Büyük Ölçekli Doğrusal Denklem Sistemleri için Hızlı ve Gürbüz Çözüm Teknikleri
Fast and Robust Solution Techniques for Large Scale Linear System of Equations

İbrahim K. Özaslan
Elektrik ve Elektronik Mühendisliği
Bilkent Üniversitesi
Ankara, Türkiye
iozaslan@ee.bilkent.edu.tr

Mert Pilancı
Elektrik Mühendisliği
Stanford Üniversitesi
Calfornia, USA
pilanci@stanford.edu

Orhan Arıkan
Elektrik ve Elektronik Mühendisliği
Bilkent Üniversitesi
Ankara, Türkiye
oarikan@ee.bilkent.edu.tr

***Özetçe* —Büyük ölçekli doğrusal sistemlerin veri matrisi, sütunlar arası yüksek ilintiye ve genellikle yüksek durum numaralarına sahiptir. Bilinmeyenlerin, ölçümlerden En Küçük Kareler (EKK) tekniğiyle üretilmesi, ölçüm gürültüsünün, sonucu kabul edilemez şekilde etkilemesine neden olmaktadır. Bu nedenle gürbüz çözüm tekniklerine ihtiyaç duyulmaktadır. Bu bildiride, yüksek durum numarasına sahip büyük ölçekli ölçüm matrislerinin yer aldığı doğrusal sistemlerin, Momentum-Yinelemeli Hessian Krokileme (Momentum - Iterative Hessian Sketch (M-IHS)) çözücüsü kullanılarak nasıl düzenlileştirilebileceği incelenmiştir. Önerilen çözücü, tüm iterasyonlar için tek bir düzenlileştirme parametresi bulmak yerine, her bir iterasyon için düzenlileştirme parametresini başka bir parametre ayarı yapmadan otomatik olarak bulmakta ve daha sonra hızlı yaklaşım sağlayan momentum parametrelerini buna göre belirlemektedir. Yapılan analizde her ne kadar Genelleştirilmiş Çapraz Doğrulama (GCV) tekniği kullanılmış olsa da, M-IHS, bildiride açıklanan adımlar kullanılarak, herhangi bir risk tahmini ile düzenlileştirilebilir.**

Anahtar Kelimeler—En küçük Kareler Yöntemi, Tikhonov Regularizasyonu, Rastlantısal Boyut Küçültme, Momentum

***Abstract*—The data matrix of large scale linear systems generally have correlated columns and high condition numbers. Finding unknowns from measurements by using Least Square technique results in noise enhancement. For this reason, robust solution techniques are needed. In this article, we investigate how to regularize the Momentum-Iterative Hessian Sketch (M-IHS) solver for solving ill-posed linear systems including large scale data matrices. Instead of using a single regularization parameter for all iterations, the proposed solver automatically finds a separate regularization parameter in each iteration without requiring any other parameter tuning, and then adjusts momentum parameters accordingly. Although Generalized Cross Validation (GCV) technique is used in the analysis, any risk estimator can be incorporated into the steps explained in the article for the regularization of M-IHS.**

Keywords—Least Squares, Tikhonov Regularization, Random Projection, Momentum

I. GİRİŞ

Bir çok problemin çözümünün temelini oluşturan doğrusal denklem sistemleri aşağıdaki gibi modellenebilir:

$$b = Ax_0 + \omega. \tag{1}$$

Burada, $A \in \mathbf{R}^{n\times d}$ veri matrisini, $b \in \mathbf{R}^n$ gürültülü ölçüm vektörünü, $\omega \in \mathbf{R}^n$ gürültü veya ölçüm hataları vektörünü ve $x_0 \in \mathbf{R}^d$ ise verilen veri matrisi ve ölçümlerden geri elde edilmeye çalışılan bilinmeyen yani parametre vektörünü temsil etmektedir. Bu çalışmada ölçüm sayısının bilinmeyen sayısından çok daha fazla olduğu $(n \gg d)$ durum üzerine odaklanılmıştır. Bu durumda denklem sistemi gürültü sebebiyle kararsızdır ve tahmin edilen çözümün, $\hat{x}$, iyiliği teorik olarak $\|x_0 - \hat{x}\|$ (*kahin*) metriği ile belirlenmektedir; ancak bu metrik pratikte ulaşılabilir olmadığı için (2)'deki hata metriği aracılığıyla, EKK yöntemi kullanılarak, doğrusal sisteme en iyi uyan çözüm analitik olarak bulunabilir:

$$\begin{aligned} x^{\mathrm{LS}} &= \underset{x\in\mathbf{R}^d}{\operatorname{argmin}}\ \|b - Ax\|_2^2 & (2)\\ &\equiv \underset{x\in\mathbf{R}^d}{\operatorname{argmin}}\ \frac{1}{2}\|Ax\|_2^2 - < A^Tb, x > & (3)\\ &= (A^TA)^{-1}A^Tb. & (4) \end{aligned}$$

Pratik uygulamalarda oldukça yaygın olan, koşul numarasının veya sütunlar arası ilintinin yüksek olması durumunda, (4)'teki matris tersinin alınması, eğer ki A matrisinin tekil değerleri 1'den küçük ise, gürültü artırımına sebep olmaktadır. Bu artırımı engellemek için (2)'deki maliyet fonksiyonuna, $\hat{x}$'in büyük değerlerini cezalandıran yeni bir terim eklenir. Eklenen terimde l_2-normu kullanıldığı takdirde bu işlem Tikhonov Düzenlileştirmesi olarak adlandırılır ve tıpkı orijinal EKK çözümü gibi analitik olarak bulunabilir:

$$\begin{aligned} x(\lambda) &= \underset{x\in\mathbf{R}^d}{\operatorname{argmin}}\ \|b - Ax\|_2^2 + \lambda\|x\|_2^2 & (5)\\ &\equiv \underset{x\in\mathbf{R}^d}{\operatorname{argmin}}\ \frac{1}{2}\|Ax\|_2^2 - < A^Tb, x > + \frac{\lambda}{2}\|x\|_2^2 & (6)\\ &= (A^TA + \lambda I_d)^{-1}A^Tb, & (7) \end{aligned}$$

burada $\lambda \in \mathbf{R}$ düzenlileştirme parametresi olarak adlandırılır. Bu parametrenin optimal olarak seçilmesi için geliştirilen yöntemlere II. Bölüm'de detaylı olarak yer verilmiştir.

Günümüzde, büyük ölçekli veri kullanan uygulamaların da yaygınlaşmasıyla birçok yöntemin temel çekirdeğini oluşturan doğrusal denklemlerin çözümü büyük önem kazanmıştır. Ölçüm sayısının 10^6'ları aşabildiği bu uygulamalarda, çözücünün performansını belirleyen en önemli kriter çoğu zaman çözücünün optimal değere ne kadar ulaştığı değil, problemi ne kadar sürede çözdüğü ve hatta verilen işlemci hafızasının çözüm için yeterli olup olmamasıdır.

Denklem (4) veya (7)'deki çözümleri, λ verildiği takdirde elde etmenin en etkili yolu, QR veya LU ayrışımlarını kullanarak, üçgen matrislerin düşük hesaplama karmaşıklığından yararlanmaktır [1]. Fakat asimptotik karmaşıklık, kullanılan ayrışımların karmaşıklığı baskın çıktığı için $O(nd^2)$ olmaktadır ki, büyük ölçekli veri kullanan uygulamalarda bu karmaşıklık başa çıkılmaz durumdadır. Gerekli olan hesaplama miktarını düşürmenin bir yolu tüm matris-matris boyutundaki hesaplamaları uzaklaştırmak ve birinci dereceden yinelemeli çözüler kullanmaktır [2].

Conjugate Gradient (CG) [3], LSQR [4] ve Chebyshev Semiconvergence (CS) [5] yakınsama hızı yüksek ve hesaplama karmaşıklığı az olan, pratikte yaygınca kullanlan birinci dereceden yinelemeli çözücülerdendir [6]. Fakat, bu yöntemlerin yakınsama hızı, veri matrisinin spektral özelliklerine yüksek derecede bağlı olup, başarılı bir önkoşullandırma yapılmazsa oldukça düşük seviyelere inmektedir [1], [7]. Etkin bir önkoşullandırma yapmak ise hesaplama karmaşıklığını istenmeyen değerlere çıkartabilir. G benzeri birinci dereceden yinelemeli çözücülerin, etkin önkoşullandırma sorununa ek olarak, bir diğer problemi de, ölçüm vektörü b ve tahmin vektörü $A\hat{x}$ arasındaki dönüşüm ya doğrusal olmamakta yada doğrusal ise dönüşüm matriksi analitik olarak bilinememektedir [8]. Bu durum düzenlileştirme parametresinin seçimi için geliştirilen, II.Bölüm'de bahsedilen metotların başarımının azalmasına sebep olmaktadır.

Birinci dereceden yinelemeli çözücülerin dezavantajlarından kaçınan bir diğer çözüm yolu ise Rastlantısal Projeksiyon(RP)'ların kullanılmasıdır. Ranstlantısal çalışan bu boyut düşürme yöntemleri, Blendenpik ve LSRN gibi LAPACK çözücülerinden daha hızlı olduğu gösterilen birinci dereceden yinelemeli çözücülerde, önkoşullandırma matrislerinin hesaplanması için başarılı bir şekilde kullanılmaktadır [9], [10]. RP yöntemleri, doğrudan (7)'de verilen EKK çözümünün yaklaştırılması için de kullanılabilir [11]. Büyük ölçekli verinin, daha önce bahsedilen hesaplama açısından oluşturduğu problemler, pratikteki veri matrislerinin şartları ve yakınsama hızları gibi konular göz önünde bulundurulduğunda, RP'ye bağlı çözücülerin başarılı çözümler ürettiği söylenebilir. Analiz edilen M-IHS yöntemi RP'ye bağlı olup, III. Bölüm'de boyut küçültme işleminin ayrıntılarına yer verilmiştir.

II. Düzenlileştirme Parametresinin Seçimi

Literatürde bulunan Ayrım Prensibi (Discrepancy Principle (DP)) yöntemi λ parametresini, ölçüm hatasının enerjisinin ölçüm sayısına oranı gürültünün varyansına, σ_ω^2, eşit olacak şekilde seçer. DP tekniğinin λ'yı optimal değerden daha büyük seçtiği rapor edilmiştir [12], [13]. Bu durum gürültüye karşı daha gürbüz sonuçlar üretse de çözücünün başarımını düşürmektedir. Diğer bir yöntem, L-Curve (LC), λ'yı $\|Ax(\lambda)-b\|_2^2$ ve $\|x(\lambda)\|_2^2$ terimlerinin oluşturduğu Pareto Eğrisinin eğriliğinin en yüksek olduğu değer olarak seçmektedir [12]. Bir diğer metot, Genelleştirilmiş Stein Tarafsız Risk Tahmini (GSURE) ise $\|x_0-\hat{x}(\lambda)\|$ metriğinin yansız bir kestiricisini bularak bu kestiriciyi en aza indiren λ değerini seçmektedir [14]. GSURE ve LC yöntemleri λ'yı, DP'nin aksine, olması gerekenden küçük seçmektedir ki bu gürültünün büyütülerek çözümü etkilemesine neden olmaktadır [15], [13]. GSURE yöntemi de, gürültünün varyasyonunun bilindiğini varsaymaktadır.

GSURE'de kullanılan $\|x_0-\hat{x}(\lambda)\|$ metriği yerine $\|Ax_0-A\hat{x}(\lambda)\|$ metriğinin yansız kestiricisini kullanmak da mümkündür. Bu kestiriciyi en aza indiren λ değerini seçen UPRE adı altında toplayabileceğimiz yöntemler, GSURE ve DP gibi gürültü hakkında istatiksel bilgi gerektirmekte olup, şimdiye kadar bahsedilen üç yöntemden de daha gürbüz sonuçlar üretmektedir [15]. Son olarak, pratikte yaygın bir biçimde kullanılan, Çapraz Geçerlilik yöntemlerinin rotasyon-değişimsiz hali olan Geneleştirilmiş Çapraz Geçerlilik (GCV) yöntemi, hem gürültü varyasyonu gibi herhangi ek bir bilgi gerektirmemekte hem de asimptotik olarak optimal λ değerine yakınsamaktadır [16]. GCV yöntemi aynı zamanda RP uygulamalarında oldukça kararlı sonuçlar vermiştir. RP'ye dayalı M-IHS çözücüsü, bahsedilen tüm düzenlileştirme yöntemleri ile kullanılabilir. Fakat diğerlerinden daha kararlı sonuçlar üretmesinden ve daha az bilgi gerektirmesinden dolayı, analizlerde GCV yönteminin kullanımı tercih edilmiştir. GCV yöntemi λ parametresini (8)'daki maliyet fonsiyonunu en aza indiren λ olarak seçmektedir:

$$G(\lambda)=\frac{\|b-Ax(\lambda)\|^2}{\left[n-\mathbf{Tr}\left(P_A(\lambda)\right)\right]^2}, \tag{8}$$

burada $x(\lambda)$ λ'ya bağlı çözümü, $P_A(\lambda)$ ise ölçüm vektörü b ve tahmin vektörü $Ax(\lambda)$ arasındaki doğrusal dönüşümün λ'ya bağlı olan matrisini temsil etmektedir. Denklem (7)'de görülen EKK çözümü için $P_A(\lambda)=A(A^TA+\lambda I_d)A^T$ olmaktadır. Pratikte GCV fonksiyonu Tekil Değer Ayrışımı(SVD) aracılığıyla en aza indirgenmektedir.

III. Rastlantısal Projeksiyon İle Boyut Küçültme

Rastlantısal boyut düşürme teknikleri Johnson-Lindenstrauss (JL) lemmasına dayalı olarak geliştirilmiştir [17]. EKK çözümü üzerinde kullanılan iki türü mevcutur. İlki, Basit Kroki olarak adlandırılan yöntemlerdir, ve (A,b) çifti yerine, (SA,Sb) çiftinin gözlenmesine dayanmaktadır. Burada $S\in\mathbf{R}^{m\times n}$, JL lemmasının yüksek olasılıkla geçerli olmasını sağlayan, rastantısal olarak oluşturulan ve *Krokileme Matrisi* olarak adlandırlan boyut düşürme dönüşüm matrisidir [18]. Denklem sisteminin çözümü, (4) veya λ'nın verildiği durumda (7)'de görülen maliyet fonksiyonu yerine, (9)'da görülen maliyet fonksiyonu en aza indirgenerek elde edilir.

$$x^{\mathrm{RP}}=\underset{x\in\mathbf{R}^d}{\operatorname{argmin}}\ \|Sb-SAx\|_2^2+\lambda\|x\|_2^2 \tag{9}$$

$$=\left((SA)^T(SA)+\lambda I_d\right)^{-1}(SA)^TSb. \tag{10}$$

Bu yöntemin *sub-optimal* olduğu [19]'de gösterilmiştir. İkinci RP türü ise, veri matrisi ile beraber ölçüm vektörünün de boyutunu düşürmekten kaynaklanan yapay gürültü artırımını engellemek ve daha küçük Kroki matrisinin kullanımına olanak sağlamak amacıyla, (SA, Sb) yerine $(SA, A^T b)$ çiftini gözlemlemektir. Bunu gerçekleştirmenin bir yolu Hessian Krokileme (Hessian Sketch HS) tekniğinin kullanılmasıdır. Bu yöntem, (9)'u kullanmak yerine, (3) ve (6)'da açıkça görülebilen karesel normu, (11)'de görüldüğü üzere Kroki matrisini kullanarak yaklaşık olarak hesaplamaktadır.

$$x(\lambda) = \underset{x\in\mathbf{R}^d}{\text{argmin}}\ \frac{1}{2}\|SAx\|_2^2 - < A^T b, x > + \frac{\lambda}{2}\|x\|_2^2 \tag{11}$$

$$= \left((SA)^T(SA) + \lambda I_d\right)^{-1} A^T b. \tag{12}$$

Ancak, yalnızca karesel normun RP kullanılarak bir kere yaklaştırılması, HS tekniğinin eniyi sonuca ulaşması için yeterli olmamaktadır. Bu yüzden HS tarafından üretilen çözümün başarımı, Newton yöntemine benzer bir şekilde yapılan yinelemeler sayesinde artırılır. Bu yöntem, "*Heavy Ball*" yöntemi ile birleştirildiğinde M-IHS yöntemi elde edilmektedir [7]. M-IHS yönteminde tek bir Kroki matrisi tüm yinelemeler için kullanılmaktadır ve bu yöntemin yakınsama hızı, CG benzeri tekniklerin aksine, veri matrisinin spektral özelliklerine bağlı değildir. M-IHS yönteminin iterasyonları (13)'te görülebilir:

$$x^{\text{k+1}} = x^{\text{k}} + \alpha(A^T S^T SA)^{-1}A^T(b - Ax^{\text{k}}) + \beta(x^{\text{k}} - x^{\text{k-1}}), \tag{13}$$

burada $\alpha = (1-d/m)^2$ ve $\beta = d/m$ momentum parametreleridir [20]. Bu bildiride, M-IHS yönteminin, GCV tekniği kullanarak nasıl düzenlileştirilmesi gerektiği ve eniyi momentum parametrelerinin bu durum için nasıl tahmin edilmesi gerektiği gösterilecektir.

IV. Önerilen Düzenlileştirme Yöntemi

Denklem (1)'de verilen doğrusal denklem sisteminin, λ'nın sağlandığı durumda, aşağıdaki yineleme kullanılarak çözülmesi önerilmektedir:

$$x^{\text{k+1}}(\lambda) = x^{\text{k}}(\lambda) + \alpha_\lambda(A^T S^T SA + \lambda I)^{-1}(A^T(b - Ax^{\text{k}}(\lambda)) - \lambda x^{\text{k}}) + \beta_\lambda(x^{\text{k}}(\lambda) - x^{\text{k-1}}(\lambda)), \tag{14}$$

burada, optimal momentum parametreleri, orjinal M-IHS tekniğinden farklı olarak, $\alpha_\lambda = (1-r_\lambda)^2$, $\beta_\lambda = r_\lambda$ şeklinde kestirilmiştir. Burada r_λ, λ'ya göre belirlenen etkin kerte ve kroki boyutu m arasındaki orandır:

$$r_\lambda = \frac{\sum_{i=1}^d \frac{\sigma_i^2}{\sigma_i^2+\lambda}}{m},$$

σ_i A matrisinin tekil değerleridir. Düzgelenmiş M-IHS tekniğinin yakınsama hızı $\sqrt{r_\lambda}$ olarak bulunabilir. Düzenlileştirme parametresi verilmediği durumlarda, GCV kullanılarak λ bulunabilir. Denklem (8)'de görülen dönüşüm matrisi $P_A(\lambda)$, tüm yinelemeler için ortak olarak yazılamayacağı için her yinelemede yeni bir λ bulma yolu seçilmiştir. GCV fonksiyonunun payında yer alan $x(\lambda)$ için (14) kullanılırken, $P_A(\lambda)$ şu şekilde yaklaşık olarak hesaplanmaktadır:

$$P_{\tilde{A}}(\lambda) = WA(A^T S^T SA + \lambda I_d)^{-1}A^T W^T$$

$$\lambda_k = \underset{\lambda\in\mathbf{R}}{\text{argmin}}\ G(\lambda) = \underset{\lambda\in\mathbf{R}}{\text{argmin}}\ \frac{\|b - Ax(\lambda)\|^2}{\left[n - \mathbf{Tr}\left(P_{\tilde{A}}(\lambda)\right)\right]^2}. \tag{15}$$

Burada $W \in \mathbf{R}^{m\times d}$ ikinci bir Kroki matrisidir. Önerilen yöntemin, tüm veriyi kullanan EKK-GCV çözümlerinden daha avantajlı oluşu $x(\lambda)$ ve $P_A(\lambda)$ hesaplanırken, A yerine boyutu küçültülmüş SA ve WA matrisinin kullanılmasından kaynaklanmaktadır. Kroki boyutu tipik olarak $2d - 8d$ arasında seçilmektedir. Kroki matrisinin kullanımı asimptotik karmaşıklığı $O(nd^2)$'den $O(md^2 + nd\log(m))$'ye düşürmektedir. VI. Bölüm'de görülebileceği gibi, $n \gg d$ olan uygulamalarda, bu fark önemli bir kazanç sağlamaktadır.

V. Sayısal Algoritma

Önerilen algoritmanın genel hali Algoritma 1'de görülebilir. Burada *sketch(·)* kroki matrisini oluşturma algoritmasıdır, ayrıntılar [11], [18]'da bulunabilir. Benzetimlerde Kroki matrisi üretilirken, Rastlantısal Birim Dikgen Sistemler(ROS) kullanılmıştır. 10. adımdaki eniyileme Altın Oran Arama ve Ardışık Parabol İnterpolasyonu algoritmalarının karışımı olan ALGOL 60 prosedürü ile gerçekleştirilebilir [21].

Algorithm 1 Düzgelenmiş M-IHS

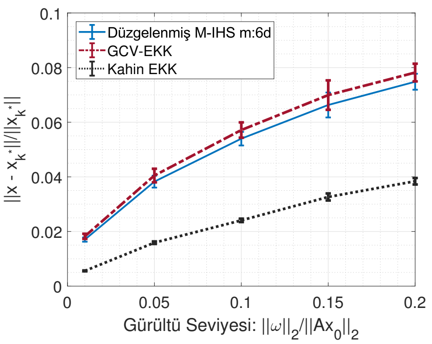
Parameters: X_{TOL}, m
Data: $A \in \mathbf{R}^{n\times d}$, b, x_0 : ilk tahmin
1: $[SA, WA]$= (sketch$(A, m))/m$
2: $[\Sigma_s, V_s]$= svd(SA)
3: $\lambda_0 = \sigma_{s,1}$
4: $k = 0$
5: $x_1 = x_0$
6: $x_0 = \underline{0}$
7: **while** $||x_k - x_{k-1}|| \geq dX_{TOL}$ **and** $k < 10 + \lceil\log n\rceil$ **do**
8: $k = k + 1$
9: $g = A^T(b - Ax_k)$
10: $\lambda_k = \underset{\lambda}{\arg\min}\ G(\lambda)$ (15)'de tanımlandığı gibi
11: $d_\lambda = \sum_{i=1}^d \frac{\sigma_{s,i}^2}{\sigma_{s,i}^2+\lambda}$
12: $\beta_\lambda = d_\lambda/m$
13: $\alpha_k = (1-\beta_\lambda)^2$
14: $\Delta x_k = V_s(\Sigma_s^2 + \lambda I)^{-1}V_s^T(g - \lambda_k x_k)$
15: $x_{k+1} = x_k + \alpha_k\Delta x_k + \beta_k(x_k - x_{k-1})$
16: **end while**

VI. Benzetim Sonuçları

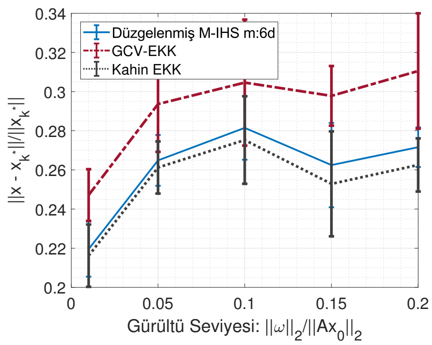
Önerilen yöntemin başarımı, Hansen tarafından yayınlanan Düzenlileştirme Araçkutusu [12] kullanılarak MATLAB dilinde hazırlanan benzetim ortamında incelenmiştir. Bu amaçla veri matriksi $A \in \mathbf{R}^{65536\times1000}$, $\mathcal{N}(1_d, \Sigma)$ olasılık dağılımından örneklenmiştir. Kovaryans matrisi Σ'nın girdileri $\Sigma_{ij} = 5\cdot 0.9^{|i-j|}$ şeklinde belirlenmiştir. Bu sayede sütunların ilintili olması sağlanmıştır. Daha sonra bu matrisin tekil değerleri Hansen'nin *baart, philiphs*, ve *heat* tekil değer profilleri ile değiştirilmiştir. Tüm profillerde durum numarası 10^6 olacak şekilde ölçeklendirme kullanılmıştır. Tekil değerlerin azalma oranlarına göre, bu profiller ileri, orta ve düşük seviyeli yüksek durum numaralı problem olarak sınıflandırılabilir [22]. Algoritmaların başarımı normalleştirilmiş etkin hata metriği, $\|\hat{x} - x_k^*\|_2/\|x_k^*\|_2$ baz alınarak ölçülmüştür. Burada x_k^* şu şekilde tanımlanır:

$$k^* = \underset{k}{\text{argmin}} \left\|\sum_{i=1}^k \frac{u_i^T b}{\sigma_i}v_i - x_0\right\|_2, \quad x_{k^*} = \sum_{i=1}^{k^*}(v_i^T x_0)v_i.$$

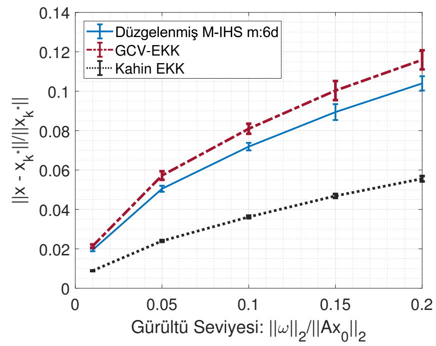
Burada, u_i sol tekil vektörler, v_i sağ tekil vektörlerdir. Farklı gürültü ve hastalık seviyeleri için elde edilen hata grafikleri Şekil 1'de görülebilir.



(a) İleri derece



(b) Orta Derece



(c) Düşük Derece

Şekil 1: Önerilen yöntemin çeşitli gürültü seviyesinde elde ettiği sonuçlar. Her bir sonuç 10 MC benzetimi üzerinden ortalama alınarak elde edilmiştir. Dikey çizgiler bir standart sapmanın miktarını göstermektedir. *GCV-EKK*, (8)'deki fonksiyonu en aza indiren λ değerini, *Kahin EKK* ise *kahin* metriğini, $\|x_k^* - \hat{x}(\lambda)\|_2$, en aza indiren λ değerini kullanarak (7)'deki denklem ile çözümü üretmektedir. Üç farklı derecede de, önerilen düzenlileştirme yöntemi ile *M-IHS*, verinin tamamının kullanıldığı *EKK-GCV* çözümünden daha iyi sonuç vermiştir. *GCV-EKK* yöntemi bir çözümü ortalama 43.32 saniyede hesaplarken, önerilen yöntem 15.62 saniyede hesaplamıştır. Süreler *tictoc* fonsiyonu ile tutulmuştur.

VII. Sonuç

Bu bildiride, veri matrisin boyutlarının yüksek olduğu doğrusal denklem sistemlerini çözmek için kullanılan, birinci dereceden yinelemeli çözülerin aksine, yakınsama hızı veri matrisinin spektral özelliklerine bağlı olmayan M-IHS yönteminin, Tikhonov Düzenlileştirme tekniği kullanılarak gürültüye karşı nasıl gürbüz hale getirileceği incelenmiştir. M-IHS yönteminde momentum parametreleri tüm iterasyonlarda sabittir. Ancak bu yönteme düzenlileştirme uygulanırken, her iterasyonda GCV tekniği kullanılarak yeni bir düzenlileştirme parametresi tahmin edilmiş ve bu parametreye bağlı olarak momentum parametreleri de değiştirilmiştir. Önerilen teknik ile düzenlileştirilen M-IHS yönteminin, standart bir dizüztü bilgisayarda yapılan benzetim sonuçlarında, GCV kullanılarak düzenlileştirilmiş EKK yönteminden ortalama üç kat daha hızlı olduğu görülmüştür. Ayrıca, önerilen teknik tüm yüksek durum numaralarında, GCV-EKK'den daha az hata vermiştir.

Kaynaklar

[1] Å. Björck, *Numerical methods in matrix computations*. Springer, 2015, vol. 59.

[2] S. Wright and J. Nocedal, "Numerical optimization," *Springer Science*, vol. 35, no. 67-68, p. 7, 1999.

[3] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*. NBS Washington, DC, 1952, vol. 49, no. 1.

[4] C. C. Paige and M. A. Saunders, "Lsqr: An algorithm for sparse linear equations and sparse least squares," *ACM Transactions on Mathematical Software (TOMS)*, vol. 8, no. 1, pp. 43–71, 1982.

[5] M. H. Gutknecht and S. Röllin, "The chebyshev iteration revisited," *Parallel Computing*, vol. 28, no. 2, pp. 263–283, 2002.

[6] R. Barrett, M. W. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst, *Templates for the solution of linear systems: building blocks for iterative methods*. Siam, 1994, vol. 43.

[7] I. K. Ozaslan, M. Pilanci, and O. Arikan, "Iterative hessian sketch with momentum," *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019.

[8] P. Favati, G. Lotti, O. Menchi, and F. Romani, "Generalized cross-validation applied to conjugate gradient for discrete ill-posed problems," *Applied Mathematics and Computation*, vol. 243, pp. 258–268, 2014.

[9] H. Avron, P. Maymounkov, and S. Toledo, "Blendenpik: Supercharging lapack's least-squares solver," *SIAM Journal on Scientific Computing*, vol. 32, no. 3, pp. 1217–1236, 2010.

[10] X. Meng, M. A. Saunders, and M. W. Mahoney, "Lsrn: A parallel iterative solver for strongly over-or underdetermined systems," *SIAM Journal on Scientific Computing*, vol. 36, no. 2, pp. C95–C118, 2014.

[11] P. Drineas, M. W. Mahoney, S. Muthukrishnan, and T. Sarlós, "Faster least squares approximation," *Numerische mathematik*, vol. 117, no. 2, pp. 219–249, 2011.

[12] P. C. Hansen, "Regularization tools: A matlab package for analysis and solution of discrete ill-posed problems," *Numerical algorithms*, vol. 6, no. 1, pp. 1–35, 1994.

[13] F. Lucka, K. Proksch, C. Brune, N. Bissantz, M. Burger, H. Dette, and F. Wübbeling, "Risk estimators for choosing regularization parameters in ill-posed problems-properties and limitations," *Inverse Problems & Imaging*, vol. 12, no. 5, pp. 1121–1155, 2018.

[14] Y. C. Eldar, "Generalized sure for exponential families: Applications to regularization," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 471–481, 2009.

[15] C. R. Vogel, *Computational methods for inverse problems*. Siam, 2002, vol. 23.

[16] G. H. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, no. 2, pp. 215–223, 1979.

[17] W. B. Johnson and J. Lindenstrauss, "Extensions of lipschitz mappings into a hilbert space," *Contemporary mathematics*, vol. 26, no. 189-206, p. 1, 1984.

[18] M. Pilanci and M. J. Wainwright, "Randomized sketches of convex programs with sharp guarantees," *IEEE Transactions on Information Theory*, vol. 61, no. 9, pp. 5096–5115, 2015.

[19] ——, "Iterative hessian sketch: Fast and accurate solution approximation for constrained least-squares," *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 1842–1879, 2016.

[20] B. T. Polyak, "Some methods of speeding up the convergence of iteration methods," *USSR Computational Mathematics and Mathematical Physics*, vol. 4, no. 5, pp. 1–17, 1964.

[21] R. P. Brent, *Algorithms for minimization without derivatives*. Courier Corporation, 2013.

[22] S. Gazzola, P. C. Hansen, and J. G. Nagy, "Ir tools: a matlab package of iterative regularization methods and large-scale test problems," *Numerical Algorithms*, pp. 1–39, 2019.