

If $a = b$, the LR-type fuzzy number is symmetric and can be expressed as $\tilde{A} = (m, a)_{LR}$. The set of all LR-type fuzzy numbers is denoted by $\tilde{F}_{LR}(\mathbb{R})$.

In particular, if

$$\mu_{\tilde{A}}(x) = \begin{cases} 1 - \frac{m-x}{a} & ; \quad m - a \le x \le m \\ 1 - \frac{x-m}{b} & ; \quad m \le x \le m + b \\ 0 & ; \quad otherwise \end{cases} \tag{3.26}$$

where $L(x) = R(x) = 1 - x$, then the fuzzy number $\tilde{A}$, illustrated in Eq. 3.26, is called a triangular fuzzy number.

These fuzzy numbers possess the following properties. Let $\tilde{A} = (a_m, a_l, a_r)_{LR}, \tilde{B} = (b_m, b_l, b_r)_{LR} \in \tilde{F}_{LR}(\mathbb{R})$, and suppose that $\tilde{A}$ and $\tilde{B}$ share the same L and R functions. Then the following properties hold:

- $-\tilde{A} = (-a_m, a_l, a_r)_{LR}$
- $\tilde{A} + \tilde{B} = (a_m + b_m, a_l + b_l, a_r + b_r)_{LR}$
- $\tilde{A} - \tilde{B} = (a_m - b_m, a_l + b_l, a_r + b_r)_{LR}$
- $t\tilde{A} = (ta_m, |t|a_l, |t|a_r)_{LR}, t \in R$
- $\tilde{A}\tilde{B} = (a_m b_m, |a_m|b_l + |b_m|a_l, |a_m|b_r + |b_m|a_r)_{LR}$

The fuzzy least squares method is a fuzzy extension of the OLS estimation method, initially introduced by Celminš [17] and Diamond [18], with its core principle being the minimization of the overall squared error between observed and estimated values. Since its inception, fuzzy least squares has undergone numerous revisions and extensions, establishing itself as a pivotal method in fuzzy regression analysis. Researchers have significantly contributed to its development and applications across various domains. For example, Chang and Ayyub [116] conducted a comprehensive comparative assessment of fuzzy regression methods and introduced a hybrid fuzzy least squares regression analysis, while Hong et al. [119] enhanced model fit by employing shape-preserving operations in fuzzy least squares linear regression. Mohammadi and Taheri [120] demonstrated the effectiveness of a specific fuzzy least squares method in soil science, using fuzzy numbers to represent the response variable.

Coppi et al. [121] developed a model analyzing the dependence of an LR-type fuzzy response variable on crisp independent variables and proposed an iterative estimation procedure. Further advancements include D'Urso and Santoro's [122] development of goodness-of-fit indices and variable selection methods tailored for fuzzy multiple regression models, as well as Arabpour and Tata's [123] improved coefficients estimation techniques for fuzzy linear regression. Kim et al. [124] explored the asymptotic properties of least squares estimators for fuzzy regression models with fuzzy inputs, outputs, and random error terms, while Chen and Hsueh [125] utilized distance metrics to construct fuzzy regression models via the least squares method. Other noteworthy contributions include Diamond and Körner's [126] extension of fuzzy linear models and least squares techniques, Kao and Chyu's [127] analysis of least squares estimators in fuzzy regression, and Xu and Li's [40] study of fuzzy multidimensional regression models using the least squares fitting approach.

3.6.2 Definition: Let $\tilde{A} = (a_m, a_l, a_r)_{LR}$ and $\tilde{B} = (b_m, b_l, b_r)_{LR}$ be two fuzzy numbers. The Euclidean distance between them is defined by Eq. 3.27, as follows [35, 123]:

$$D_E^2(\tilde{A}, \tilde{B}) = (a_m - b_m)^2\omega_m + (a_l - b_l)^2\omega_l + (a_r - b_r)^2\omega_r \tag{3.27}$$

where $\omega_m > 0, \omega_l > 0, \omega_r > 0$ are arbitrary weights.

In particular, if $\tilde{A} = (a, \alpha)_{LR}$ and $\tilde{B} = (b, \gamma)_{LR}$ are symmetric fuzzy numbers, the distance in Eq. 3.27 simplifies to Eq. 3.28:

$$D_E^2(\tilde{A}, \tilde{B}) = (a - b)^2 + (\alpha - \gamma)^2 \tag{3.28}$$

From the properties of Euclidean distance, it can be shown that it is non-negative, symmetric, and satisfies the triangle inequality, making the fuzzy number space a complete metric space.

Based on previous definitions and the findings of Zeng and Zhang [35] attention now turns to fuzzy least squares method using the Euclidean distance [41].

The fuzzy linear regression model is expressed by Eq. 3.29, as follows:

$$\tilde{y}_i = \tilde{A}_0 + \tilde{A}_1 x_{i1} + \tilde{A}_2 x_{i2} + \cdots + \tilde{A}_p x_{ip} \qquad i = 1,2,\ldots,n \tag{3.29}$$

where x_{ij} represent crisp values, and $\tilde{y}_i (i = 1,2,\ldots,n)$ and $\tilde{A}_j (j = 0,1,\ldots,p)$ are fuzzy numbers of LR-type. Without loss of generality, it is assumed that $x_{ij} > 0$; for cases where $x_{ij} < 0$, the model can be equivalently transformed to ensure $x_{ij} > 0$. Furthermore, to simplify the analysis, both $\tilde{y}_i$ and $\tilde{A}_j$ LR-type fuzzy numbers.

For the fuzzy least squares method, the traditional OLS method is applied to estimate the regression coefficients $\tilde{A}_j$ in the previous model. Suppose that $\tilde{y}_i = (c_i, s_i)_{LR}$ and $\tilde{A}_j = \left(a_j, r_j\right)_{LR}$ have the same membership functions. The model in Eq. 3.29 can then be expressed as in Eq. 3.30 as follows:

$$(c_i, s_i) = (a_0, r_0) + (a_1, r_1)x_{i1} + \cdots + \left(a_p, r_p\right)x_{ip} \tag{3.30}$$

Using the Euclidean distance, the sum of squared errors can be written as shown in Eq. 3.31:

$$D_E^2 = \sum_{i=1}^{n} \left[\left(c_i - \left(a_0 + a_1 x_{i1} + \cdots + a_p x_{ip} \right) \right)^2 + \left(s_i - \left(r_0 + r_1 x_{i1} + \cdots + r_p x_{ip} \right) \right)^2 \right] \tag{3.31}$$

Let: $a = \left(a_0, a_1, \ldots, a_p\right)^T, r = \left(r_0, r_1, \ldots, r_p\right)^T, c = (c_1, c_2, \ldots, c_n)^T, s = (s_1, s_2, \ldots, s_n)^T$ and $X = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix}$. Then the error can be minimized as follows:

$D_E^2 = (Xa - c)^T (Xa - c) + (Xr - s)^T (Xr - s)$, and

$$\begin{cases} \frac{\partial D_E^2}{\partial a} = 2X^T Xa - 2X^T c = 0 \\ \frac{\partial D_E^2}{\partial r} = 2X^T Xr - 2X^T s = 0 \end{cases} \Rightarrow \begin{cases} \hat{a} = (X^T X)^{-1} X^T c \\ \hat{r} = (X^T X)^{-1} X^T s \end{cases} \tag{3.32}$$

For the error term in the estimation using Euclidean distance, the simplest method to evaluate the performance of a fuzzy regression model is by using the residual as a performance index.

Denote the residual as $\varepsilon_i = \tilde{y}_i - \hat{\tilde{y}}_i$ or $\varepsilon_i^2 = \left|\tilde{y}_i - \hat{\tilde{y}}_i\right|^2$ and the sum of squared errors $SST_E = \sum_{i=1}^{n} \varepsilon_i^2$, where $\hat{\tilde{y}}_i$ represents the estimated value. Using Euclidean distance, the SST can be expressed in Eq. 3.33, as follows:

$$SST_E = \sum_{i=1}^{n} D_E^2\left(\tilde{y}_i - \hat{\tilde{y}}_i\right) = (X\hat{a} - c)^T(X\hat{a} - c) + (X\hat{r} - s)^T(X\hat{r} - s) \tag{3.33}$$

3.7. α-Cut Based Fuzzy Regression

Multicollinearity, characterized by a high degree of linear dependence among explanatory variables, poses a significant challenge in classical linear regression models, often rendering the OLS estimator undefined or highly variable. Ridge regression addresses this issue by shrinking the coefficients of less influential variables toward zero, thereby stabilizing the variance of the OLS estimator. Similarly, multicollinearity frequently arises in fuzzy linear regression models, complicating the estimation process. Given the limited number of covariates in fuzzy data, removing explanatory variables is often impractical. To address this issue, several studies have developed fuzzy Ridge estimators using methods based on the distance between fuzzy numbers [38, 42, 128], though the applicability of these approaches remains limited [49].

Kim and Jung [49] introduced α-cut-based estimation method using Ridge and OLS methods as an extension of the classical Ridge and OLS estimation methods initially developed by Choi et al. [48]. This method utilizes the α-cut, a fundamental concept in fuzzy set theory, to enhance the estimation of fuzzy coefficients. By employing the α-cut approach, the model evaluates multiple sections of the fuzzy set, rather than limiting the analysis to its lower and upper bounds. This flexibility enables the estimation process to accommodate a broader range of potential values for the dependent variables, which are typically represented as triangular or other forms of fuzzy numbers. The α-cut-based estimation method is summarized as follows:

Consider a set of observations $\{\tilde{y}_i, \tilde{x}_{i1}, \tilde{x}_{i2}, \ldots, \tilde{x}_{ip}\}_{i=1}^{N} = \{(y_{il}, y_{im}, y_{ir})_T, (x_{1l}, x_{1m}, x_{1r})_T, \ldots, (x_{p1}, x_{pm}, x_{pr})_T\}_{i=1}^{N}$, where the dependent variable and the independent are triangular fuzzy

numbers, assuming a linear relationship between the dependent and independent variables, as expressed in the model in Eq. 3.34:

$$\tilde{y}_i = \tilde{A}_0 \oplus \tilde{A}_1 \odot \tilde{x}_{i1} \oplus \cdots \oplus \tilde{A}_p \odot \tilde{x}_{ip} \oplus \tilde{\varepsilon}_i, \qquad i = 1, \ldots, N \tag{3.34}$$

Here, $\{\tilde{A}_j\}_{j=0}^{p} = \left\{\left(A_{jl}, A_{jm}, A_{jr}\right)_T\right\}_{j=0}^{p}$ are the fuzzy regression coefficients, and $\{\tilde{\varepsilon}_i\}_{i=1}^{N} = \{(\varepsilon_{il}, \varepsilon_{im}, \varepsilon_{ir})_T\}_{i=1}^{N}$ are the fuzzy error terms. The operations $\oplus$ and $\odot$ denote addition and multiplication between fuzzy numbers, respectively. These N equations are often written in matrix form as in Eq. 3.35:

$$\tilde{\boldsymbol{Y}} = \tilde{\boldsymbol{X}} \odot \tilde{\boldsymbol{A}} \oplus \tilde{\boldsymbol{\varepsilon}} \tag{3.35}$$

For a comprehensive discussion on arithmetic operations involving fuzzy numbers, refer to [15, 63, 129]. It is important to note that fuzzy variables can be either symmetric or asymmetric. Given that crisp sets are a specific instance of fuzzy sets, it is also possible to combine fuzzy inputs with either fuzzy or crisp outputs.

This section illustrates the α-cut-based fuzzy estimation method using triangular membership functions. The data is divided into training and test sets, with the training set denoted as $\{\tilde{y}_i, \tilde{x}_{i1}, \tilde{x}_{i2}, \ldots, \tilde{x}_{ip}\}_{i=1}^{n}$ and the test set as $\{\tilde{y}_i, \tilde{x}_{i1}, \tilde{x}_{i2}, \ldots, \tilde{x}_{ip}\}_{i=1}^{m}$, where N represents the total number of observations, n is the number of observations in the training set, and m is the number of observations in the test set, such that $n + m = N$. The α-cut-based regression model is estimated using the training data, and its performance is evaluated on the test set. The estimation process consists of 6 steps, five of which are the core steps introduced by Choi et al. [48] and Kim and Jung [49].

Step 1: Create α-cuts sets for the triangular fuzzy input and output, as illustrated in Figure 3.1. For any α-cut in the range [0,1], the following applies to the dependent and independent variables:

The dependent variable can be expressed using the shape shown in Eq. 3.36, as follows:

$$\tilde{y}_i(\alpha) = \left[l_{y_i}(\alpha), r_{y_i}(\alpha)\right] = \left[(\alpha-1)s_{il}^{y} + y_{im}, -(\alpha-1)s_{ir}^{y} + y_{im}\right] \quad (3.36)$$

Similarly, the independent variable can be expressed using Eq. 3.37:

$$\tilde{x}_{ij}(\alpha) = \left[l_{x_{ij}}(\alpha), r_{x_{ij}}(\alpha)\right] = \left[(\alpha-1)s_{ijl}^{x} + x_{ijm}, -(\alpha-1)s_{ijr}^{x} + x_{ijm}\right] \quad (3.37)$$

for $i = 1, \dots, n$ and $j = 1, \dots, p$.

Here, $s_{il}^{y}, s_{ir}^{y}, s_{ijl}^{x}$ and $s_{ijr}^{x} \geq 0$ represent the left and right spreads of the dependent and independent variables, respectively. The α-cuts are represented by the sequence $(\alpha_q)_{q=0}^{Q}$ for some Q, where $\alpha_q \in [0,1]$.

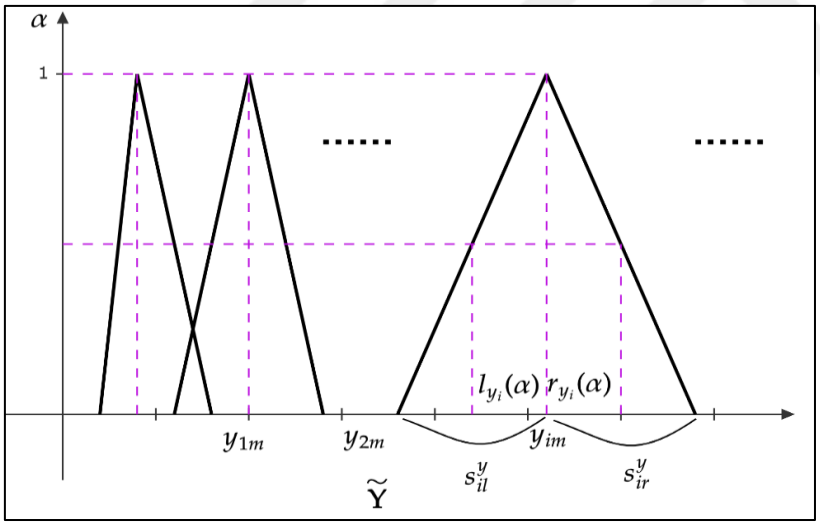


Figure 3.1. Fuzzy output data.

Step 2: Perform classical estimation methods for each crisp value obtained through the α-cut sequence, which were derived and organized in Step 1. Ridge and OLS methods were applied in [48, 49].

Here, assuming that the endpoints of the α-cut set of $\tilde{\boldsymbol{Y}}$ have been centered and the endpoints of the α-cut set of $\tilde{\boldsymbol{X}}$ have been standardized, following the convention in classical Ridge regression [82]. This step is explained and summarized in Figure 3.2.

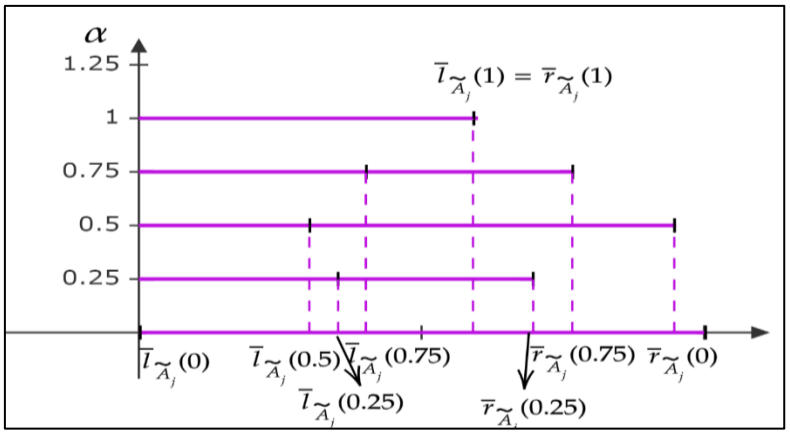


Figure 3.2. Intermediate estimators $l_{\tilde{A}}(\alpha_q)$ and $r_{\tilde{A}}(\alpha_q)$ for the α-cut sequence (0, 0.25, 0.5, 0.75, 1).

Step 3: Obtain the estimators $\ddot{l}_{\tilde{A}}(\alpha_q)$ and $\ddot{r}_{\tilde{A}}(\alpha_q)$ of $l_{\tilde{A}}(\alpha_q)$ and $r_{\tilde{A}}(\alpha_q)$ by modifying the intermediate estimators $\bar{l}_{\tilde{A}}(\alpha_q)$ and $\bar{r}_{\tilde{A}}(\alpha_q)$ to ensure that the estimated coefficients form the membership function of a triangular fuzzy number. The following operations are performed using Eq. 3.38-3.40, as follows (see Figure 3.3):

$$\ddot{l}_{\tilde{A}}(\alpha_q) = \ddot{r}_{\tilde{A}}(\alpha_q) = \bar{l}_{\tilde{A}}(\alpha_q) = \bar{r}_{\tilde{A}}(\alpha_q) \tag{3.38}$$

$$\ddot{l}_{\tilde{A}}(\alpha_q) = min\{\bar{l}_{\tilde{A}}(\alpha_q), \ddot{l}_{\tilde{A}}(\alpha_{q+1})\}, \quad q = Q-1, Q-2, \dots, 0 \tag{3.39}$$

$$\ddot{r}_{\tilde{A}}(\alpha_q) = max\{\bar{r}_{\tilde{A}}(\alpha_q), \ddot{r}_{\tilde{A}}(\alpha_{q+1})\}, \quad q = Q-1, Q-2, \dots, 0 \tag{3.40}$$

This modification ensures the estimators satisfy the properties of a triangular fuzzy number, where $\ddot{l}_{\tilde{A}}(\alpha_q)$ and $\ddot{r}_{\tilde{A}}(\alpha_q)$ represent the left and right endpoints, respectively, and both converge at α_Q.

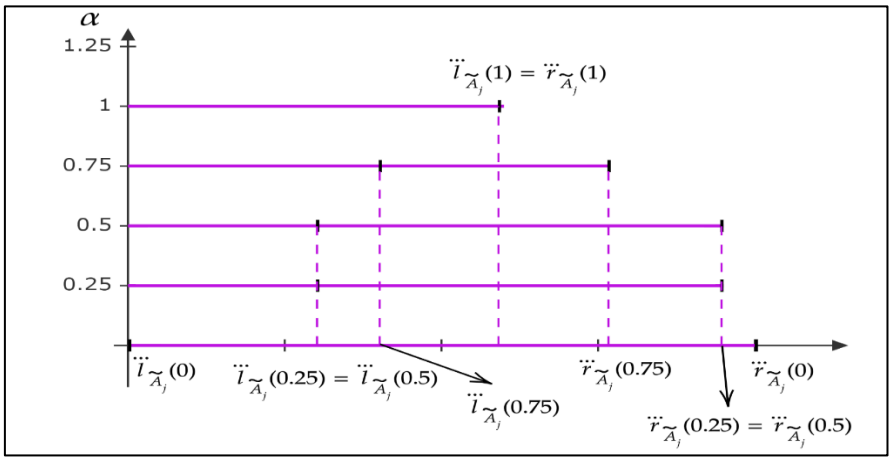


Figure 3.3. Modified estimators $\ddot{l}_{\tilde{A}}(\alpha_q) = \ddot{r}_{\tilde{A}}(\alpha_q)$ for the α-cut sequence (0, 0.25, 0.5, 0.75, 1).

Step 4: Estimate the triangular fuzzy coefficient $\hat{\tilde{A}} = (\hat{A}_l, \hat{A}_m, \hat{A}_r)_T$ and its membership function $\mu_{\hat{\tilde{A}}}$ by fitting a linear regression line through the points $\ddot{l}_{\tilde{A}}(\alpha_q)$ and $\ddot{r}_{\tilde{A}}(\alpha_q)$ for $q = 0, \dots, Q$, respectively.

To ensure the estimated fuzzy coefficient adheres to the properties of a triangular fuzzy number, the membership function $\mu_{\hat{\tilde{A}}}$ is constrained to satisfy Eq. 3.41:

$$\mu_{\hat{\tilde{A}}}\left(\hat{l}_{\tilde{A}}(1)\right) = \mu_{\hat{\tilde{A}}}\left(\hat{r}_{\tilde{A}}(1)\right) = 1 \tag{3.41}$$

This constraint ensures that the maximum membership is achieved at the endpoints of the triangular fuzzy number (see Figure 3.4). The triangular fuzzy coefficient $\hat{\tilde{A}}$ is characterized by its left endpoint $\hat{A}_l$, center $\hat{A}_m$, and right endpoint $\hat{A}_r$.

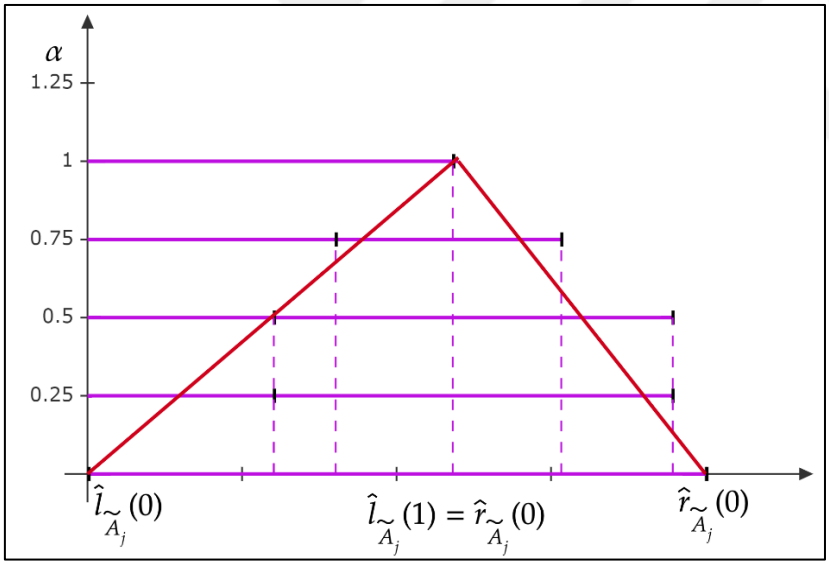


Figure 3.4. Estimated fuzzy coefficient $\hat{\tilde{A}} = (\hat{A}_l, \hat{A}_m, \hat{A}_r)_T$ and its membership function.

Step 5: Symmetric fuzzy inputs or outputs do not always guarantee that the estimated membership function $\mu_{\hat{\tilde{A}}}$ will also be symmetric. To reduce the difference between the true values and the fitted values, we consider the following candidates for $\mu_{\hat{\tilde{A}}}$:

$$\mu_{\hat{\tilde{A}}}(x) = \begin{cases} \frac{x - \hat{l}_{\tilde{A}}(0)}{\hat{l}_{\tilde{A}}(1) - \hat{l}_{\tilde{A}}(0)} & ; \quad x \le \hat{l}_{\tilde{A}}(1) = \hat{r}_{\tilde{A}}(1) \\ \frac{\hat{r}_{\tilde{A}}(0) - x}{\hat{r}_{\tilde{A}}(0) - \hat{r}_{\tilde{A}}(1)} & ; \quad x > \hat{l}_{\tilde{A}}(1) = \hat{r}_{\tilde{A}}(1) \end{cases} \quad for\ x \in R \tag{3.42}$$

Alternatively,

$$\mu_{\hat{\tilde{A}}}(x) = \begin{cases} \frac{x - \hat{l}_{\tilde{A}}(0)}{\hat{l}_{\tilde{A}}(1) - \hat{l}_{\tilde{A}}(0)} & ; \quad x \leq \hat{l}_{\tilde{A}}(1) \\ \frac{\hat{l}_{\tilde{A}}(0) + x}{\hat{l}_{\tilde{A}}(1) - \hat{l}_{\tilde{A}}(0)} & ; \quad x > \hat{l}_{\tilde{A}}(1) \end{cases} \qquad for\ x \in R \tag{3.43}$$

where $\hat{r}_{\tilde{A}}(0)$ is chosen as $\hat{r}_{\tilde{A}}(0) = 2 \times \hat{l}_{\tilde{A}}(1) - \hat{l}_{\tilde{A}}(0)$.

To evaluate the proposed fuzzy estimators, a performance criteria based on Diamond's [18] fuzzy distance measure is introduced. Let the dependent variable be $\tilde{y}_i = (y_{il}, y_{im}, y_{ir})_T$, for $i = 1, \dots, n$, and its predicted value as $\hat{\tilde{y}}_i = (\hat{y}_{il}, \hat{y}_{im}, \hat{y}_{ir})_T = \left(\boldsymbol{X}_{il}^t \widehat{\boldsymbol{A}}_l, \boldsymbol{X}_{im}^t \widehat{\boldsymbol{A}}_m, \boldsymbol{X}_{ir}^t \widehat{\boldsymbol{A}}_r\right)_T$. Here, n is the number of observations in the training set. The root mean square error for fuzzy numbers ($RMSE_F$) is defined as:

$$RMSE_F = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \{(y_{il} - \hat{y}_{il})^2 + (y_{im} - \hat{y}_{im})^2 + (y_{ir} - \hat{y}_{ir})^2\}} \tag{3.44}$$

To conclude, compute the $RMSE_F$ for each membership function and select the one that minimizes the criterion.

The previous steps can be simplified in the following flowchart in Figure 3.5:

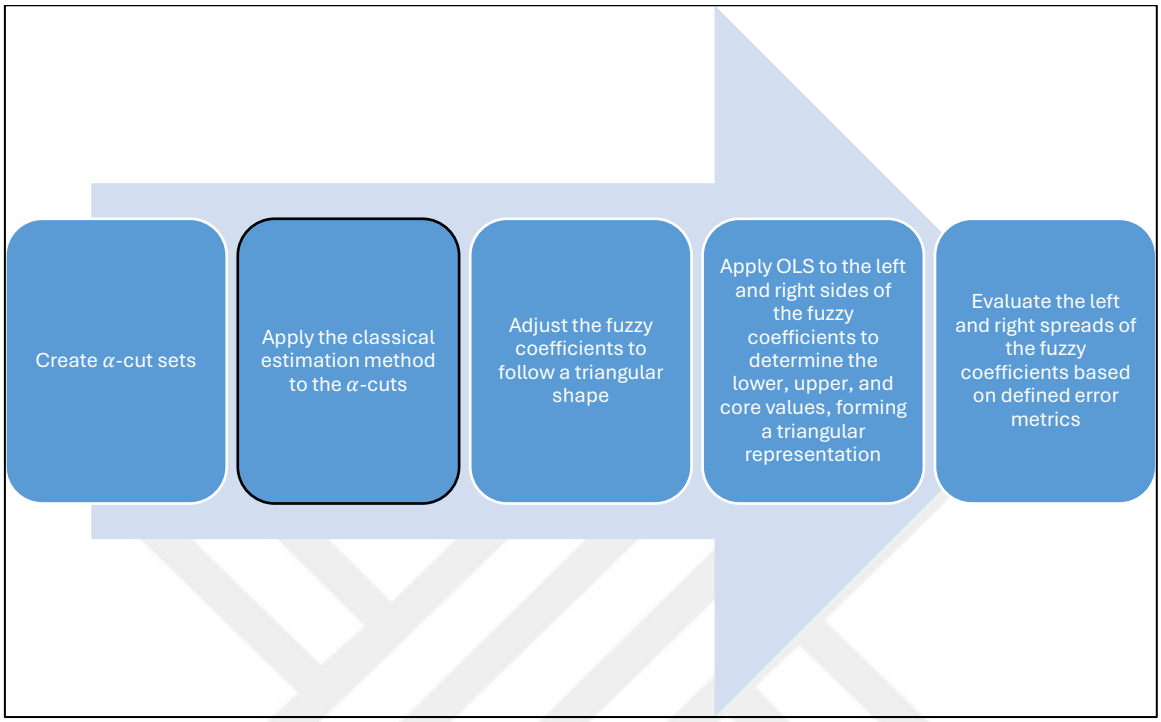


Figure 3.5. α-cut-based fuzzy estimation algorithm.

As a summary, the α-cut-based fuzzy estimation algorithm, as illustrated in Figure 3.5, can be summarized in six key steps. In the first step, the α-cuts of the fuzzy dependent variable are determined. This study considers datasets with crisp input variables and a fuzzy output, meaning that for each selected α-cut, the corresponding crisp values on the left and right sides of the fuzzy number are extracted. For instance, if $\alpha \in \{0,0.25,0.5,0.75,1\}$, this yields four values on the left, four on the right, and one at the center. In the second step, classical estimation methods—Ridge, Liu or OLS—are applied to the data. These methods are executed according to the selected α-cuts: four times on the left, four times on the right, and twice at the core (which produces the same result for the center). Depending on the configuration of α-cuts, the estimation process may require fewer or more iterations. The third step involves verifying whether the resulting coefficients satisfy the definition of a triangular fuzzy number, as defined in Eq. 2.6 and illustrated in Figure 2.3. If the triangular condition is not met, modifications are made as demonstrated in Figures 3.2 and 3.3. In the fourth step, the classical OLS method is employed to determine the lower and upper bounds as well as the core values of the fuzzy coefficients. The fifth step, although optional, is important for enhancing the accuracy and symmetry of the coefficients. Since the results of Step 4 may yield asymmetrical fuzzy coefficients, Step 5 allows for the examination and

adjustment of the left and right spreads separately to identify optimal values. Step 6, also considered optional, aims to further refine and optimize the final results and will be discussed in detail in the following step.

Step 6: Repeat Steps 1–5 for a chosen sequence of α-cuts, $\{\alpha_q\}_{q=0}^{Q}$, where each α_q is equally spaced between 0 and 1. After performing these steps for each α-cut, select the optimal set of α-cuts that minimizes the $RMSE_F$. Using this optimal sequence, compute the final fuzzy coefficient estimate, $\widehat{\widetilde{\boldsymbol{A}}}$, ensuring that it reflects the best fit based on the selected levels. This approach refines the fuzzy coefficient estimation by carefully balancing between accuracy and computational efficiency, ensuring that the selected α-cuts yield the best possible outcome.

3.8. Hybrid α-Cut-Based Estimation Framework

This research extends the α-cut-based fuzzy regression method originally introduced by Choi et al. [48] and later applied by Kim and Jung [49], by addressing several key limitations and introducing novel components to the framework. While previous studies focused only on a single sample size and compared fuzzy Ridge and fuzzy OLS methods, this thesis introduces a broader, more comprehensive evaluation of the α-cut-based approach. By analyzing multiple sample sizes, this work investigates the sensitivity of the algorithm's performance to data volume—an important aspect not considered in earlier studies.

A central contribution of this research is the systematic treatment of biased parameter selection in fuzzy Ridge regression. Previous works applied Ridge within the fuzzy context but did not explore or explain how to select the biased parameter k, which is critical for the estimator's performance. This thesis fills that gap through a detailed survey and implementation of various techniques for determining k, leading to more efficient and reliable results without relying on computationally expensive methods like K-fold cross-validation.

Importantly, this study is the first to incorporate the fuzzy Liu estimator within the α-cut-based fuzzy regression framework. Unlike fuzzy Ridge, which was already used in prior works, Liu regression had not yet been applied in this context. Here, it is introduced not as

a replacement but as a novel modeling alternative that brings a different biased structure, potentially offering better performance in certain conditions. Various methods for selecting the fuzzy Liu biasing parameter d are tested, and a new hybrid modeling framework is proposed to facilitate fair comparisons between fuzzy Ridge, fuzzy Liu, and fuzzy OLS estimators under consistent conditions.

In addition to methodological innovations, this thesis proposes a clear and structured approach for interpreting both fuzzy and crisp outputs. Given that fuzzy regression often produces results that are difficult to interpret directly, this work introduces a visually guided and intuitively accessible format for presenting outcomes—making it easier for researchers and practitioners to understand and apply the model's results in practical settings.

4. SIMULATION STUDY

In this section, a simulation study inspired by the work of Choi et al. [48] and Kim and Jung [49] was conducted with significant modifications to evaluate the performance of the proposed approach. In the study by Kim and Jung [49], K-fold cross-validation method was applied to the α-cut-based fuzzy Ridge regression. For simplicity, this method will be referred to as fuzzy Ridge in this work. Similarly, fuzzy OLS and fuzzy Liu will be used to represent their respective fuzzy regression methods throughout the text. For the fuzzy Ridge regression, 77 proposed methods, summarized in Table 3.1, will be employed to identify candidate values for the biased parameter k, which have not been previously tested with the α-cut-based estimation algorithm or other fuzzy regression algorithms under such a large number of scenarios. Additionally, the fuzzy Liu estimation method, consisting of 13 candidate values summarized in Table 3.2, will be tested for the first time under this extensive set of scenarios. This study aims to compare the performance of fuzzy Ridge, fuzzy OLS, and fuzzy Liu estimations using the fuzzy α-cut-based algorithm.

Samples were generated with $N = \{20,50,100,200\}$ observations for each of the $p = 4$ crisp independent variables. The dimensionality of the data aligns with commonly observed characteristics in fuzzy datasets. Following the approach of Gibbons [130], the independent variables x_{ij} were generated using Eq. 4.1, as follows:

$$x_{ij} = (1 - \rho^2)^{1/2} z_{ij}, \qquad i = 1,2,\dots,n \qquad j = 1,2,\dots,p \tag{4.1}$$

where ρ is a constant and z_{ij} are drawn from independent normal distributions with a mean of 50 and variance of one. The values x_{ij} are assumed to be non-negative to reflect the properties of real-world fuzzy data. The degree of linear association between independent variables is controlled by ρ, where ρ^2 represents the correlation between any two independent variables. 3 correlation levels were examined, corresponding to $\rho = 80$, $\rho = 90$, and $\rho = 0.99$, representing moderate, high, and very high correlations, respectively.

Observations for the fuzzy dependent variable were generated using Eq. 4.2-4.4, as follows:

$$y_{il} = A_{0l} + A_{1l}x_{i1} + \cdots + A_{pl}x_{ip} + \varepsilon_i^1 \quad (4.2)$$

$$y_{im} = A_{0m} + A_{1m}x_{i1} + \cdots + A_{pm}x_{ip} + \varepsilon_i^2 \quad (4.3)$$

$$y_{ir} = A_{0r} + A_{1r}x_{i1} + \cdots + A_{pr}x_{ip} + \varepsilon_i^3 \quad (4.4)$$

where ε_i^1, ε_i^2, and ε_i^3 are independent normal random variables with a mean of 0 and a variance of σ^2. 4 values of σ (0.25,0.5,1, and 1.5) were investigated to represent different levels of variability in the spreads of the fuzzy dependent variable. Larger values of σ correspond to greater variability.

The left and right spreads of the fuzzy dependent variable were determined using Eq. 4.5:

$$s_{il}^y = y_{im} - y_{il} \quad ,s_{ir}^y = y_{ir} - y_{im} \quad i = 1, \dots, N \quad (4.5)$$

Both symmetric $s_l^y = s_r^y$ and asymmetric $s_l^y \neq s_r^y$ cases were considered. The coefficients for the fuzzy coefficients were set as follows: $A_l = (0, 0.1, 0.15, 0.2, 0.25)$, $A_m = (0, 0.4, 0.45, 0.5, 0.55)$, and $A_r = (0, 0.7, 0.75, 0.8, 0.85)$.

To analyze the effects of the factors ρ, σ and N, the α-cut sequences were fixed as (0,0.25,0.5,0.75,1). Instead of the 200 replicates used in the original study, 1000 replicates were generated for each scenario. The independent variables and fuzzy coefficients were held constant across replicates, while the error terms, and consequently the fuzzy dependent variable, varied.

The simulated data were split into training and testing sets. The fuzzy Ridge, fuzzy OLS and fuzzy Liu estimation methods were fitted to the training data, and performance metrics—$RMSE_F$ —was computed using the testing data. For replicate t, the performance measures $RMSE_F^t$ was calculated, and Eq. 4.6 derives the average of the errors for each scenario.

$$RMSE_F = \frac{1}{1000}\sum_{t=1}^{1000} RMSE_F^t \quad (4.6)$$

The results of the simulation study are summarized in Tables 4.1-4.6, as follows:

NOTE: The results for the scenarios involving 77+1 methods for fuzzy Ridge regression, 13 methods for fuzzy Liu regression, and the fuzzy OLS method are extensive and would require a substantial number of pages to present comprehensively. As a result, the findings are summarized by emphasizing the best-performing and least-effective methods, offering an overall evaluation of their performance. In the table, "K" refers to the K-fold cross-validation method. The parameters d for Liu typically yields similar or very close results across different scenarios, and thus the results are represented as "d" for brevity. "L" denotes the fuzzy OLS method in this study. The numbers listed under each method indicate the ranking based on performance, alongside the corresponding error values.

Table 4.1. The performance measures when $\rho = 0.80$ and asymmetric data.

$\sigma = 0.25$								$\sigma = 0.5$							
$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 20$		$n = 50$		$n = 100$		$n = 200$	
k_{54} (1)	0.87	K (1)	0.63	K (1)	0.50	K (1)	0.47	k_{54} (1)	1.40	k_{54} (1)	1.16	k_{54} (1)	1.13	K (1)	0.95
K (6)	1.06	k_{50} (2)	0.76	k_{50} (2)	0.66	k_{24} (2)	0.57	K (70)	5.56	K (13)	1.94	K (2)	1.13	k_{15} (2)	1.07
d (11)	1.13	d (6)	0.84	d (7)	0.71	d (3)	0.57	d (71)	6.85	d (15)	3.33	d (12)	1.46	d (9)	1.14
L (38)	3.64	L (30)	0.96	L (27)	0.76	L (23)	0.59	k_{34} (90)	20.71	L (90)	5.02	L (33)	1.60	L (28)	1.16
k_7 (91)	12.27	k_7 (91)	14.08	k_7 (91)	14.39	k_7 (91)	11.40	L (91)	20.83	k_7 (91)	12.19	k_7 (91)	12.92	k_7 (91)	11.16
$\sigma = 1.0$								$\sigma = 1.5$							
$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 20$		$n = 50$		$n = 100$		$n = 200$	
k_{12} (1)	2.64	k_{54} (1)	2.37	k_{54} (1)	1.94	k_{54} (1)	1.92	k_{62} (1)	4.59	k_{12} (1)	3.59	k_{12} (1)	3.26	k_{54} (1)	2.76
K (70)	24.07	K (70)	10.48	K (69)	4.43	K (13)	2.36	K (70)	46.45	K (70)	23.53	K (70)	11.88	K (68)	4.97
d (72)	31.79	d (72)	17.93	d (71)	7.63	d (14)	3.59	d (71)	67.22	d (71)	41.95	d (72)	20.47	d (70)	8.50
k_{34} (90)	70.64	k_{34} (90)	24.02	L (90)	8.96	L (35)	3.87	k_{34} (90)	132.46	k_{34} (90)	53.11	k_{34} (90)	23.33	k_{34} (90)	9.24
L (91)	70.86	L (91)	24.07	k_7 (91)	9.90	k_7 (91)	9.82	L (91)	132.79	L (91)	53.19	L (91)	23.36	L (91)	9.26

In Table 4.1, a summarized version of the extensive results is presented. These results include 77+1 formulas for fuzzy Ridge regression, the fuzzy OLS method, and 13 formulas for Liu regression, totaling 92 methods. Listing all results for 96 scenarios would occupy substantial

space; thus, this table provides a concise summary of methods for determining the optimal k values in fuzzy Ridge regression, fuzzy OLS, and the d values in fuzzy Liu regression. For the fuzzy Liu regression, as previously mentioned, the results across all 13 formulas are generally similar or close so we provide the results of fuzzy Liu as d. Notably, the k_2 formula for the fuzzy Ridge often produced negative k value according to the simulated data, disrupting the simulation study, and was therefore excluded from the results.

The Table 4.1 focuses on scenarios where the ρ value is 0.80, and the dependent variable consists of asymmetric triangular fuzzy numbers. In most scenarios, using the proposed formulas for fuzzy Ridge regression yielded better results than methods like K-fold cross-validation. While certain scenarios showed optimal results with the K-fold method, the proposed formulas achieved comparable outcomes within significantly shorter computation times. In contrast, K-fold requires numerous iterations and may not always identify the optimal k values for the fuzzy Ridge regression. Additionally, when σ values are 0.25 and 0.5, the K-fold method performs better with larger sample sizes, but such cases demand more complex and larger matrices, increasing computation time. For fuzzy Liu regression, with $\sigma = 0.25$ across different sample sizes and $\sigma = 0.5$ for larger sample sizes, the fuzzy Liu formulas provided satisfactory results close to the optimal, but still slightly less effective than fuzzy Ridge regression.

Regarding fuzzy OLS, fuzzy Ridge regression converges to fuzzy OLS when $k = 0$. Hence, fuzzy Ridge results can serve to derive both of the fuzzy Ridge and the fuzzy OLS outcomes. Notably, k_{54} for the fuzzy Ridge consistently yielded the best results, with other formulas also showing comparable or nearly equivalent performance. the α set sequence was kept constant to isolate the effects of other factors. Testing different α sequences is expected to enhance the performance of other formulas, as will be shown in the next section, which involves using multiple sequences with real datasets. The data in this study were generated using a specific system distinct from real datasets. Overall, larger sample sizes logically result in smaller errors.

Table 4.2. The performance measures when $\rho = 0.80$ and symmetric data.

$\sigma = 0.25$								$\sigma = 0.5$							
$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 20$		$n = 50$		$n = 100$		$n = 200$	
k_{54} (1)	1.00	k_{54} (1)	0.95	k_{54} (1)	0.95	K (1)	1.00	k_{54} (1)	1.83	k_{54} (1)	1.49	k_{54} (1)	1.51	k_{54} (1)	1.53
K (12)	3.60	K (9)	2.42	K (6)	1.36	k_{54} (2)	1.12	K (69)	8.95	K (66)	4.77	K (10)	3.33	K (9)	2.42
d (71)	4.62	d (18)	3.41	d (15)	2.40	d (15)	1.90	d (71)	14.38	d (71)	8.44	d (71)	5.11	d (17)	3.55
L (90)	9.87	L (90)	4.04	L (34)	2.62	L (34)	2.00	k_{34} (90)	35.99	L (90)	10.97	L (90)	5.62	L (36)	3.68
k_7 (91)	15.35	k_7 (91)	14.66	k_7 (91)	14.92	k_7 (91)	11.67	L (91)	36.48	k_7 (91)	16.46	k_7 (91)	15.12	k_7 (91)	11.41
$\sigma = 1.0$								$\sigma = 1.5$							
$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 20$		$n = 50$		$n = 100$		$n = 200$	
k_{72} (1)	4.32	k_{54} (1)	3.06	k_{54} (1)	2.77	k_{54} (1)	2.74	k_{74} (1)	5.26	k_{72} (1)	5.07	k_{54} (1)	4.53	k_{54} (1)	4.05
K (69)	28.43	K (68)	14.26	K (68)	8.22	K (65)	5.18	K (66)	45.44	K (68)	25.24	K (68)	15.01	K (68)	8.95
d (72)	51.89	d (72)	32.63	d (71)	17.33	d (70)	8.82	d (72)	91.13	d (72)	62.41	d (72)	33.83	d (71)	19.03
k_{34} (90)	110.44	k_{34} (90)	41.88	L (90)	19.65	L (90)	9.34	k_{34} (90)	178.64	k_{34} (90)	77.74	k_{34} (90)	37.96	k_{34} (90)	20.35
L (91)	111.20	L (91)	42.08	k_7 (91)	20.01	k_7 (91)	13.80	L (91)	179.51	L (91)	77.97	L (91)	38.06	L (91)	20.42

In Table 4.2, the same scenarios as in Table 4.1 were examined; however, the dependent variable here consists of symmetric triangular fuzzy numbers. Notably, using formulas for the fuzzy Ridge regression continues to produce the best results in comparison to other methods in most scenarios. For σ values of 0.25 and 0.5, the K-fold method yields satisfactory results, although they differ from the outcomes in Table 4.1.

Regarding the fuzzy Liu method, the formulas did not perform as effectively in certain scenarios compared to those in Table 4.1 but still close in all scenarios. For the fuzzy Ridge regression, k_{54} remains the best-performing formula based on the simulated data and the chosen α sequence. It is important to note that using different α-cut sequences may alter the results. Finally, fuzzy OLS remains the least effective method in terms of performance according to the observed ranking.

Table 4.3. The performance measures when $\rho = 0.90$ and asymmetric data.

$\sigma = 0.25$								$\sigma = 0.5$							
$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 20$		$n = 50$		$n = 100$		$n = 200$	
k_{54} (1)	0.84	K (1)	0.59	K (1)	0.49	K (1)	0.46	k_{39} (1)	1.46	k_{54} (1)	1.12	k_{54} (1)	1.08	K (1)	0.94
K (2)	0.90	k_{54} (2)	0.77	k_{50} (2)	0.64	k_{15} (2)	0.56	K (69)	4.62	K (11)	1.92	K (2)	1.13	k_{49} (2)	1.07
d (7)	1.02	d (9)	0.90	d (4)	0.67	d (4)	0.57	d (71)	6.23	d (71)	4.59	d (15)	2.27	d (13)	1.27
L (90)	7.19	L (34)	1.44	L (27)	0.74	L (23)	0.60	k_{34} (90)	34.48	L (90)	8.96	L (34)	3.18	L (32)	1.36
k_7 (91)	11.11	k_7 (91)	13.59	k_7 (91)	13.68	k_7 (91)	10.88	L (91)	34.92	k_7 (91)	11.44	k_7 (91)	12.67	k_7 (91)	10.47
$\sigma = 1.0$								$\sigma = 1.5$							
$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 20$		$n = 50$		$n = 100$		$n = 200$	
k_{13} (1)	2.58	k_{63} (1)	2.33	k_{54} (1)	1.96	k_{54} (1)	1.89	k_{12} (1)	4.20	k_{12} (1)	3.37	k_{63} (1)	3.14	k_{54} (1)	2.76
K (69)	22.50	K (68)	10.28	K (69)	4.87	K (12)	2.29	K (69)	47.72	K (68)	24.41	K (69)	12.54	K (67)	5.32
d (71)	32.53	d (72)	23.57	d (71)	13.03	d (70)	5.36	d (71)	67.15	d (72)	52.33	d (71)	31.22	d (71)	14.72
k_{34} (90)	108.72	k_{34} (90)	39.67	L (90)	17.54	L (90)	6.46	k_{34} (90)	193.49	k_{34} (90)	79.25	k_{34} (90)	39.94	k_{34} (90)	17.34
L (91)	109.58	L (91)	39.86	k_7 (91)	17.65	k_7 (91)	9.14	L (91)	194.51	L (91)	79.51	L (91)	40.09	L (91)	17.43

In Table 4.3, the scenarios align with those in the previous tables, particularly Table 4.1, with the dependent variable remaining asymmetric but with $\rho = 0.90$, representing a high correlation degree. Notably, the K-fold method demonstrates greater effectiveness in identifying optimal results across more scenarios compared to Tables 4.1 and 4.2. However, similar or very close results are achieved using specific formulas, such as k_{54},k_{50},k_{15},k_{49}, and others, which are not included in the table to conserve space. These formulas offer faster results, particularly when comparing the computational time required for K-fold method with numerous candidates and iterations versus testing 77 formulas for the fuzzy Ridge regression.

For fuzzy Liu regression, the performance remains consistent, yielding good results for $\sigma = 0.25$ and large sample sizes with $\sigma = 0.5$, though the performance declines for larger σ values. Fuzzy OLS shows comparable results in certain scenarios; however, it remains unreliable as a primary method.

Table 4.4. The performance measures when $\rho = 0.90$ and symmetric data.

$\sigma = 0.25$								$\sigma = 0.5$							
$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 20$		$n = 50$		$n = 100$		$n = 200$	
k_{54} (1)	1.02	k_{54} (1)	0.90	k_{54} (1)	0.88	K (1)	0.92	k_{12} (1)	2.34	k_{54} (1)	1.50	k_{54} (1)	1.46	k_{54} (1)	1.50
K (9)	3.05	K (8)	2.21	K (7)	1.33	k_{54} (2)	0.95	K (69)	8.44	K (65)	4.65	K (9)	3.03	K (8)	2.38
d (22)	3.96	d (16)	3.28	d (15)	2.26	d (15)	1.75	d (71)	13.41	d (71)	10.27	d (71)	5.55	d (20)	3.98
k_{34} (90)	15.70	L (90)	4.96	L (34)	2.63	L (34)	1.92	k_{34} (90)	56.55	k_{34} (90)	18.27	L (90)	7.28	L (51)	4.36
L (91)	16.42	k_7 (91)	14.01	k_7 (91)	14.13	k_7 (91)	11.01	L (91)	58.18	L (91)	18.68	k_7 (91)	14.32	k_7 (91)	11.32
$\sigma = 1.0$								$\sigma = 1.5$							
$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 20$		$n = 50$		$n = 100$		$n = 200$	
k_{74} (1)	4.35	k_{54} (1)	3.42	k_{54} (1)	2.80	k_{54} (1)	2.74	k_{29} (1)	5.52	k_{74} (1)	5.06	k_{54} (1)	4.57	k_{54} (1)	4.06
K (69)	27.00	K (68)	14.58	K (68)	8.67	K (66)	5.00	K (68)	48.59	K (68)	26.32	K (68)	15.74	K (67)	9.02
d (71)	50.27	d (72)	39.08	d (71)	24.40	d (70)	12.33	d (71)	90.86	d (72)	75.40	d (71)	47.49	d (71)	26.45
k_{34} (90)	163.64	k_{34} (90)	63.20	k_{34} (90)	31.46	k_{34} (90)	14.24	k_{34} (90)	253.92	k_{34} (90)	113.78	k_{34} (90)	59.99	k_{34} (90)	30.47
L (91)	166.60	L (91)	63.92	L (91)	31.82	L (91)	14.44	L (91)	256.93	L (91)	114.64	L (91)	60.41	L (91)	30.77

In Table 4.4, the scenarios mirror those in Table 4.3, but the dependent variable's values are symmetric. As observed in Table 4.2, the performance of the K-fold method for the fuzzy Ridge regression declines compared to the formulas. Specifically, k_{54} and other formulas continue to provide the best results for the fuzzy Ridge regression based on the simulated data and the chosen α sequence.

Similarly, the fuzzy Liu and fuzzy OLS methods perform comparably to the K-fold method. Furthermore, it is evident that for the fuzzy Ridge regression, formulas such as k_7,k_{34},k_{36}, and some others consistently yield the poorest results in comparison to all other methods and formulas.

Table 4.5. The performance measures when $\rho = 0.99$ and asymmetric data.

$\sigma = 0.25$								$\sigma = 0.5$							
$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 20$		$n = 50$		$n = 100$		$n = 200$	
K (1)	0.62	K (1)	0.49	K (1)	0.45	K (1)	0.45	k_{14} (1)	1.30	K (1)	1.15	K (1)	0.96	K (1)	0.89
d (2)	0.67	k_{42} (2)	0.66	k_{40} (2)	0.55	k_{40} (2)	0.52	K (10)	1.93	k_{14} (2)	1.17	k_{14} (2)	1.05	k_{40} (2)	0.98
k_{14} (15)	0.79	d (6)	0.74	d (6)	0.63	d (6)	0.55	d (11)	2.71	d (67)	4.45	d (69)	4.86	d (15)	3.39
k_{36} (90)	56.28	k_{36} (90)	18.44	L (34)	6.36	L (34)	1.73	k_{32} (90)	146.95	k_{36} (90)	61.17	k_{36} (90)	29.51	k_{36} (90)	13.67
L (91)	61.52	L (91)	20.08	k_7 (91)	11.18	k_7 (91)	9.42	L (91)	159.06	L (91)	65.21	L (91)	31.63	L (91)	14.96
$\sigma = 1.0$								$\sigma = 1.5$							
$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 20$		$n = 50$		$n = 100$		$n = 200$	
k_{13} (1)	2.51	k_{39} (1)	2.10	k_{14} (1)	2.05	k_{14} (1)	1.91	k_{13} (1)	3.77	k_{13} (1)	3.13	k_{39} (1)	2.96	k_{14} (1)	2.80
K (65)	11.38	K (59)	4.80	K (10)	2.89	K (2)	1.92	K (67)	29.04	K (65)	13.52	K (64)	7.87	K (10)	3.69
d (67)	14.68	d (70)	23.82	d (70)	26.27	d (71)	18.29	d (69)	33.81	d (69)	47.83	d (70)	53.00	d (72)	41.25
k_{34} (90)	325.25	k_{34} (90)	163.71	k_{34} (90)	93.03	k_{34} (90)	48.30	k_{34} (90)	531.81	k_{34} (90)	265.4	k_{34} (90)	155.5	k_{50} (90)	94.81
L (91)	345.82	L (91)	172.18	L (91)	97.61	L (91)	50.79	L (91)	559.49	L (91)	275.7	L (91)	160.9	L (91)	97.62

In Table 4.5, the ρ value is set to 0.99, indicating a very high correlation and significant multicollinearity. The K-fold method demonstrates the best performance when the σ values are 0.25 and 0.5. However, comparable results can still be achieved using certain formulas such as k_{14},k_{40},k_{42}, and others for the fuzzy Ridge regression.

For σ values of 1.0 and 1.5, the results diverge, particularly when the dependent variable's values are asymmetric. Additionally, the fuzzy Liu regression shows some promising outcomes when $\sigma = 0.25$ and, to a lesser extent, when $\sigma = 0.5$. As observed in the previous tables, the fuzzy OLS method consistently fails to deliver satisfactory performance across all scenarios.

Table 4.6. The performance measures when $\rho = 0.99$ and symmetric data.

$\sigma = 0.25$								$\sigma = 0.5$							
$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 20$		$n = 50$		$n = 100$		$n = 200$	
k_{63} (1)	1.33	k_{39} (1)	1.12	k_{54} (1)	0.87	K (1)	0.76	k_{13} (1)	2.21	k_{13} (1)	1.93	k_{63} (1)	1.80	k_{54} (1)	1.61
K (7)	2.12	K (8)	1.50	K (2)	0.98	k_{54} (2)	0.88	K (65)	5.65	K (10)	3.42	K (8)	2.63	K (8)	1.80
d (8)	2.48	d (12)	2.39	d (12)	1.99	d (14)	1.48	d (67)	7.53	d (69)	9.71	d (69)	10.89	d (70)	7.24
k_{34} (90)	73.21	k_{34} (90)	26.77	k_{7} (90)	11.97	L (34)	5.78	k_{34} (90)	186.70	k_{34} (90)	87.72	k_{34} (90)	45.51	k_{50} (90)	22.12
L (91)	87.95	L (91)	32.85	L (91)	14.15	k_{7} (91)	9.52	L (91)	218.98	L (91)	99.69	L (91)	52.48	L (91)	26.38
$\sigma = 1.0$								$\sigma = 1.5$							
$n = 20$		$n = 50$		$n = 100$		$n = 200$		$n = 20$		$n = 50$		$n = 100$		$n = 200$	
k_{12} (1)	4.38	k_{12} (1)	3.61	k_{12} (1)	3.35	k_{13} (1)	3.11	k_{62} (1)	5.43	k_{29} (1)	5.23	k_{12} (1)	4.98	k_{13} (1)	4.54
K (67)	18.09	K (66)	10.16	K (65)	6.83	K (9)	4.25	K (67)	37.43	K (67)	22.32	K (66)	13.99	K (61)	7.24
d (69)	25.05	d (70)	36.71	d (70)	39.96	d (72)	31.46	d (69)	49.39	d (69)	70.05	d (70)	75.94	d (72)	61.31
k_{34} (90)	424.77	k_{34} (90)	211.0	k_{34} (90)	123.9	k_{34} (50)	76.98	k_{34} (90)	672.86	k_{34} (90)	349.73	k_{34} (90)	205.9	k_{50} (90)	132.6
L (91)	489.25	L (91)	234.3	L (91)	136.6	L (91)	81.18	L (91)	753.96	L (91)	377.66	L (91)	220.9	L (91)	135.6

For the final table, Table 4.6, the dependent variable's values are symmetric. The K-fold method continues to deliver good results when $\sigma = 0.25$ and $\sigma = 0.5$, though it does not achieve optimal performance in the fuzzy Ridge regression compared to formulas such as k_{12}, k_{13},k_{39},k_{54},k_{63}, and others. The performance of d-based methods follow a similar trend, yielding reasonable results but in a different order compared to Table 4.5.

The fuzzy OLS method remains consistent with previous observations, providing suboptimal performance based on the generated data and the chosen σ sequence. Additionally, across all the tables, it is evident that errors are generally larger when the dependent variable is symmetric compared to when it is asymmetric.

5. EMPIRICAL STUDY

In this section, the performance of various methods for determining the optimal k value for fuzzy Ridge regression, fuzzy OLS, and the fuzzy Liu regression using α-cut-based estimation algorithm is evaluated using 3 datasets widely utilized in studies involving fuzzy logic and multicollinearity:

Dataset 1: Shanghai Housing Price Data

This dataset, presented in an asymmetric fuzzy format by Zhou et al. [131], consists of 147 samples and 6 crisp independent variables: housing size, mortgage interest rate, real estate tax, down payment ratio, annual household income, and family population. The dependent variable is an asymmetric fuzzy representation of house price.

To assess multicollinearity, the VIF was calculated, yielding the following values: $VIF_1 = 1.69, VIF_2 = 1.12, VIF_3 = 1.09, VIF_4 = 1.16, VIF_5 = 1.69, VIF_6 = 1.05$.

Generally, variables with $VIF > 5$ indicate problematic levels of multicollinearity and when $VIF > 10$ indicate a high prevalence of multicollinearity. In this dataset, the low VIF values suggest either an absence of multicollinearity or negligible levels that are unlikely to impact the analysis.

Dataset 2: Tanaka's Dataset

This dataset, introduced by Tanaka et al. [15], includes 5 crisp independent variables: material rank, first-floor space (m^2), second-floor space (m^2), number of rooms, and number of Japanese-style rooms. The dependent variable is represented as a symmetric fuzzy house price, and the dataset comprises 15 observations.

The VIF values are: $VIF_1 = 9.50, VIF_2 = 6.05, VIF_3 = 1.67, VIF_4 = 3.87, VIF_5 = 6.13$. These results indicate problematic levels of multicollinearity, requiring careful consideration in model development.

Dataset 3: Cheese Taste Data

The third dataset, analyzed by Kashani et al. [132], involves three independent variables: acetic acid, hydrogen sulfide, and lactic acid. The dependent variable is a symmetric fuzzy representation of cheese taste. This dataset contains 15 observations.

The VIF values are: $VIF_1 = 12.83$, $VIF_2 = 8.77$, $VIF_3 = 2.63$. These values reflect high multicollinearity.

In the sixth step of the estimation process (Section 3.7), steps 1–5 are repeated with user-defined α-cut sequences. Table 5.1 shows these sequences, equally spaced between 0 and 1, with step sizes set by the researchers.

The results, evaluated using performance measures such as $RSME_F$, are compared to identify the best α-cut sequence for model construction. Researchers can select smaller step sizes to refine the α-cut sets further.

Table 5.1. Suggested candidates for α-cut sequences.

No	α-cut seq.	No	α-cut seq.	No	α-cut seq.
1	0:(0.01):1	10	0:(0.06):1	19	0:(0.35):1
2	0:(0.015):1	11	0:(0.07):1	20	0:(0.4):1
3	0:(0.02):1	12	0:(0.08):1	21	0:(0.45):1
4	0:(0.025):1	13	0:(0.09):1	22	0:(0.5):1
5	0:(0.030):1	14	0:(0.1):1	23	0:(0.6):1
6	0:(0.035):1	15	0:(0.15):1	24	0:(0.7):1
7	0:(0.040):1	16	0:(0.2):1	25	0:(0.8):1
8	0:(0.045):1	17	0:(0.25):1	26	0:(0.9):1
9	0:(0.05):1	18	0:(0.3):1	27	0:(1.0):1

Table 5.1 allows researchers to customize step sizes and total iterations to meet the specific needs of their analysis. The iterative procedure ensures the selection of the most appropriate α-cut sequence, contributing to the construction of reliable fuzzy regression models. For instance, the sequence 0:(0.15):1 incorporates all crisp numbers from the dependent variable's values with membership levels of 0,0.15,0.30,0.45,0.60,0.75, and 0.90. Consequently, fuzzy Ridge, fuzzy Liu, and fuzzy OLS methods are each applied 14 times.

Additionally, it is important to note that 77+1 methods are used for k in the fuzzy Ridge regression, 13 for the fuzzy Liu regression, and one for the fuzzy OLS.

The outcomes obtained from the application of the fuzzy Ridge, fuzzy OLS, and fuzzy Liu are presented in Table 5.2.

Table 5.2. Results of applying different methods to the 3 datasets.

Method	Optimal Biased Parameters	Optimal α-Cut Seq.	The Values of the Optimal Biased Parameters	$RMSE_F$
Dataset 1: Shanghai Housing Price (Low multicollinearity)	k_{70}	25	Left-Side: 1.04, 7.14 Right-Side: 10.30, 32.81	157.49
	K-fold	15	Left-Side: 19.34, 18.61, 17.88, 17.16, 16.43, 20.75, 26.16 Right-Side: 20.16, 20.65, 21.14, 21.64, 22.13, 22.62, 23.12	169.08
	Liu	15	Left side: 0.88, 0.95, 0.97, 0.98, 0.99, 0.96, 0.39 Right side: 0.58, 0.26, 0, 0, 0, 0, 0	171.32
	OLS	27		172.18
Dataset 2: Tanaka's Dataset (Moderate multicollinearity)	k_{10}	26	Left side: 439.28, 382.44. Right side: 450.42, 504.01	1246.94
	K-fold	16	Left side: 320.96, 320.72, 320.47, 320.23, 319.99, 319.74 Right side: 320.96, 321.21, 321.45, 321.70, 321.94, 322.19	1352.25
	Liu	9	Left side: 0 (21 terms) Right side: 0 (21 terms)	1430.48
	OLS	22		2321.24
Dataset 3: Cheese Taste (High multicollinearity)	k_{47}	27	Left side: 7.20 and 3.76 Right side: 7.20 and 5.45	15.82
	K-fold	25	Left side: 3.44 and 3.37 Right side: 1.43 and 1.58	22.94
	Liu	1	Left side: 0 (101 terms) Right side: 0 (101 terms)	34.61
	OLS	17		76.75

From the results presented in Table 5.2, the discussion of Tanaka's dataset will be deferred to the end, as it is the most extensively used dataset in the fuzzy regression field, warranting more detailed analysis. Regarding the first dataset, Shanghai Housing Price Data, which exhibits low VIF values indicating no multicollinearity issues, the results show that the best outcome was achieved using k_{70}, with an error of 157.49. Notably, the K-fold cross-validation method did not provide the best results. It is well-known that the K-fold method might yield optimal results by selecting a large candidate set of k-values and repeating the computation process many times, but this increases computational demand.

The Shanghai Housing Price dataset contains 147 observations and 6 crisp independent variables, as previously mentioned. Additionally, several formulas for the fuzzy Ridge regression, such as k_{29},k_{69},k_{77} and k_{79}, produced better results than the K-fold method. While these formulas do not always yield optimal results, they provide outcomes very close to the best solution.

In the simulation study, formulas like k_{12},k_{13}, and particularly k_{54} were found to perform optimally. However, for this dataset, they produced suboptimal results, with k_{54} yielding an error of 390.92. This highlights the importance of testing a wide range of α-cut sequences, as these can significantly influence the method's performance. In the simulation study, the α-cut sequences were fixed across scenarios, which may have limited their adaptability.

The optimal results for k_{70} were obtained using an α-cut sequence of 0:0.8:1. This approach applies the classical Ridge method four times—twice on the left side of the triangular fuzzy number and twice on the right side—by utilizing the crisp values of the dependent variable corresponding to membership values of 0 and 0.8. Similarly, for the K-fold method, the α-cut sequence 0:0.15:1 (i.e., $\alpha = 0,0.15,0.30, \dots ,0.90$) was used, requiring Ridge regression to be applied 14 times. However, due to K-fold's inherent iterative process, the total number of Ridge applications becomes significantly large, which may not always lead to optimal results.

For the fuzzy Liu regression, the results of the 13 d-formulas were close, with the best result achieved using the α-cut sequence 0:0.15:1. Consistent with the simulation study, the fuzzy

Liu regression's performance depended on leveraging the fuzzy Ridge regression coefficients effectively, demonstrating competitive results compared to the K-fold method.

For the fuzzy OLS, the results were less effective than the previously discussed approaches. The α-cut sequence 0:1:1 was used, applying the method only at $\alpha = 0$ and $\alpha = 1$. These values correspond to the crisp bounds (minimum and maximum) and core (membership value of 1) for each fuzzy number.

For the Cheese Taste Data, which has high VIF values, the same performance order observed in the first dataset was found across the approaches, with the optimal k value achieved using k_{47}. This dataset contains 15 observations with 3 independent variables. In fuzzy Ridge regression, 14 formulas outperformed the K-fold method, including k_{10},k_{27},k_{75},k_{76}, and others, which produced similar results, consistent with observations in the simulation study. Fuzzy Liu and fuzzy OLS methods performed similarly to each other for this dataset.

For the Tanaka's dataset, which consists of 15 observations, one symmetric dependent variable, and five crisp independent variables, the order of approach performances remained consistent with the previous datasets. The best result was obtained using k_{10}, which also performed well in Dataset 3, as previously mentioned. In addition to k_{10}, formulas k_8,k_{11},k_{48}, and k_{59} outperformed the K-fold method. In simulation studies, formulas such as k_{12},k_{13}, and k_{54}, and others achieved optimal results when all independent variables had the same degree of correlation with the one of them. However, for the three datasets used here, the correlation among independent variables varied, which lowered the performance of these formulas. This highlights the importance of utilizing diverse α-cut sequences, which enhance the ability of formulas to detect optimal k values for fuzzy Ridge regression. Fuzzy Liu regression provided strong results with d_1, d_2, d_3, d_4, d_5, d_{10}, d_{11}, d_{12}, and d_{13}, while fuzzy Liu produced the worst results when using d_6, d_7, d_8, and d_9 and fuzzy OLS were better than them.

Taking into account the use of α-cut sequences can significantly improve results by enhancing the ability to select optimal formulas or methods for determining the biased d or k values, even for the fuzzy OLS. Figure 5.1 presents the $RMSE_F$ values corresponding to

each sequence from Table 5.1, alongside the optimal sequences identified for each approach in Table 5.2, as follows:

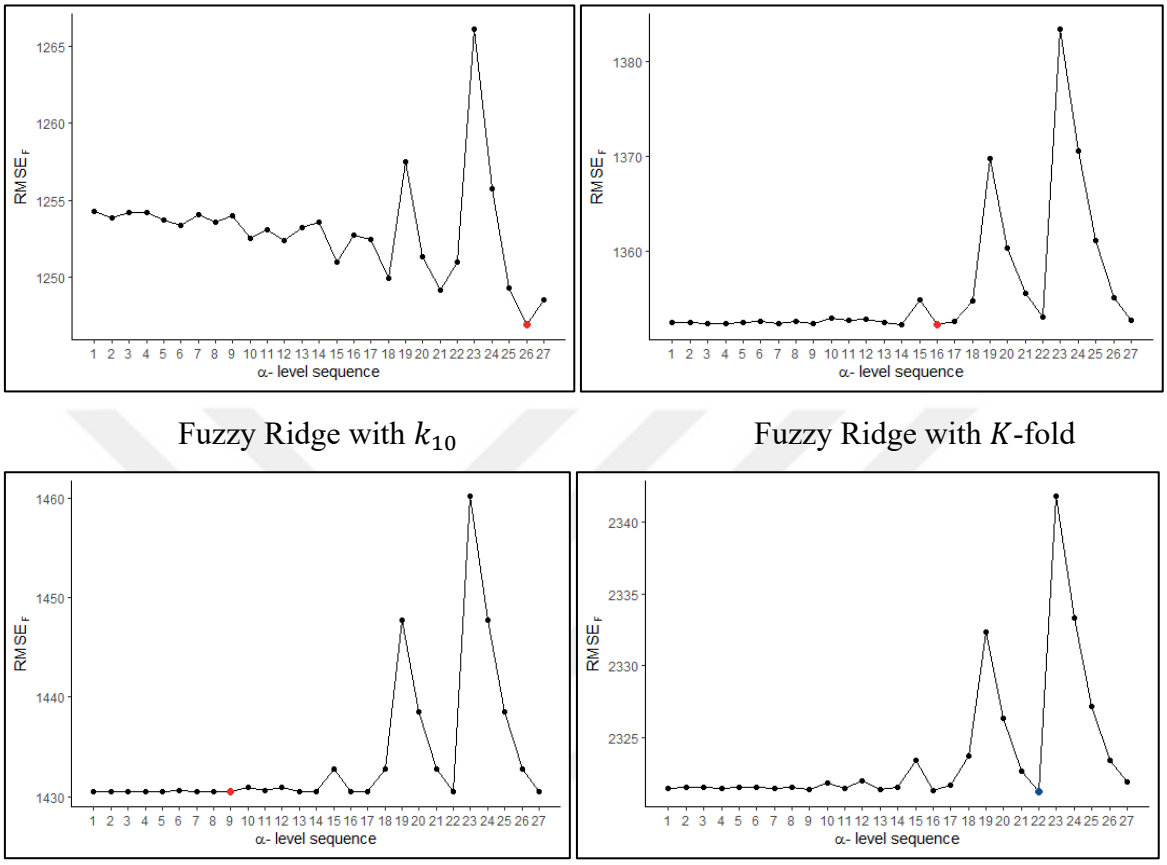


Figure 5.1. $RMSE_F$ values across α-cut sequences for various approaches.

Using the optimal methods, the regression models can be constructed as fuzzy regression models as follows:

α-cut based fuzz Ridge with k_{10}:

$$\tilde{y} = (-1857.85, -1162.91, -467.97)_T + 1820.59x_1 + 72.59x_2 + (57.95,62.35,66.76)_T x_3 - 111.85x_4 + (586.85,665.11,743.38)_T x_5$$

α-cut based fuzz Ridge with K-fold:

$$\tilde{y} = (-1864.95, -1174.00, -483.05)_T + 1899.23x_1 + (74.59,76.36,78.12)_T x_2 + (59.87,66.61,73.35)_T x_3 - 162.39x_4 + (507.71,552.38,597.05,)_T x_5$$

α-cut based fuzz Liu:

$$\tilde{y} = (-1924.26, -1207.75, -491.24)_T + 1966.43x_1 + (76.74,79.00,81.24)_T x_2 + (61.99,69.19,76.40)_T x_3 - 193.80x_4 + (406.73,456.29,505.85)_T x_5$$

α-cut based fuzz OLS:

$$\tilde{y} = (-2038.12, -1129.60, -221.09)_T + 2386.56x_1 + (87.13,93.36,99.60)_T x_2 + (71.72,82.13,92.54)_T x_3 - 376.35x_4 + (-285.03, -188.25, -91.48)_T x_5$$

As observed from the models across all approaches, X_1 and X_4 consistently exhibit crisp coefficient values. Additionally, using the optimal method k_{10}, X_2 also shares the same crisp coefficient behavior alongside X_1 and X_4. This highlights that fuzzy regression methods are instrumental in identifying the fuzzy relationships between the dependent and independent variables, as well as their ranges, by examining the spreads of the fuzzy model coefficients. By comparing the left and right spreads from the models, insights can be gained into whether the coefficients lean more to the left or right, indicating the direction and strength of these relationships. This demonstrates the importance of incorporating fuzzy logic, particularly with alternative shapes beyond triangular fuzzy numbers, such as Gaussian and other forms, to capture more nuanced relationships.

The estimated and observed values for the four approaches applied to Tanaka's dataset are presented in Figure 5.2 as follows:

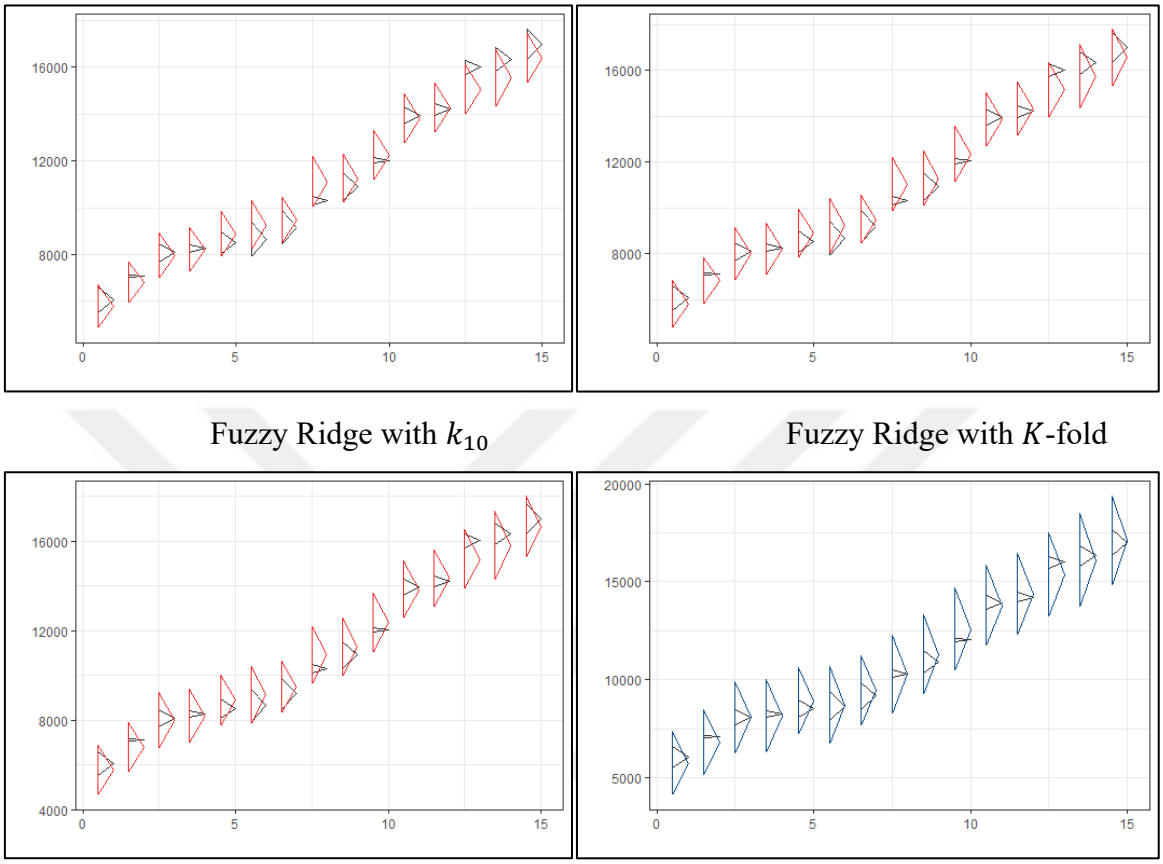


Figure 5.2. Expected and observed values for Tanaka's dataset (Large Triangles Represent the Estimated Values).

As shown in Figure 5.2, k_{10} was identified as the optimal value based on the candidate values and methods used. In contrast, fuzzy OLS did not exhibit a small spread when compared to the other methods, with most of the k candidates for the fuzzy Ridge and d for the fuzzy Liu.

The observed values and the corresponding relative errors of the estimated fuzzy outputs for the four models, as illustrated in Figure 5.2, are presented in Table 5.3 below:

Table 5.3. Relative errors of cores and spreads components across different methods, with corresponding MAPE values.

	Cores c_i					Spreads s_i				
i	Observed	k_{10}	K-fold	Liu	OLS	Observed	k_{10}	K-fold	Liu	OLS
1	6060	4.29%	4.28%	4.69%	5.36%	550	69.75%	90.63%	102.59%	194.90%
2	7100	4.06%	4.08%	4.35%	4.44%	50	1679.80%	1947.96%	2093.42%	3236.46%
3	8080	1.53%	1.15%	1.13%	0.26%	400	142.52%	187.41%	207.91%	367.04%
4	8260	0.57%	0.66%	0.89%	1.41%	150	527.29%	649.36%	705.43%	1144.19%
5	8650	7.03%	6.43%	5.59%	0.39%	750	37.84%	59.01%	71.13%	167.05%
6	8520	3.98%	4.34%	4.24%	4.28%	450	115.14%	133.88%	150.15%	279.58%
7	9170	3.26%	3.45%	3.30%	2.91%	700	38.30%	52.86%	64.01%	152.87%
8	10310	7.63%	6.83%	5.85%	0.42%	200	437.79%	487.75%	532.94%	896.33%
9	10920	2.89%	3.23%	3.24%	3.26%	600	73.51%	100.84%	116.19%	237.46%
10	12030	1.72%	2.45%	2.76%	4.42%	100	939.24%	1127.20%	1225.90%	2008.54%
11	13940	1.02%	0.74%	0.75%	1.13%	350	205.06%	239.31%	266.78%	488.33%
12	14200	0.44%	0.86%	0.96%	1.21%	250	320.58%	371.85%	411.47%	732.16%
13	16010	6.06%	5.44%	5.17%	3.98%	300	250.48%	299.52%	334.26%	615.72%
14	16320	4.77%	3.75%	3.30%	1.39%	500	144.79%	183.06%	204.96%	377.66%
15	16990	3.58%	2.56%	2.01%	0.69%	650	64.88%	93.12%	110.56%	251.64%
MAPE		3.52%	3.35%	3.22%	2.37%	MAPE	336.47%	401.58%	439.85%	743.33%

All models are represented using symmetric triangular fuzzy numbers, as shown in Figure 5.2. Table 5.3 is organized into two main sections:

- The left section reports the relative errors (expressed as percentages) for the core (center) values.
- The right section reports the relative errors for the spreads, which capture the fuzziness around the cores.

The relative error for each observation is computed using Eq. 5.1 and Eq. 5.2, respectively:

$$Relative\ Error_{\ core} = \left|\frac{\hat{c}-c}{c}\right| \tag{5.1}$$

$$Relative\ Error_{\ spread} = \left|\frac{\hat{s}-s}{s}\right| \tag{5.2}$$

Additionally, the Mean Absolute Percentage Error ($MAPE$), a widely used metric for assessing prediction accuracy, is calculated for each method and presented in the final row

of Table 5.3. $MAPE$ represents the average of the absolute percentage differences between observed and predicted values, and is computed using *Eq. 5.3 and Eq. 5.4:*

$$MAPE_{\ core} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{c}_i - c_i}{c_i}\right| \tag{5.3}$$

$$MAPE_{\ spread} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{s}_i - s_i}{s_i}\right| \tag{5.4}$$

Summary of results:

- For the core errors, the fuzzy OLS model achieved the best accuracy with a $MAPE$ of approximately 2.37%, followed by fuzzy Liu, fuzzy Ridge with K-fold cross-validation, and fuzzy Ridge with the k_{10} formula ($MAPE$s ranging from 3.22% to 3.52%). This indicates that fuzzy OLS is the most effective at estimating the central tendency of the fuzzy outputs.
- In contrast, for the spread errors, which quantify the uncertainty or fuzziness around the core, fuzzy Ridge using k_{10} outperformed all other methods with a $MAPE$ of approximately 336.47%, while fuzzy OLS resulted in the highest spread error at around 743.33%. This suggests that although fuzzy OLS excels in predicting the center, it is less reliable in estimating the spread compared to fuzzy Ridge with k_{10}.
- Fuzzy Ridge with K-fold and fuzzy Liu exhibited intermediate performance in both core and spread estimation, indicating a trade-off between center and spread accuracy.

Note: Note: All computations and analyses in this thesis were performed using Microsoft Excel and RStudio (version 2024.09.0). Custom scripts based on the fuzzy regression methodology of Kim and Jung [49] were used, but the full code is not included to respect copyright and because of its length. Functions for estimating the shrinkage parameters k (Ridge) and d (Liu) were developed by the author and will be shared online after completing ongoing research. Readers can refer to [49] for foundational implementation details.

6. CONCLUSION AND RECOMMENDATIONS

Fuzzy logic has proven to be a pivotal tool across various disciplines, particularly in mathematics and statistics, where it empowers decision-making and model-building. One of the key advantages of fuzzy logic is its ability to generalize classical logic, allowing results from conventional methods to be seamlessly integrated. This versatility is rooted in its foundational processes—fuzzification and defuzzification—which play a critical role in achieving accurate results. Selecting appropriate fuzzification and defuzzification methods requires a deep understanding of the dataset and the objectives of the analysis.

As discussed in Chapter 2, fuzzy sets and operations on fuzzy numbers are extensions of their crisp counterparts. This relationship makes fuzzy logic a generalized framework, with fuzzy numbers serving as an inclusive case of crisp numbers. Classical statistical methods have been extended through fuzzy logic to accommodate the inherent uncertainties in real-world data, providing more comprehensive insights. Among these methods, regression analysis has received significant attention, with numerous studies expanding it into the fuzzy domain.

Fuzzy regression methods differ fundamentally from classical regression by not requiring the assumptions of normality, homoscedasticity, or linearity. Instead, they inherently account for errors and uncertainties due to the complex relationships among variables. Fuzzy Ridge regression, fuzzy Liu regression, and fuzzy OLS methods were introduced and adapted for fuzzy environments, particularly to address the multicollinearity problem. Existing literature highlights several fuzzy regression methods, with a notable focus on approaches that account for not only the lower and upper bounds of fuzzy numbers but also additional crisp components.

In this study, triangular fuzzy numbers were utilized as the dependent variable, and the α-cut-based estimation algorithm was employed. The α-cut-based approach is integral to fuzzy regression, as it transforms fuzzy numbers into crisp sets, enabling systematic computations. Previous research has explored fuzzy Ridge regression using methods like K-fold cross-validation to determine the bias parameter. While effective, K-fold cross-validation often requires extensive iterations and computational resources and may not consistently yield

optimal results. Additionally, the fuzzy Liu regression method had not been previously investigated using the α-cut-based fuzzy estimation approach.

This work evaluated the fuzzy Ridge regression with 77+1 candidate formulas for determining the bias parameter k, the fuzzy Liu regression with 13 candidate formulas for d, and the fuzzy OLS across a wide range of simulated scenarios and 3 real-world datasets.

Key findings

1. The fuzzy Ridge regression using candidate formulas for k proved to be the most effective approach overall. While K-fold cross-validation occasionally delivered optimal results, several formulas achieved comparable performance with significantly lower computational time. Testing all 77 formulas highlighted their practicality, as many provided robust results across different scenarios.
2. The fuzzy Liu regression demonstrated relatively stable performance across various scenarios. Although its results were generally similar to fuzzy Ridge regression, fuzzy Ridge consistently outperformed fuzzy Liu, particularly when using K-fold cross-validation or most of the 77 candidate formulas. This makes Ridge a more reliable choice in practice.
3. The fuzzy OLS consistently underperformed in most scenarios, even when multicollinearity was low. As fuzzy Ridge regression reduces to fuzzy OLS when $k = 0$, the fuzzy Ridge effectively subsumes the fuzzy OLS, offering superior flexibility and performance.
4. Incorporating diverse α-cut sequences enhanced the robustness of the results. By systematically varying the α-cuts, it was possible to refine the ranking of formulas and methods, demonstrating the value of this approach in both simulated and real-world datasets.

Implications

The findings underscore the efficiency and versatility of α-cut-based fuzzy regression methods. Testing multiple candidate values for k and d, coupled with diverse α-cut sequences, allows for more accurate and computationally efficient modeling. While

triangular fuzzy numbers were used in this study, other shapes can be explored depending on the nature of the dataset, number of independent variables, and sample size. These factors may influence the performance and applicability of different estimation methods.

In conclusion, the α-cut-based approach offers a robust framework for addressing uncertainties in data, regardless of the degree of multicollinearity. This flexibility paves the way for extending other estimation methods within the fuzzy regression paradigm, as evidenced by the results of this study. Future research may explore alternative fuzzy number shapes and further optimize candidate formulas to enhance performance across diverse application domains.

Data and code availability

The original and utilized codes, methods, simulated data, and datasets were cited and referenced in the text according to their appearance. Modifications and improvements were also detailed in the thesis in the order in which they appear.

REFERENCES

1. Łukasiewicz, J. (1930). Philosophische Bemerkungen zu mehrwertigen Systemen des Aussagenkalküls [Philosophical Remarks on Many-Valued Systems of Propositional Logic]. *Comptes Rendus Des Séances De La Société Des Sciences Et Des Lettres De Varsovie Cl III*, 23, 51–77.

2. Dubois, D. and Prade, H. (1979). Operations in a fuzzy-valued logic. *Information and Control*, 43(2), 224-240.

3. Heisenberg, W. (1927). Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik [On the Physical Content of Quantum Theoretical Kinematics and Mechanics]. *Zeitschrift für Physik*, 43(3-4), 172-198.

4. Zadeh, L.A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338-353.

5. Vargeloğlu, A.A. (2020). *Sezgisel kümelere dayalı bulanık regresyon analizi ve uygulamaları [Fuzzy Regression Analysis Based on Intuitionistic Sets and Its Applications]*. M. Sc. Thesis, Gazi University Graduate School Of Natural And Applied Sciences, Ankara.

6. Zadeh, L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning—I. *Information Sciences*, 8(3), 199-249.

7. Zadeh, L.A. (1988). Fuzzy logic. *Computer*, 21(4), 83-93.

8. Zadeh, L.A. (1995). Discussion probability theory and fuzzy logic are complementary rather than competitive. *Technometrics*, 37(3), 271-276.

9. Zadeh, L.A. (1996). Fuzzy logic = Computing with words. *IEEE Transactions on Fuzzy Systems*, 4(2), 103-111.

10. Ross, T.J. (2010). Classical sets and fuzzy sets, *in Fuzzy Logic with Engineering Applications*, T.J. Ross (Ed. 2), 25-47.

11. Maurya, V. (2020). Chapter 5 defuzzification methods. *Debasis Samanta, Indian Institute of Technology Kharagpur*.

12. Hoerl, A.E. and Kennard, R.W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55-67.

13. Kejian, L. (1993). A new class of biased estimate in linear regression. *Communications in Statistics - Theory and Methods*, 22(2), 393-402.

14. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1), 267-288.

15. Tanaka, H., Uejima, S., and Asai, K. (1982). Linear regression analysis with fuzzy model. *IEEE Transactions on Systems, Man, and Cybernetics*, 12(6), 903-907.

16. Celmiņš, A. (1987). Multidimensional least-squares fitting of fuzzy models. *Mathematical Modelling*, 9(9), 669-690.

17. Celmiņš, A. (1987). Least squares model fitting to fuzzy vector data. *Fuzzy Sets and Systems*, 22(3), 245-269.

18. Diamond, P. (1988). Fuzzy least squares. *Information Sciences*, 46(3), 141-157.

19. Watada, J. (1992). Fuzzy time-series analysis and its forecasting of sales volume. *Fuzzy Regression Analysis*, 211-227.

20. Sakawa, M. and Yano, H. (1992). Multiobjective fuzzy linear regression analysis for fuzzy input-output data. *Fuzzy Sets and Systems*, 47(2), 173-181.

21. Moskowitz, H. and Kim, K. (1993). On assessing the H value in fuzzy linear regression. *Fuzzy Sets and Systems*, 58(3), 303-327.

22. Chang, P.T. and Lee, E.S. (1994). Fuzzy least absolute deviations regression and the conflicting trends in fuzzy parameters. *Computers & Mathematics with Applications*, 28(5), 89-101.

23. Peters, G. (1994). Fuzzy linear regression with fuzzy intervals. *Fuzzy Sets and Systems*, 63(1), 45-55.

24. Chang, P.T. and Lee, E.S. (1996). A generalized fuzzy weighted least-squares regression. *Fuzzy Sets and Systems*, 82(3), 289-298.

25. Redden, D.T. and Woodall, W.H. (1996). Further examination of fuzzy linear regression. *Fuzzy Sets and Systems*, 79(2), 203-211.

26. Kim, K.J., Moskowitz, H., and Koksalan, M. (1996). Fuzzy versus statistical linear regression. *European Journal of Operational Research*, 92(2), 417-434.

27. Kim, K.J. and Chen, H.R. (1997). A comparison of fuzzy and nonparametric linear regression. *Computers & Operations Research*, 24(6), 505-519.

28. D'Urso, P. and Gastaldi, T. (2000). A least-squares approach to fuzzy linear regression analysis. *Computational Statistics & Data Analysis*, 34(4), 427-440.

29. Ishibuchi, H. and Nii, M. (2001). Fuzzy regression using asymmetric fuzzy coefficients and fuzzified neural networks. *Fuzzy Sets and Systems*, 119(2), 273-290.

30. Chang, Y.H.O. (2001). Hybrid fuzzy least-squares regression analysis and its reliability measures. *Fuzzy Sets and Systems*, 119(2), 225-246.

31. Lee, H.T. and Chen, S.H. (2001). Fuzzy regression model with fuzzy input and output data for manpower forecasting. *Fuzzy Sets and Systems*, 119(2), 205-213.

32. Nasrabadi, M.M., Nasrabadi, E., and Nasrabady, A.R. (2005). Fuzzy linear regression analysis: A multi-objective programming approach. *Applied Mathematics and Computation*, 163(1), 245-251.

33. Buckley, J.J. (2005). Fuzzy statistics: Regression and prediction. *Soft Computing*, 9(10), 769-775.

34. Kim, K.J., Kim, D.H., and Choi, S.H. (2005). Least absolute deviation estimator in fuzzy regression. *Journal of Applied Mathematics and Computing*, 18, 649–656.

35. Zeng, J. and Zhang, A. (2018). Statistical analysis of fuzzy linear regression model based on different distances. *International Journal of Fuzzy Logic Systems*, 8(1/2), 1-13.

36. Chen, L.H. and Nien, H.H. (2020). Mathematical programming approach to formulate intuitionistic fuzzy regression model based on least absolute deviations. *Fuzzy Optimization and Decision Making*, 19(2), 191-210.

37. Belhadj, B. (2022). Fuzzy simple linear regression using gaussian membership functions minimization problem. *Journal of Fuzzy Extension and Applications*, 3(4), 279–289.

38. Hong, D.H., Hwang, C., and Ahn, C. (2004). Ridge estimation for regression models with crisp inputs and Gaussian fuzzy output. *Fuzzy Sets and Systems*, 142(2), 307-319.

39. Xu, R. (1991). A linear regression model in fuzzy environment. *Advances in Modelling Simulation*, 27, 31–40.

40. Xu, R. and Li, C. (2001). Multidimensional least-squares fitting with a fuzzy model. *Fuzzy Sets and Systems*, 119(2), 215-223.

41. Yang, M.-S. and Ko, C.-H. (1996). On a class of fuzzy c-numbers clustering procedures for fuzzy data. *Fuzzy Sets and Systems*, 84(1), 49-60.

42. Hong, D.H. and Hwang, C. (2004). Ridge regression methods for fuzzy models using triangular fuzzy numbers. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 12(02), 145-159.

43. Donoso, S., Marín, N., and Vila, M.A. (2007). Fuzzy Ridge regression with non symmetric membership functions and quadratic models. *Intelligent Data Engineering and Automated Learning - IDEAL 2007*, 135-144.

44. Bas, E., Egrioglu, E., Yolcu, U., and Grosan, C. (2019). Type 1 fuzzy function approach based on Ridge regression for forecasting. *Granular Computing*, 4(4), 629-637.

45. Rabiei, M.R., Arashi, M., and Farrokhi, M. (2019). Fuzzy Ridge regression with fuzzy input and output. *Soft Computing*, 23, 12189-12198.

46. Karbasi, D., Nazemi, A., and Rabiei, M.R. (2021). An optimization technique for solving a class of Ridge fuzzy regression problems. *Neural Processing Letters*, 53(5), 3307-3338.

47. Pişkin, M.D. (2021). *Ridge regresyona dayalı tip 1 bulanık regresyon fonksiyonları ile yabancılara yapılan aylık konut satışı öngörüsü [Forecasting monthly housing sales to foreigners using type-1 fuzzy regression functions based on Ridge regression]*. M. Sc. Thesis, Giresun University Graduate School Of Natural And Applied Sciences, Giresun.

48. Choi, S.H., Jung, H.Y., and Kim, H. (2019). Ridge fuzzy regression model. *International Journal of Fuzzy Systems*, 21(7), 2077-2090.

49. Kim, H. and Jung, H.Y. (2020). Ridge fuzzy regression modelling for solving multicollinearity. *Mathematics*, 8(9), 1-15.

50. Kocasoy, G.İ. (2023). *Ridge ve Liu regresyonda tavlama benzetimi optimizasyonu kullanılarak yanlılık parametrelerinin elde edilmesi ve bazı yanlılık parametreleri ile karşılaştırılması [Obtaining the biasing parameters using simulated annealing optimization in Ridge and Liu regression and comparing them with some biasing parameters].* M. Sc. Thesis, Gazi University Graduate School Of Natural And Applied Sciences, Ankara.

51. Akgün, G. (2017). *Bulanık veriler ile regresyon analizi için bir yaşlaşım [An approach to fuzzy regression analysis with fuzzy data].* M. Sc. Thesis, Ege University Graduate School Of Natural And Applied Sciences, Izmir.

52. Mamdani, E.S. and Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Man-Machine Studies*, 7(1), 1-13.

53. Gündüz, S.Ö. (2021). *Fuzzy regression analysis and an application.* M. Sc. Thesis, Hacettepe University Graduate School Of Natural And Applied Sciences, Ankara.

54. Çetintav, B. (2012). *Fuzzy linear regression.* M. Sc. Thesis, Dokuz Eylül University Graduate School Of Natural And Applied Sciences, Izmir.

55. Hajek, P., Godo, L., and Esteva, F. (2013). Fuzzy logic and probability. *arXiv preprint* arXiv:1302.4953.

56. Gerla, G. (1994). Inferences in probability logic. *Artificial Intelligence*, 70(1), 33-52.

57. Castillo, O. and Melin, P. (2008). *Type-2 Fuzzy Logic: Theory and Applications*. 1 ed. Studies in Fuzziness and Soft Computing. Berlin, Heidelberg: Springer, 243.

58. İçen, D. (2010). *Bulanık doğrusal regresyon analizi [Fuzzy linear regression analysis].* M. Sc. Thesis, Hacettepe University Graduate School Of Natural And Applied Sciences, Ankara.

59. Jang, J.-S.R., Sun, C.-T., and Mizutani, E. (1997). *Neurofuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Englewood Cliffs: Prentice-Hall, 614.

60. Dutta, P., Boruah, H., and Ali, T. (2011). Fuzzy arithmetic with and without using α-cut method: A comparative study. *International Journal of Lastest Trends in Computing*, 2(1), 99-107.

61. Klir, G.J. and Yuan, B. (1995). *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. New Jersey: Prentice Hall, 574.

62. Dubois, D.J. and Prade, H. (1980). *Fuzzy Sets and Systems: Theory and Applications*. 1st Edition ed. Mathematics in Science and Engineering. Vol. 144, New York: Academic Press (HBJ), 393.

63. Homaida, A., Pekalp, M.H., and Ebegil, M. (2024). Bulanık küme ve bulanık sayı: Uygulamalarla aritmetik işlemler [Fuzzy set and fuzzy number: Arithmetic operations with applications]. *Dicle University Journal of the Institute of Natural and Applied Sciences*, 13(2), 223-247.

64. Dubois, D. and Prade, H. (1993). *Fuzzy Numbers: An Overview*, in *Readings in Fuzzy Sets for Intelligent Systems*, D. Dubois, H. Prade, and R.R. Yager, Editors, Morgan Kaufmann, 112-148.

65. Zimmermann, H.-J. (2001). *Fuzzy Set Theory—and Its Applications*. 4 ed. Dordrecht: Springer, 514.

66. Mogharreban, N. and DiLalla, L.F. (2006). Comparison of defuzzification techniques for analysis of non-interval data. *NAFIPS 2006 - 2006 Annual Meeting of the North American Fuzzy Information Processing Society, Montreal, QC, Canada*, 257-260.

67. Saade, J. and Diab, H. (2004). Defuzzification methods and new techniques for fuzzy. *Iranian Journal of Electrical and Computer Engineering*, 3(2), 161-174.

68. Chandramohan, A., Rao, M.V.C., and Arumugam, M.S. (2006). Two new and useful defuzzification methods based on root mean square value. *Soft Computing*, 10(11), 1047-1059.

69. Kocatürk, Y. (2007). *Bulanık değişkenler ve bulanık yenileme süreçleri [Fuzzy variables and fuzzy renewal processes]*. M. Sc. Thesis, Ankara University Graduate School Of Natural And Applied Sciences, Ankara.

70. Stefanini, L. and Guerra, M.L. (2007). On fuzzy arithmetic operations: some properties and distributive approximations. *International Journal of Applied Mathematics*, 171–199.

71. Chandrasekaran, S. and Tamilmani, E. (2015). Arithmetic operation of fuzzy numbers using α-cut method. *IJISET - International Journal of Innovative Science, Engineering & Technology*, 2(10), 299-315.

72. Shapique, A.M. (2017). Arithmetic operations on heptagonal fuzzy numbers. *Asian Research Journal of Mathematics*, 2(5), 1-25.

73. Eljaoui, E., Melliani, S., and Chadli, L.S. (2018). Multiplication operations and powers of trapezoidal fuzzy numbers. *Journal of Universal Mathematics*, 1(2), 204-215.

74. Jayasri, P. and Elavarasi, P. (2018). Fuzzy set theory and arithmetic operations on fuzzy numbers. *International Journal of Scientific Research and Management*, 6(02), 14-16.

75. Musa, I.I.M. (2015). *Investigation of basic concepts of fuzzy arithmetic.* M. Sc. Thesis, Eastern Mediterranean University Institute of Graduate Studies and Research, Gazimağusa.

76. Ali, M.Y., Sultan, A., and Khan, A.F.M.K. (2016). Comparison of fuzzy multiplication operation on triangular fuzzy number. *IOSR Journal of Mathematics*, 12(04), 35-41.

77. Chutia, R., Mahanta, S., and Baruah, H.K. (2010). An alternative method of finding the membership of a fuzzy number. *International Journal of Latest Trends in Computing*, 1(2), 69-72.

78. Montgomery, D.C., Peck, E.A., and Vining, G.G. (2012). *Introduction to Linear Regression Analysis*. 5 ed. New Jersey: John Wiley & Sons Inc, 872.

79. Fox, J. (2015). *Applied Regression Analysis and Generalized Linear Models*. 3rd ed. USA: SAGE Publications. 816.

80. Marquardt, D.W. (1970). Generalized inverses, Ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics*, 12(3), 591-612.

81. Bishop, C.M. (2006). *Pattern Recognition and Machine Learning*. 1 ed. New York: Springer, 778.

82. Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining*. 2nd ed. Inference, and Prediction. Berlin/Heidelberg, Germany: Springer, 764.

83. Hoerl, A.E., Kennard, R.W., and Baldwin, K.F. (1975). Ridge regression: Some simulations. *Communications in Statistics*, 4(2), 105-123.

84. Lawless, J.F. and Wang, P. (1976). A simulation study of Ridge and other regression estimators. *Communications in Statistics - Theory and Methods*, 5(4), 307-323.

85. Hocking, R.R., Speed, F.M., and Lynn, M.J. (1976). A class of biased estimators in linear regression. *Technometrics*, 18(4), 425-437.

86. Schaefer, R.L., Roi, L.D., and Wolfe, R.A. (1984). A Ridge logistic estimator. *Communications in Statistics - Theory and Methods*, 13(1), 99-113.

87. Nomura, M. (1988). On the almost unbiased Ridge regression estimator. *Communications in Statistics - Simulation and Computation*, 17(3), 729-743.

88. Kibria, B.M.G. (2003). Performance of some new Ridge regression estimators. *Communications in Statistics - Simulation and Computation*, 32(2), 419-435.

89. Khalaf, G. and Shukur, G. (2005). Choosing Ridge parameter for regression problems. *Communications in Statistics - Theory and Methods*, 34(5), 1177-1182.

90. Alkhamisi, M., Khalaf, G., and Shukur, G. (2006). Some modifications for choosing Ridge parameters. *Communications in Statistics - Theory and Methods*, 35(11), 2005-2020.

91. Alkhamisi, M.A. and Shukur, G. (2007). A monte carlo study of recent Ridge parameters. *Communications in Statistics - Simulation and Computation*, 36(3), 535-547.

92. Muniz, G. and Kibria, B.M.G. (2009). On some Ridge regression estimators: An empirical comparisons. *Communications in Statistics - Simulation and Computation*, 38(3), 621-630.

93. Al-Hassan, Y.M. (2010). Performance of a new Ridge regression estimator. *Journal of the Association of Arab Universities for Basic and Applied Sciences*, 9(1), 23-26.

94. Kibria, B.M.G., Månsson, K., and Shukur, G. (2012). Performance of some logistic Ridge regression estimators. *Computational Economics*, 40(4), 401-414.

95. Khalaf, G. (2012). A proposed Ridge parameter to improve the least square estimator. *Journal of Modern Applied Statistical Methods*, 11(2), 443-449.

96. Muniz, G., Kibria, B.M.G., and Shukur, G. (2012). On developing Ridge regression parameters: A graphical investigation. *Statistics & Operations Research Transactions*, 26(2), 115-138.

97. Khalaf, G. (2013). A comparison between biased and unbiased estimators in ordinary least squares regression. *Journal of Modern Applied Statistical Methods*, 12(2), 293-303.

98. Dorugade, A.V. (2014). On comparison of some Ridge parameters in Ridge regression. *Sri Lankan Journal of Applied Statistics*, 15(1), 31-46.

99. Asar, Y., Karaibrahimoğlu, A., and Genç, A. (2014). Modified Ridge regression parameters: A comparative Monte Carlo study. *Hacettepe Journal of Mathematics and Statistics*, 43(5), 827-841.

100. Dorugade, A.V. (2014). New Ridge parameters for Ridge regression. *Journal of the Association of Arab Universities for Basic and Applied Sciences*, 15(1), 94-99.

101. Khalaf, G. and Iguernane, M. (2014). Ridge regression and III-conditioning. *Journal of Modern Applied Statistical Methods*, 13(2), 355-363.

102. Göktaş, A. and Sevinç, V. (2016). Two new Ridge parameters and a guide for selecting an appropriate Ridge parameter in linear regression. *Gazi University Journal of Science*, 29(1), 201-211.

103. Bhat, S. and Vidya, R. (2016). A comparative study on the performance of new Ridge estimators. *Pakistan Journal of Statistics and Operation Research*, 12(2), 317-325.

104. Dorugade, A.V. (2016). Improved Ridge estimator in linear regression with multicollinearity, heteroscedastic errors and outliers. *Journal of Modern Applied Statistical Methods*, 15(2), 362-381.

105. Khalaf, G. and Iguernane, M. (2016). Multicollinearity and a Ridge parameter estimation approach. *Journal of Modern Applied Statistical Methods*, 15(2), 400-410.

106. Karaibrahimoğlu, A., Asar, Y., and Genç, A. (2016). Some new modifications of Kibria's and Dorugade's methods: An application to Turkish GDP data. *Journal of the Association of Arab Universities for Basic and Applied Sciences*, 20(1), 89-99.

107. Lukman, A.F., Ayinde, K., and Ajiboye, A.S. (2017). Monte carlo study of some classification-based Ridge parameter estimators. *Journal of Modern Applied Statistical Methods*, 16(1), 428-451.

108. Lukman, A.F. and Olatunji, A. (2018). Newly proposed estimator for Ridge parameter an application to the nigerian economy. *Pakistan Journal of Statistics*, 34(2), 91-98.

109. Aytekin, H.T. (2021). *Makine öğreniminin, araştırmacıların veri analizi bağlamında potansiyel önemi [The potential importance of machine learning in the context of data analysis for researchers].* Ufuk Üniversitesi, 1. Uluslararası Sosyal Bilimler Kongresi.

110. Shukur, G., Månsson, K., and Sjölander, P. (2015). Developing interaction shrinkage parameters for the Liu estimator — with an application to the electricity retail market. *Computational Economics*, 46(4), 539-550.

111. Qasim, M., Amin, M., and Omer, T. (2020). Performance of some new Liu parameters for the linear regression model. *Communications in Statistics - Theory and Methods*, 49(17), 4178-4196.

112. Yang, M.S. and Liu, H.H. (2003). Fuzzy least-squares algorithms for interactive fuzzy linear regression models. *Fuzzy Sets and Systems*, 135(2), 305-316.

113. Chang, P.T. (1994). *Fuzzy regression analysis.* Ph. D. Thesis, Kansas State University College of Engineering, Kansas.

114. Škrabánek, P. and Martínková, N. (2019). *Getting started with fitting fuzzy linear regression models in R*. Available from: https://cran.r-project.org/web/packages/fuzzyreg/vignettes/GettingStarted.html.

115. Aydoğdu, N. (2020). *Bulanık regresyon analizi ile Aydın ili konut fiyatlarını etkileyen değişkenlerin belirlenmesi [Determining the variables affecting housing prices in aydın province through fuzzy regression analysis].* M. Sc. Thesis, Aydın Adnan Menderes University Graduate School of Social Sciences, Aydın.

116. Chang, Y.H.O. and Ayyub, B.M. (2001). Fuzzy regression methods – a comparative assessment. *Fuzzy Sets and Systems*, 119(2), 187-203.

117. Shapiro, A.F. (2004). Fuzzy regression and the term structure of interest rates revisited. *In 14th International AFIR 2004, Boston, 7 – 10 November*.

118. Bardossy, A. (1999). Note on fuzzy regression. *Fuzzy Sets and Systems*, 37(11), 65-75.

119. Hong, D.H., Song, J.K., and Do, H.Y. (2001). Fuzzy least-squares linear regression analysis using shape preserving operations. *Information Sciences*, 138(1), 185-193.

120. Mohammadi, J. and Taheri, S.M. (2004). Pedomodels fitting with fuzzy least squares regression. *Iranian Journal of Fuzzy Systems*, 1(2), 45-61.

121. Coppi, R., D'Urso, P., Giordani, P. and Santoro, A. (2006). Least squares estimation of a linear regression model with LR fuzzy response. *Computational Statistics & Data Analysis*, 51(1), 267-286.

122. D'Urso, P. and Santoro, A. (2006). Goodness of fit and variable selection in the fuzzy multiple linear regression. *Fuzzy Sets and Systems*, 157(19), 2627-2647.

123. Arabpour, A. and Tata, M. (2008). Estimating the parameters of a fuzzy linear regression model. *Iranian Journal of Fuzzy Systems*, 5(2), 1-19.

124. Kim, H.K., Yoon, J.H., and Li, Y. (2008). Asymptotic properties of least squares estimation with fuzzy observations. *Information Sciences*, 178(2), 439-451.

125. Chen, L.H. and Hsueh, C.C. (2009). Fuzzy regression models using the least-squares method based on the concept of distance. *IEEE Transactions on Fuzzy Systems*, 17(6), 1259-1272.

126. Diamond, P. and Körner, R. (1997). Extended fuzzy linear models and least squares estimates. *Computers & Mathematics with Applications*, 33(9), 15-32.

127. Kao, C. and Chyu, C.L. (2003). Least-squares estimates in fuzzy regression analysis. *European Journal of Operational Research*, 148(2), 426-435.

128. Donoso, S., Marín, N., and Vila, M.A. (2006). Fuzzy regression with quadratic programming: An application to financial data. *Intelligent Data Engineering and Automated Learning* – IDEAL 2006, 4224, 1304-1311.

129. Tanaka, H. and Lee, H. (1998). Interval regression analysis by quadratic programming approach. *IEEE Transactions on Fuzzy Systems*, 6(4), 473-481.

130. Gibbons, D.G. (1981). A simulation study of some Ridge estimators. *Journal of the American Statistical Association*, 76(373), 131-139.

131. Zhou, J., Zhang, H., Gu, Y., and Pantelous, A.A. (2018). Affordable levels of house prices using fuzzy linear regression analysis: The case of Shanghai. *Soft Computing*, 22(16), 5407-5418.

132. Kashani, M., Arashi, M., Rabiei, M.R., D'Urso, P., and De Giovanni, L. (2021). A fuzzy penalized regression model with variable selection. *Expert Systems with Applications*, 175(114696), 1-16.



Gazili olmak ayrıcalıktır