S.Ü. Müh.-Mim. Fak. Derg., c.21, s.3-4, 2006
J. Fac.Eng.Arch. Selcuk Univ., v.21, n.3-4, 2006

ROBUST YÖNTEMLERLE UYUŞUMSUZ ÖLÇÜLERİN BELİRLENMESİ

Cevat İNAL, Mevlüt YETKİN
S. Ü. Müh. - Mim. Fak. Jeodezi ve Fotogrametri Müh. Böl., KONYA.
cevat@selcuk.edu.tr, myetkin@selcuk.edu.tr

ÖZET: Bu çalışmada, jeodezide en yaygın parametre kestirim yöntemi olan En Küçük Kareler Yöntemini tamamlayıcı kestiriciler örneğin M kestiriciler, BIBER kestiricisi, iteratif ağırlıklandırma yöntemi, Berberan yöntemi ve korelasyonlu ölçüler için bifaktör indirgeme modeli gibi robust yöntemler ve ek olarak robust – ridge kestiricisi incelenmiştir. Bunun yanında dengeleme öncesi veya sonrası kaba hatalı ölçülerin belirlenmesinde kullanılan test yöntemleri, güvenirlik, kırılma noktası ve kaldıraç noktası gibi kavramlar irdelenmiştir. Son olarak basit bir lineer regresyon modelinde çeşitli örnekler verilmiştir.

Anahtar kelimeler: Uyuşumsuz Ölçü, Klasik Uyuşumsuz Ölçü Testleri, Robust Yöntemler, Kırılma Noktası

Robust Methods for the Detection of Outliers

ABSTRACT: In this study, complementary estimators to Least Squares Method which is the most widely used estimation technique in geodesy, especially robust estimators such as M estimators, BIBER Estimator, iterative weighting method, Berberan Method and bifactor reduction model for correlated observations are studied. Furthermore, pre – adjustment and post – adjustment tests for outliers, reliability, breakdown point and leverage point are described. Finally, various examples in a simple linear regression model are given.

Key words: Outlier, Conventionel Tests for Outliers, Robust Methods, Breakdown Point

GİRİŞ

Jeodezide en yaygın olarak kullanılan kestirim yöntemi En Küçük Kareler Yöntemi (EKKY)'dir. Jeodezik ağlarda yapılan ölçülerin dengelenmesinde bu yöntemden yararlanılır. (EKKY); hesap algoritmasının basit oluşu, gözlemlerle ilgili istatistik dağılımların bilinmesine gerek duyulmaması, fonksiyonel ve stokastik modellerin başlangıçtaki değerlerini koruması, varyans – kovaryans dağılımı ve hata istatistiği yönünden basit ve anlaşılır olması gibi çeşitli nedenlerle lineer Gauss – Markoff yapısındaki istatistik modellerin çözümünde yaygın bir şekilde kullanılmaktadır. (Dilaver v.d. 1998).

Ölçülerde kaba veya sistematik hata yapılmadığı sürece rasgele ölçü hataları belirli bir standart sapmaya göre normal dağılımdadır. Ancak, çoğu kez ölçüyü yapan kişinin dikkatsizliği, yanlış okuma, yanlış kayıt, yanlış hedefleme, yanlış merkezleme veya bilgisayara veri girişi sırasında ölçü presizyonuna göre oldukça büyük hatalar yapılabilmektedir. Genellikle 3σ ve daha büyük hatalar kaba hata olarak tanımlanmaktadır. Kaba hatalı ölçüler ise uyuşumsuz ölçülerdir. Sanılanın aksine kaba hatalı ölçüler de rasgele hatalı ölçüler gibi normal dağılımlı olabilir. Sorun rasgele ölçü hatalarının geldiği dağılımların farklı olmasıdır. Bu nedenle ölçü hataları normal dağılım da olmaz, bu durumda ölçü hatalarımızın dağılımı karışık normal dağılım veya kirletilmiş normal dağılım olarak adlandırılmaktadır. Bu dağılımın şekli çan eğrisine benzese de olasılık yoğunluk fonksiyonu normal dağılımdan farklı olur (Wilcox 2001). Doğal olarak böyle bir dağılımdaki ölçülerle EKKY ile çözüm yapıldığında yanlış değerler elde edilir. Parametre kestiriminde, kestirilen değerle gerçek değer arasındaki fark *kayıp* olarak adlandırılır (Gross 2003). Amaç bu kaybı minimum yapmaktır. İşte ölçü hataları normal dağılımda olduğu zaman EKKY ile parametre kestiriminde bu farkın minimum olmasına *minimum varyans* ilkesi adı verilmektedir. EKKY ile parametre kestiriminin en iyi taraflarından biri budur.

EKKY ile parametre kestiriminde dikkat edilmesi gereken bir diğer noktada rasgele hataların varyanslarının belli bir parametreye göre değişip değişmemesidir. Eğer rasgele ölçü hatalarının hepsinin standart sapması belli bir normal dağılımdan geliyorsa yani varyansları eşit ise bu durum homoskedastiklik olarak adlandırılır. Aynı alet ve ölçücü tarafından yapılan doğrultu gözlemleri buna örnek gösterilebilir. Eğer rasgele hataların varyansı belli bir parametreye göre değişiyorsa diğer bir deyişle varyansları farklı ise bu durum heteroskedastiklik olarak adlandırılır. Nivelman ağlarında, kenar ağlarında ve GPS ağlarındaki durum ise budur. Yine doğruluğu farklı aletlerin kullanıldığı doğrultu ağlarında da heteroskedastiklik söz konusudur. Bazen, doğrultu – kenar ağlarında olduğu gibi heteroskedastiklik ve homoskedastiklik durumları ile aynı anda karşılaşılabilir bu durumda stokastik model heterojen olur (Erenoğlu 2003; Hekimoğlu ve Berber 2003). Stokastik modelin yanlış kurulması bilinmeyenler kadar birim ağırlıklı varyansında yanlış belirlenmesine neden olur. Bu durumda sonsal varyansa dayalı testler (Global test, Pope testi) geçerli olmaz. Bu nedenle olağan EKKY ile parametre kestirimi yerine ağırlıklı EKKY ile parametre kestirimi yapılmalıdır (Wetherill 1986).

EKKY ile parametre kestiriminde dikkat edilecek bir diğer nokta ise lineer bağımlılıktır (colinearity). A katsayılar matrisinin sütunları arasında yakın bir lineer bağımlılık olması durumda katsayılar matrisinden hesaplanacak N normal denklem katsayılar matrisi kötü kondisyonlu bir matris olur, yani bu matris elemanlarındaki çok küçük bir değişiklik inversin önemli ölçüde değişmesine neden olur. Bu durum doğal olarak normal denklem matrisinin inversinden hesaplanan bilinmeyenlerin varyans – kovaryans matrisinin hatalı olmasına yol açar. Bilindiği gibi jeodezik ağlarda kalite kontrolünde bu matristen yararlanılmaktadır. Lineer bağımlılığa karşı *ridge kestiriciler* önerilmiştir. Bazen lineer bağımlılıkla kaba hata problemi bir arada karşımıza çıkabilir. Bu durumda ridge kestiricilerle robust kestirimin bir arada kullanılması önerilir (Gross 2003).

UYUŞUMSUZ ÖLÇÜ KAVRAMI

Robust istatistikte ölçüler iyi ölçüler ve kötü ölçüler olmak üzere ikiye ayrılmaktadır. İyi ölçüler rasgele hatalı, kötü ölçüler ise kaba hatalı yani uyuşumsuz ölçülerdir. Diğer bir deyişle iyi ölçüler, aynı dağılımdan (normal dağılımdan) kötü ölçüler ise başka bir dağılımdan gelmektedir (Hekimoğlu 1998). Robust istatistikte uyuşumsuzlar ölçü uzayında ortaya çıkan uyuşumsuzlar (kaba hatalı ölçüler) ve tasarım uzayında ortaya çıkan uyuşumsuzlar (kaldıraç noktaları) olmak üzere ikiye ayrılmaktadır (Hekimoğlu 2006). Ölçü uzayındaki uyuşumsuzlara karşı Robust yöntemler, örneğin M kestiriciler, kullanılabilir. Kaldıraç noktalarındaki uyuşumsuzları belirlemek içinse Robust yöntemlerden genelleştirilmiş M Kestiriciler, BIBER kestiricisi veya LMS (Least Median Squares) gibi yüksek kırılma noktalı robust yöntemler tercih edilmelidir (Hekimoğlu 2005).

Ölçü uzayındaki uyuşumsuz ölçüler rasgele uyuşumsuz ölçüler ve ortak etkilenmiş uyuşumsuz ölçüler olarak sınıflandırılabilir. Rasgele uyuşumsuz ölçüler, ölçülerimizde tesadüfi bir şekilde yapılan kaba hatalar nedeniyle oluşur. Ortak etkilenmiş uyuşumsuz ölçüler ise belli bir kaynaktan etkilenmiş bir grup uyuşumsuz ölçüdür. Bu ölçülerdeki hatalar aynı yönlüdür (Berber 1997). Dolayısıyla bu kirletilmiş ölçüler de normal dağılımda değildir.

Uyuşumsuz ölçülerin bir diğer sınıflandırması ise konumlarına göre yapılır. Örneğin aynı bir istasyon noktasında yapılan doğrultu ölçülerinde birden fazla kaba hata yapılmışsa bu ölçüler komşu uyuşumsuz ölçülerdir. Eğer farklı istasyon noktalarında kaba hata yapılmış ise bu da rasgele dağılmış uyuşumsuz ölçüler olarak adlandırılır. Peki komşuluğun önemi nedir? Komşu ölçülerin kısmi redundansları birbirine yakındır. Kısmi redundans uyuşumsuz ölçü belirlemede en önemli kavramlardan biridir. Kısmi redundanslar bir ölçüdeki rasgele hatanın sözgelimi kaba hatanın o ölçünün düzeltmesine yansıma yüzdesini verir. Gerek uyuşumsuz ölçü belirleme testlerinde gerekse de robust yöntemlerde incelenen temel veri ise ölçülerin düzeltmeleridir. Bu nedenle ölçülerdeki hataların mümkün olduğunca düzeltmelerine yansıması istenir. Bir ölçüdeki hata düzeltmesine ne kadar fazla yansırsa o kadar kolay belirlenir. Bu nedenle jeodezik ağ tasarımında mümkün olduğunca homojen ve yüksek (0.5'in üzerinde) kısmi redundansları sağlayacak ağlar tasarlanmalıdır. Komşu ölçülerde kaba hata yapılmışsa bunları belirlemek, kısmi redundansları birbirine yakın olacağı için rasgele dağılmış uyuşumsuz ölçülere göre daha zordur (Berber 1997).

GÜVENİRLİK

Uyuşumsuz ölçü belirlemede en önemli kavramlardan birisi güvenirliktir. Güvenirlik bir ağın gözlemlerdeki kaba hatalara karşı koyma gücü olarak tanımlanabilir. Güvenirlik konusunda iç ve dış güvenilirlik ayrımı yapılır. İç güvenirlik belli bir güven seviyesi $(1-\alpha)$ ve test gücüyle $(1-\beta)$ hipotez testleri yaparak belirlenecek minimum kaba hata sınır değeriyle ilgilidir. Hipotez testlerinde doğru bir hipotezin yanlışlıkla reddedilmesi 1. tip hata olarak adlandırılır ve olasılığı α dır. Yanlış bir hipotezin kabul edilmesi ise 2. tip hata olarak adlandırılır. Bu hatanın olasılığı ise β dır. Yani iç güvenirlik ölçülerin kontrol edilebilirliği ile ilgili olarak ağa uygulanan uyuşumsuz ölçü belirleme tekniğinin duyarlılığını yansıtır. Belli bir sınır değerin altında kalan kaba hatalar belirlenemez. Dış güvenirlik ise ağın kendisinin ölçülerdeki kaba hatalara karşı duyarlılığı ile ilgilidir. Dış güvenirlik konusunda belirlenemeyen kaba hataların parametre kestirimindeki etkisinin mümkün olduğunca küçük olması istenir. Kısmi redundans sayılarının olabildiğince büyük olması yüksek bir iç ve dış güvenirlik sağlayacaktır (Kuang 1996).

UYUŞUMSUZ ÖLÇÜLERİN BELİRLENMESİ

Günümüzde uyuşumsuz ölçülerin belirlenmesinde geleneksel test yöntemleri ve robust yöntemler olmak üzere iki yaklaşım vardır. Her iki yaklaşımda da temel veri ölçülerin düzeltmeleridir. Bilindiği gibi dengelemeyle bulunan ölçü düzeltmeleri korelasyonludur, EKKY'nin yayma etkisi nedeniyle her hangi bir ölçüdeki hata diğer ölçülerin düzeltmelerine de yansıyabilmektedir. Bu durum gerek test yöntemleriyle gerekse de robust yöntemlerle bazen yanlış sonuçlar alınabilmesine sebep olabilmektedir. Sözgelimi robust yöntemlerde ilk iterasyon da düzeltmesi standart sapmasının üç katından fazla olmayan ölçünün ağırlığının değiştirilmemesi önerilir (Leick 1995). Dengeleme öncesi lup kapanmaları kontrol edilerek veya tekrarlı ölçüler yapılarak özellikle büyük kaba hatalı ölçüler test yöntemleriyle belirlenebilir. Küçük kaba hatalı ölçüler ise dengeleme sonrası bulunan düzeltmelerin hipotez testleriyle analizi ile belirlenmektedir (Kuang 1996).

Klasik Uyuşumsuz Ölçü Testleri

EKKY ile parametre kestiriminin en önemli özelliklerinden biri dengeleme sonrası etkin istatistik analizlere imkan vermesidir. Örneğin dengeleme sonucu bulduğumuz varyansın önsel varyansla istatistiksel olarak uyuşumlu olup olmadığı global testle, hangi ölçünün kaba hatalı olduğu ise klasik uyuşumsuz ölçü belirleme testleriyle belirlenebilir (Kuang 1996). Bu amaç için kullanılan yöntemler;

- Baarda yöntemi (Global Test + Data Snooping)
- Pope testi
- t testi

Bu yöntemlerle güvenilir olarak ancak bir tane uyuşumsuz ölçü belirlenebilmektedir (Berber 1997). EKKY'nin yayma etkisinden düzeltmeler önemli ölçüde etkilenmektedir. Yani dengeleme ile herhangi bir ölçünün düzelmesi oluşurken bu düzeltme sadece o ölçüde yapılan hata nedeniyle oluşmaz, izdüşüm matrisinin diagonal olmayan elamanlarının çarpanı kadar herhangi bir ölçüdeki hata bütün düzeltmeleri etkiler. İzdüşüm matrisi ise ağın geometrik ve stokastik yapısını yansıttığından ağ tasarımında uygun bir izdüşüm matrisinin oluşturulmasına dikkat edilmelidir. Test yöntemleriyle uyuşumsuz ölçü araştırması yaparken standartlaştırılmış düzeltmesi en büyük olan ölçünün test istatistiği kritik değeri aşıyorsa uyuşumsuz ölçü olarak değerlendirilir. Yani diğer ölçülerin standartlaştırılmış düzeltmesi kritik değeri aşsa bile uyuşumsuz ölçü olarak değerlendirilmezler. Çünkü iteratif çözümde her iterasyon da standartlaştırılmış düzeltmesi en büyük olan ölçü, ölçü kümesinden atılacağı için yayma etkisi azalacak böylece ilk iterasyonda kritik değeri aşan iyi ölçü artık doğru bir şekilde kritik değerin altına düşebilecektir. Bu yöntemler arasındaki temel fark ölçülerin standartlaştırılmasında farklı varyans faktörlerinin kullanılmasıdır. Örneğin Baarda yönteminde önsel varyans, Pope yönteminde sonsal varyans, t testinde ise düzeltmesi en büyük olan ölçü dışındaki ölçülerden yararlanarak hesaplanan sonsal varyans değeri kullanılmaktadır (Demirel 2005).

Baarda Yöntemi

Baarda yöntemi, Global test ve Data Snooping olmak üzere iki aşamadan oluşur. Global testte önsel varyansla sonsal varyans değeri karşılaştırılarak model test edilir (Kuang 1996). Global test geçilememişse üç şık üzerinde durulur;

- Fonksiyonel model yanlıştır.
- Stokastik model yanlıştır.
- Ölçülerimizde kaba hata yapılmıştır.

EKKY ile parametre kestiriminde stokastik model büyük önem taşır. Stokastik model ise ölçü ağırlıkları, varyansları ve kovaryanslarının belirlenmesi aşamasıdır. Dengelemede hangi ölçünün ne kadar düzeltme alacağı ağırlığı ile ilişkilidir. Bu nedenle ağırlıkların mümkün olduğunca doğru bir şekilde belirlenmesi gerekir. Ölçü ağırlıklarının yanlış belirlenmesi düzeltmelerin yanlış hesaplanmasına, bu da dengeleme ile bulunan sonsal varyansın hatalı olmasına neden olur. Global testin geçilememesinin nedenlerinden birisi bu olabilir. Varyans, kaba hatalardan kestirilen bilinmeyenlere göre daha fazla etkilenir (Wilcox 2001). Ölçülerde kaba hata varsa sonsal varyans hataya bağlı olarak artacaktır. Bu durumda da global test geçilemeyebilir. Global test geçilememiş ise ilk olarak stokastik model kontrol edilmelidir. Eğer stokastik model doğru olduğu halde global test hala geçilemiyorsa ölçülerimizde büyük bir olasılıkla kaba hata yapılmıştır. Sonraki aşama Data Snooping Yöntemi ile uyuşumsuz ölçülerin yerelleştirilmesi yani hangi ölçünün kaba hatalı olduğunun belirlenmesidir (Kuang 1996).

Data Snooping yönteminde sıfır hipotezi ve alternatif hipotez sırasıyla,

$$\begin{aligned} H_0 &: E(\nabla l_i) = 0 \\ H_A &: E(\nabla l_i) \neq 0 \end{aligned} \tag{1}$$

şeklinde kurulur.

Baarda test istatistiği; v_i mutlak değer olarak ölçü düzeltmesi ve σ_{vi} ölçü düzeltmesinin standart sapması olmak üzere,

$$w_i = \frac{|v_i|}{\sigma_{vi}} \tag{2}$$

şeklinde hesaplanır. Düzeltmenin standart sapması, σ_{li} ilgili ölçünün standart sapması ve r_i kısmi redundans sayısı olmak üzere;

$$\sigma_{vi} = \sqrt{r_i}\,\sigma_{li} \tag{3}$$

formülüyle hesaplanır. Buradan Baarda yönteminde ölçü duyarlıklarının iyi bilinmesinin diğer bir deyişle ölçülerin doğru bir şekilde ağırlıklandırılmasının ne kadar önemli olduğu anlaşılmaktadır. Baarda, w_i test istatistiklerinin standart normal dağılımda olduğunu göstermiştir. α yanılma olasılığı 0.001 olarak alınmışsa sınır değer 3.29 olarak bulunur. Bir ölçünün uyuşumsuz olarak değerlendirilebilmesi için,

$$|w_i| > 3.29 \tag{4}$$

veya

$$\left|\hat{v}_i\right| > 3.29\sigma_{vi} \quad (5)$$

olması gerekir (Kuang 1996).

Baarda yöntemiyle belirlenebilir kaba hata sınır değeri;

$$\nabla_0 l_i = \frac{\delta_0 \sigma_{li}}{\sqrt{r_i}} \quad (6)$$

dir. δ_0, hipotez testlerindeki 1. ve 2. tip hata olasılıklarına bağlı olarak değişen dış merkezlik parametresidir. (6) formülünden görüleceği gibi bir ölçünün kısmi redundans sayısı ne kadar büyükse yani o ölçü diğer ölçüler tarafından ne kadar fazla kontrol edilebiliyorsa o ölçüde o kadar küçük kaba hata Baarda testiyle belirlenebilir. Belirlenemeyen kaba hataların bilinmeyen parametrelerin kestirimindeki etkisi ise,

$$\nabla_{0,i} X = \left(A^T PA\right)^{-1} A^T P \nabla_{0,i} l \quad (7)$$

dir. Burada A ; n x u boyutunda katsayılar matrisi, X ; u x 1 boyutunda bilinmeyenler vektörü, P ölçülerin ağırlık matrisi, $\nabla_{0,i}X$ ise etki vektörüdür. Buna göre eğer kısmi redundansı küçük bir ölçüde kaba hata yapılmışsa bu kaba hatanın sonuçlar üzerindeki etkisi, kısmi redundansı büyük olan ölçüdeki kaba hataya göre daha fazla olur (Seemkooei 2001). Bunun dışında, Baarda yönteminin başarılı olması için ölçü duyarlıklarının çok iyi bilinmesi gerekir (Kuang 1996). Baarda yönteminde düzeltmelerin standartlaştırılmasında birim ağırlıklı ölçünün önsel varyans değeri kullanılır. Bu değer ölçülerin kovaryans matrisinin tersi ile çarpılarak ağırlık matrisi elde edilir. Ölçülerden yararlanarak hesaplanan birim ağırlığın sonsal varyans değeridir. Önsel ve sonsal varyansların istatistiksel olarak eşit olması durumunda dengelemenin doğru olduğu söylenir. Genellikle birim ağırlığın önsel varyansı 1 olarak alınır. Bu durumda ağırlık matrisi ölçülerin varyans – kovaryans matrisinin inversine eşittir (Leick 1995).

Pope Testi

Önsel varyansın bilinemediği durumlarda düzeltmelerden hesaplanan varyans kullanılarak uyuşumsuz ölçü araştırması yapılabilir. Ülkemizde yaygın olarak kullanılan Pope yönteminde test istatistiği (8) formülü ile sonsal varyansa göre hesaplanmaktadır. Hesaplanan değer kaba hatalardan etkilenmektedir. Bu nedenle Pope yöntemi iyi bir yöntem değildir. Pope test istatistiği;

$$T_i = \frac{\left|v_i\right|}{\hat{\sigma}_0 \sqrt{\left(Q_{vv}\right)_{ii}}} \quad (8)$$

şeklinde verilir. (8) eşitliğinde varyans paydada yer almaktadır. Ne kadar büyük kaba hata yapılırsa varyans o kadar artar ve bunun sonucu T_i test istatistiği de o oranda azalır. Belirli bir değerden sonra uyuşumsuz bir ölçü bu yöntemle iyi bir ölçüymüş gibi değerlendirilebilir (Kuang 1996).

t Testi

Sonsal varyansın hesabında kullanılan $V^T PV$ ifadesine her bir ölçünün ayrı bir katkısı vardır. Doğal olarak kaba hatalı ölçülerin iyi ölçülere göre daha büyük bir katkısı olacaktır. İşte bu ifadeye en büyük katkı sağlayan ölçü şüpheli ölçüdür. Bu ölçü dışındaki ölçülerden yararlanarak hesaplanan sonsal varyans değeri büyük bir olasılıkla kaba hatalardan daha az etkilenmiş olacaktır. Bu özelliği itibariyle t testinin Pope yönteminden daha iyi bir yöntem olduğu söylenebilir. f serbestlik derecesi, v_i düzeltme değeri, Q_{vivi} düzeltmelerin ağırlık katsayıları matrisi olmak üzere student dağılımlı test istatistiği,

$$T_i = \frac{\left|v_i\right|}{\sigma_{0,i} \sqrt{\left(Q_{vv}\right)_{ii}}} \sim t_{f-1} \quad (9)$$

$$\sigma_{0,i}^2 = \left(v^T Pv\right) - \frac{v_i^2}{Q_{vivi}} / f - 1 \quad (10)$$

(Demirel 2005) olur.

KIRILMA NOKTASI KAVRAMI

Robust istatistiğin en önemli iki kavramından birisi kırılma noktası diğeri ise etki fonksiyonudur. Robust istatistikte amaç kırılma noktası yüksek kestiriciler veya test yöntemleri geliştirmektir. Kırılma noktası bir kestiricinin güvenilirlik ölçütüdür. Farklı kestiricilerden kırılma noktası daha yüksek olan daha güvenilirdir. Kırılma noktası bir kestiricinin baş edebileceği veya tolerans gösterebileceği uyuşumsuz ölçü sayısı ile ilgilidir. Örneğin bir kestirici bir tane bile uyuşumsuz ölçüye duyarlıysa bu kestiricinin kırılma noktası sıfırdır. Olabilecek en kötü kırılma noktası budur. Bu duruma en güzel örnek EKKY'dir. Bilindiği gibi en basit EKKY ortalama almaktır. Örneğin ölçülerimizin, 10,11,12,13 ve 14 olduğunu varsayalım. Bu ölçülerin ortalaması 12 dir ve ölçüler normal dağılımda olduğu için EKKY iyi sonuç vermiştir. Şimdi son ölçüyü 14 yerine 50 yapalım ve buna göre ortalama hesaplayalım. Yeni ortalama 19.2 olacaktır. Yani ortalama değer, diğer bir deyişle EKKY kırılmıştır, yani yöntem yanlış sonuç vermiştir. Ortalama bir tek uyuşumsuz ölçüye bile duyarlık göstermiştir. Öyleyse kırılma noktası sıfırdır yani güvenilir olmayan kestiricidir. Şimdi ise son ölçüyü 14 yerine 100 alarak medyanı hesaplayalım. Bu durumda medyan yine 12'dir yani bir tane uyuşumsuz ölçü medyanı kıramamıştır. Medyanın kırılma noktası 0.50 dir. Olabilecek en yüksek kırılma noktası budur. Buradan medyanın ortalamaya göre daha güvenilir olduğu söylenebilir (Wilcox 2001).

Uyuşumsuz Ölçü Belirleme Testlerinin Kırılma Noktası

Klasik uyuşumsuz ölçü testleriyle güvenilir olarak ancak bir tane uyuşumsuz ölçü belirlenebilir. Bunu bir örnekle açıklayalım. Bilindiği gibi normal dağılımda ortalamaya (μ) ve standart sapmaya (σ) göre tanımlanan aralıkların olasılıkları bulunabilir. Sözgelimi $\mu \pm \sigma$ aralığının olasılığı 0.68 dir. Normal dağılımın bu özelliği uyuşumsuz ölçü belirleme testlerinin temelini oluşturur. Uyuşumsuz ölçü testlerinde düzeltmeler standart sapmaya bölünerek 2, 2.5 veya 3.29 gibi seçilen bir sınır değerle karşılaştırılır, standartlaştırılmış düzeltmesi bu sınır değeri aşan ölçü uyuşumsuz olarak kabul edilir. Ancak gerek EKKY ile kestirilen değerin gerekse de standart sapmanın kırılma noktası sıfır olduğu için test yöntemleriyle ancak bir tane uyuşumsuz ölçü güvenle belirlenebilir. Örneğin 2,3,4,5,6,7,8,9,10,50 gibi değerlerimiz olsun. Uyuşumsuz ölçü belirleme kriterimizi $|X - \mu| > 2\sigma$ olarak kabul edelim. Yani standartlaştırılmış düzeltmesi 2 den büyük olan ölçü uyuşumsuz olsun. Bu değerlerle hesaplanan ortalama 10.4, standart sapma ise 14.15 dir. Buna göre 50 değeri uyuşumsuzdur. Şimdi 10 değerini de 50 yapalım yani oda bir uyuşumsuz olsun. Buna göre hesapladığımız ortalama ve standart sapma sırasıyla 14.4 ve 18.89 olacaktır. Ancak bu durumda 50 değerinin standartlaştırılmış düzeltmesi 2'den değil de 1.88'den büyük olacaktır; $|50 - 14.4| > 1.88\sigma$. Ölçü uyuşumsuz olduğu halde düzeltmesi standart sapmanın 2 katından küçük olacağı için uyuşumsuz ölçü olarak belirlenemeyecektir. Yani uyuşumsuz ölçü testi 2 tane uyuşumsuz ölçü olduğu zaman kırılmış olacaktır (Wilcox 2001).

KALDIRAÇ NOKTASI KAVRAMI

Robust istatistikte tasarım uzayındaki uyuşumsuzlar kaldıraç noktası olarak adlandırılır (Rousseeuw ve Leroy 1987). Kaldıraç noktalarıyla ilgili olarak iki tür problemle karşılaşılabilir. İlkinde kaba hata yapılmasa bile kaldıraç noktası yüzünden En Küçük Kareler Yöntemi kırılır. İkinci problem ise kaldıraç noktalarında kaba hata yapılmasıdır. Kaldıraç noktalarının en karakteristik özelliği kısmi redundanslarının küçük olmasıdır (Hekimoğlu 2005). Kısmi redundansı küçük olan ölçülerde kaba hata yapılmışsa bu tür uyuşumsuz ölçüleri herhangi bir yöntemle belirlemek kısmi redundansı büyük olan ölçülere göre daha zordur. Bir robust kestiricisi olan M kestiriciler de bir kaldıraç noktası olması durumunda bile kırılmaktadır. Kaldıraç noktası olması durumunda Eş redundanslı tasarım (genelleştirilmiş M kestirimi) uygulanır. Kaldıraç noktalarına karşı LMS gibi yüksek kırılma noktalı robust yöntemlerde işe yaramaktadır.

Bunların yanında Wicki (2001) tarafından sunulan BIBER kestiricisi de kaldıraç noktalarındaki uyuşumsuzları belirlemede kullanılabilir.

Genelleştirilmiş M kestiriminde normal denklemler,

$$\frac{1}{\sigma}\sum_{i=1}^{n}\eta(x_i)w(v_i/\sigma)a_{ij}=0,\quad \mathrm{j}=1,2,...,\mathrm{n} \qquad (11)$$

şeklinde verilir (Hampel vd 1986). Genelleştirilmiş M Kestirimindeki amaç, $\eta(x_i)$ ağırlık fonksiyonu yardımıyla kaldıraç noktasındaki uyuşumsuzların etkisini sınırlandırmaktır. $\eta(x_i)$ için Huber $\sqrt{r_i}$, Koch $t/2=8$ olmak üzere $r_d=(1/n)\sum_{i=1}^{n}r_i^{t/2}$, Hekimoğlu (1998) de ise p_i^* denkleşmiş ağırlıkların alınmasını önerilmiştir. r_i ölçülerin kısmi redundanslarıdır (Hekimoğlu 2005).

Eş redundanslı tasarım iteratif olarak şu şekilde gerçekleştirilebilir. İlk olarak izdüşüm matrisi,

$$H=A\left(A^TPA\right)^{-1}A^TP \qquad (12)$$

eşitliği ile hesaplanır. Bütün ölçülerin geometrik ve stokastik etkisi farklı olacağı için izdüşüm matrisinin diagonal elemanları farklı olacaktır. Bu matrisin birim matristen farkı redundans matrisidir. Redundans matrisinin diagonal elemanları ise ölçülerin kısmi redundanslarıdır. Eğer kaldıraç noktası varsa onun kısmi redundansı diğerlerine göre oldukça küçüktür. Yani kaldıraç noktası olup olmadığını belirlemede en temel etken izdüşüm matrisidir (Gross 2003). Eşredundanslı tasarımın birinci aşamasında yeni ağırlıklar,

$$p_i^*=p_i\frac{r_i}{1-r_i} \qquad (13)$$

ile belirlenir (Hekimoğlu 1998). Yeni ağırlıklara göre izdüşüm matrisi yeniden hesaplanır. Bu işleme yakınsama sağlayıncaya kadar devam edilir. Eğer eş redundanslı tasarım, robust bir yöntem olan M kestiriminin her aşamasında kullanılırsa redundansları denkleştirilmiş genelleştirilmiş M kestirimi elde edilir (Ata 1999).

ROBUST M KESTİRİM YÖNTEMLERİ

Robust kestirim yöntemleri içinde en yaygın kullanılanlardan biri M kestiricilerdir. M kestiriciler maksimum olasılık kestiricisinin genelleştirilmiş biçimidir (Hampel v.d. 1986, Huber 1981). M kestiricisi olarak çok sayıda yöntem sunulmuştur. Bu yöntemlerin her biri farklı amaç (kayıp), etki ve ağırlık fonksiyonu ile tanımlanır. Parametre kestiriminde bilinmeyen parametrelerin gerçek değeri ile kestirilen değerleri arasındaki fark kayıp fonksiyonu ile ifade edilir. Amaç bu fonksiyonu minimum yapmaktır. Kayıp fonksiyonu $\rho(v)$, düzeltmelere göre tanımlanan bir fonksiyondur. Bu fonksiyonunun düzeltmelere göre türevinin alınmasıyla da etki fonksiyonu elde edilir. Etki fonksiyonu kırılma noktasından sonra Robust istatistiğin bir diğer önemli kavramıdır. Düzeltmeleri büyük olan ölçülerin (bir anlamda kaba hatalı ölçülerin) parametre kestirimindeki etkilerinin az olması istenir. Etki fonksiyonu,

$$\frac{\partial\rho(v_i)}{\partial v_i}=\psi(v_i) \qquad (14)$$

ile belirlenir. Etki fonksiyonunun düzeltmelere bölünmesiyle de ağırlık fonksiyonu elde edilir. Ağırlık fonksiyonu ile düzeltmelere göre ölçülerimizin ağırlıkları belirlenir. Ağırlık fonksiyonu,

$$P(v)=\frac{\psi(v)}{v} \qquad (15)$$

şeklinde verilir. M kestirimi, iteratif olarak yeniden ağırlıklandırmalı en küçük kareler algoritmasıyla gerçekleştirilmektedir. Her iterasyonda standartlaştırılmış düzeltmeler bir sınır değer ile karşılaştırılır, seçilen ağırlık fonksiyonuna göre ölçülerin yeni ağırlıkları belirlenir. Bu iteratif çözümde kaba hatalı ölçülerin yeni ağırlıkları gittikçe küçülmektedir. İstenen yakınsama sağlayıncaya kadar iterasyona devam edilir (Berber 1997). Normal denklemler ve bilinmeyenler;

$$A^T P(v) = A^T P(AX - L) = 0 \qquad (16)$$

$$X = \left(A^T PA\right)^{-1} A^T PL \qquad (17)$$

şeklinde yazılabilir.

En yaygın olarak kullanılan ağırlık fonksiyonları aşağıda sıralanmıştır,

Huber

$$\overline{p}_i = \begin{cases} 1 & |v_i| \le c \\ \dfrac{c}{|v_i|} & |v_i| > c \end{cases} \quad \text{c=1.5 veya 2.0} \qquad (18)$$

Hampel

$$\overline{p}_i = \begin{cases} p_i & |v_i| \le a \\ p_i a / |v_i| & a < |v_i| \le b \\ p_i a \dfrac{c - |v_i|}{c - b} \Big/ |v_i| & b < |v_i| \le c \\ 0 & |v_i| > c \end{cases}$$

$$a = 1.7, b = 3.4, \mathrm{c} = 8.5 \qquad (19)$$

Beaton – Tukey

$$\overline{p}_i = \begin{cases} p_i \left(1 - \left(\dfrac{v_i}{c}\right)^2\right) & |v_i| \le c \\ 0 & |v_i| > c \end{cases}$$

$$c = 1.5 \text{ veya } 2.0 \qquad (20)$$

Andrews

$$\overline{p}_i = \begin{cases} p_i \sin(v_i / c) / (v_i / c) & |v_i| \le c\pi \\ 0 & |v_i| > c\pi \end{cases}$$

$$c = 1.5 \text{ veya } 1.339 \qquad (21)$$

Danimarka (Csapo, Kis, Völgyesi 2003),

$$\overline{p}_i = \frac{1}{1 + a_k v_{j-1}^2}, \qquad a_k = 3 / v_k^2$$

$$\begin{array}{lll} v_k = 3\sigma_0 & v_{\max} > 3\sigma_0 & \\ v_k = 2\sigma_0 & 2\mu_0 < v_{\max} < 3\sigma_0 & (22) \\ v_k = \sigma_0 & \mu_0 < v_{\max} < 2\sigma_0 & \end{array}$$

Danimarka yönteminde σ_0 değeri olarak, biliniyorsa önsel varyans bilinmiyorsa robust standart sapma kullanılmalıdır. Düzeltmelerin standartlaştırılmasında mümkün olduğunca En Küçük Kareler Yöntemindeki sonsal varyans değeri kullanılmamalıdır. Çünkü bu değer kaba hatalardan etkilenmiş bir değerdir. Robust standart sapma V ölçü düzeltmesi olmak üzere,

$$\hat{\sigma}_{rob} = \frac{1}{0.6745} \mathrm{medyan}\left(\left| \mathrm{V_i} - \mathrm{medyan}(\mathrm{V_i}) \right|\right) \qquad (23)$$

dir (Gross 2003).

DİĞER ROBUST YÖNTEMLER

Aşağıda iteratif ağırlıklandırma yöntemi, Berberan yöntemi, BIBER kestiricisi, Robust – Ridge kestiricisi ve korelasyonlu ölçüler için robust yöntemler olmak üzere farklı robust yöntemler verilmiştir.

İteratif Ağırlıklandırma Yöntemi

EKKY'nin iyi yanlarından birisi ölçülerin dağılımlarının bilinmesine gerek duyulmamasıdır. Ağırlık matrisinin ölçülere ilişkin varyans – kovaryans matrisinin tersi alınarak oluşturulması minimum varyans çözümünü sağlar. EKKY'deki ağırlık matrisi,

$$P = \sigma_0^2 \Sigma_{ll}^2 \qquad (24)$$

eşitliği ile hesaplanır (Demirel 2005). Burada P ağırlık matrisi, Σ_{ll} ölçülerin varyans – kovaryans matrisidir. EKKY ile dengelemede yapılan budur. Bir ölçüde kaba hata yapılması o ölçünün düzeltmesinin artmasına yol açar. Yayma etkisinin fazla olmadığı bir sistemde kaba hatalı ölçünün düzeltmesi diğerlerinkinden belirgin bir şekilde büyüktür. Bu özellikten yararlanarak, ölçü ağırlıkları varyanslarından başka düzeltmelerine de bağlı olarak belirlenerek uyuşumsuz ölçülerin parametre kestirimindeki etkisi azaltılmaya çalışılır. Bu yöntem iteratif ağırlıklandırma yöntemi olarak adlandırılır.

İteratif ağırlıklandırma yönteminde, ağırlık matrisi aşağıdaki gibi

$$P = \left(E^2 + \Sigma_{ll}\right)^{-1} \quad (25)$$

belirlenir. E matrisi, v_i'ler ölçü düzeltmesi olmak üzere,

$$E = \begin{bmatrix} v_1 & 0 & . & 0 \\ 0 & v_2 & . & 0 \\ . & . & . & . \\ 0 & 0 & . & v_n \end{bmatrix} \quad (26)$$

şeklinde oluşturulur (Aduol 1994).

Berberan Yöntemi

Hem Robust yöntemlerin hem de uyuşumsuz ölçü testlerinin birbirlerine göre avantajları vardır. Berberan yöntemi her iki yöntemin avantajlarından birlikte istifade edebilmek için sunulmuştur. Berberan yönteminde Baarda test istatistikleri belirli bir yanılma olasılığına (α) göre normal dağılım tablosundan alınan kritik değerle karşılaştırılır. Ölçülerin ağırlıkları ise Danimarka yöntemi ile belirlenir. σ_{li} ölçülerimizin standart sapmasıdır. Demek ki bu yöntemde ölçü duyarlıkları çok iyi bilinmelidir. w_i ise Baarda yönteminde hesaplanan test istatistikleridir. k sınır değeri olarak sırasıyla (her bir iterasyonda) $\alpha = 0.1$, $\alpha = 0.01$ ve $\alpha = 0.001$ yanılma olasılıklarına karşılık gelen ve normal dağılım tablosundan alınan 1.645, 2.58 ve 3.29 değerleri kullanılır. Berberan yöntemi 3 iterasyon adımından oluşmaktadır. Berberan yönteminde ağırlıklar,

$$(p_l)_{ii} = \begin{cases} \dfrac{1}{\sigma_{li}^2 w_i} & |w_i| > k \textbf{ ise} \\ \dfrac{1}{\sigma_{li}^2} & |w_i| \le k \textbf{ ise} \end{cases} \quad (27)$$

fonksiyonuna göre elde edilir. Berberan yönteminin sağladığı ilk avantaj c sınır değerinin daha güvenli bir şekilde belirlenmesidir. Bilindiği gibi diğer Robust yöntemlerde c = (1.5 – 2)σ gibi değerler kullanılmaktadır. Bunlar istatistikçiler tarafından deneyimlere bağlı olarak verilmiş değerlerdir. Bu sınır değerler, jeodezik problemlerde her zaman ve her koşulda geçerli olacak diye bir kural yoktur. Berberan yönteminin diğer bir avantajı ise iterasyon sayısının 3 olmasıdır. Bazı Robust yöntemler daha büyük iterasyon sayısı gerektirebilmektedir (Berberan 1995).

BIBER (Bounded Influence By standardizEd Residuals) Kestiricisi

BIBER (Bounded Influence By standardizEd Residuals) kestiricisi jeodezik ağların dengelenmesi için geliştirilmiş bir kestiricidir, Schweppe tipli maksimum olasılık kestiricisi olan BIBER kestiricisi EKKY'ne benzer bir kayıp fonksiyonuna sahiptir. Ölçüler normal dağılımda olduğu zaman EKKY ile özdeş sonuç vermektedir, aksi takdirde algoritması standartlaştırılmış düzeltmeler ile uyuşumsuz ölçülerin etkisini sınırlandırmaktadır. Bu sayede kaldıraç noktalarının etkisi azaltılabilmektedir. Sınır değer, düzeltmenin standart sapması bir c sabiti ile çarpılarak elde edilmektedir. BIBER kestiricisinin etki fonksiyonu aşağıdaki gibi verilmiştir.

$$\psi_{ki}(v_i) = \begin{cases} v_i & |v_i| < k_i \text{ ise} \\ sign(v_i)k_i & |v_i| \ge k_i \text{ ise} \end{cases} \quad (28)$$

Sınır değer,

$$k_i = c\sigma_{vi} = c\sigma_0\sqrt{r_i} \quad (29)$$

şeklinde hesaplanabilir. c değerinin (2.5 – 4) arasında alınması uygun görülmektedir. BIBER Kestiricisi için normal denklem aşağıdaki gibidir.

$$\sum_{i=1}^{n} \psi_{ki}(v_i)a_{ij} = 0 \quad (30)$$

BIBER kestiricisinde birinci aşamada EKKY ile çözüm daha sonra her bir ölçü için aşağıdaki formüle göre standartlaştırılmış düzeltmeler hesaplanır;

$$w_i = \frac{v_i}{\sigma_{vi}} = \frac{v_i}{\sigma_0\sqrt{(q_{vv})_{ii}}} \quad (31)$$

σ_{vi} ilgili ölçünün standart sapmasıdır. Standartlaştırılmış düzeltmesi en büyük olan ölçü şüpheli ölçüdür. Bundan sonraki işlem adımı bir önceki aşamada seçilen ölçünün etkisini azaltacak şekilde bilinmeyen parametrelerin ve düzeltmelerin kestirimi aşamasıdır. Maksimum standart sapmalı ölçünün kurgusal ağırlığı p_i^*,

$$p_i^* = \frac{k_i}{|v_i|} \tag{32}$$

ile bulunur. Buna göre bilinmeyen parametrelerin robust kestirimi ve düzeltmeler sırasıyla aşağıdaki gibi bulunur;

$$\begin{aligned} \overline{X}_{ROB} &= \left(A^T P^* A\right)^{-1} A^T P^* L \\ V_{TOT} &= A\overline{X}_{ROB} - L \\ V_{ROB} &= P^* V_{TOT} \end{aligned} \tag{33}$$

Yukarıdaki gibi bulunan V_{ROB} vektörü; $|v_i| < k_i$ olan ölçüler için ölçülen değer ile dengelenmiş değer arasındaki farkı, $|v_i| \geq k_i$ olan ölçüler ise k_i sınır değeri ifade etmektedir. Bütün işlem tüm düzeltmeler $\left[-k_i \leq v_{ROB_i} \leq k_i\right]$ aralığında oluncaya kadar devam eder (Wicki 2001). BIBER kestiricisi ölçülerin iteratif olarak ağırlıklandırılması ilkesine dayanır.

Ridge Kestiriciler

Ridge kestiriciler aslında Robust kestiriciler değildir. Sözgelimi Danimarka yöntemi gibi kaba hatalı ölçülere karşı kullanılmazlar. A Katsayılar matrisinin sütunları arasında yakın lineer bir bağımlılık olması durumunda Ridge kestiricisi iyi sonuç vermektedir. Ridge kestiriminde amaç normal denklem matrisinin inversinin kondisyonunu arttırmaktır. Lineer bağımlılık problemiyle kaba hata probleminin birlikte olması durumunda bilinmeyen parametreler;

$$X_k^{rob} = \left(A'A + kI_u\right)^{-1} A'AX_{ROB} \tag{34}$$

şeklinde kestirilir. k değeri için Hoerl, Kennard ve Baldwin,

$$k = \frac{u\sigma^2}{X'X} \tag{35}$$

alınmasını önermiştir. Burada u bilinmeyen sayısıdır. σ^2 varyans değeri, X ise bilinmeyen parametrelerin kestirim değeridir. Robust – Ridge kestiriminde önce seçilen bir Robust yönteme göre parametre kestirimi yapılarak X_{rob} değeri elde edilir. Daha sonra yukarıdaki formül iteratif olarak çözülerek hem kaba hatalardan hem de lineer bağımlılıktan etkilenmeden parametre kestirimi yapılabilir (Gross 2003).

Korelasyonlu Ölçüler için Robust Yöntemler

Korelasyonlu ölçülerle jeodezide sıkça karşılaşılabilir. Örneğin doğrultu ölçüleri korelasyonlu olmamasına rağmen doğrultu farklarından elde edilen açılar korelasyonludur. Bir diğer örnek GPS ölçmelerinden verilebilir. Bilindiği gibi GPS gözlemlerinde, ikili farklar oluşturma bir nevi korelasyonlu ölçüler demektir. Şimdiye kadar verilen robust yöntemler bağımsız ölçüler için uygundur. Gözlemlerde kaba hata yapılmışsa ve korelasyonlu ölçüler söz konusu ise indirgeme faktörlerinin kullanılması gerekir. Huber′ in etki fonksiyonu seçilmişse indirgeme faktörleri,

$$\gamma_{ii} = \begin{cases} 1 & |\tilde{v}_i| \leq c \\ \dfrac{c}{|\tilde{v}_i|} & |\tilde{v}_i| > c \end{cases} \tag{36}$$

dir. c, 1.0 veya 1.5 gibi bir sabittir. Ağırlık elemanları için bir diğer indirgeme faktörü IGG IIIM,

$$\gamma_{ii} = \begin{cases} 1 & |\tilde{v}_i| \leq k_0 \\ \dfrac{k_0}{|\tilde{v}_i|}\left(\dfrac{k_1 - |\tilde{v}_i|}{k_1 - k_0}\right) & k_0 < |\tilde{v}_i| \leq k_1 \\ 0 & |\tilde{v}_i| > k_1 \end{cases} \tag{37}$$

ile de belirlenebilir (Yang v.d. 2002). Bu ifadelerdeki $\tilde{v}_i$ ler standartlaştırılmış düzeltmelerdir. k_0 ve k_1 sırasıyla 2.0 – 3.0 ve 4.5 – 8.5 olarak seçilen sabitlerdir. Bifaktör indirgeme modelinde i. ve j. ölçülere ait indirgeme faktörleri sırasıyla γ_{ii} ve γ_{jj} olmak üzere,

$$\gamma_{ij} = \sqrt{\gamma_{ii}\gamma_{jj}} \tag{38}$$

dir. Ölçülerin ağırlık matrisinin tüm elemanları için hesaplanan indirgeme faktörleri kullanılarak eşdeğer ağırlık matrisi aşağıdaki gibi verilir.

$$\overline{P} = \begin{bmatrix} \gamma_{11}p_{11} & \gamma_{12}p_{12} & \cdots & \gamma_{1n}p_{1n} \\ \gamma_{21}p_{21} & \gamma_{22}p_{22} & \cdots & \gamma_{2n}p_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{n1}p_{n1} & \gamma_{n2}p_{n2} & \cdots & \gamma_{nn}p_{nn} \end{bmatrix} \tag{39}$$

Bilinmeyenlerin yeni robust kestirimi,

$$\hat{X} = \left(A^T \overline{P} A\right)^{-1} A^T \overline{P} L \tag{40}$$

Bilinmeyenlerin varyans – kovaryans matrisi,

$$\Sigma_{\hat{X}} = \hat{\sigma}_0^2 \left(A^T \overline{P} A\right)^{-1} \tag{41}$$

Birim ağırlığın sonsal değeri ise

$$\hat{\sigma}_0^2 = \frac{V^T \overline{P} V}{n-u} \tag{42}$$

eşitliği ile hesaplanır (Yang v.d. 2002).

SAYISAL UYGULAMALAR

Yukarıda izah edilen robust yöntemlere ait çeşitli uygulamalar aşağıda verilmiştir. Uygulamalarda $Y = AX + \varepsilon$ şeklinde verilen lineer bir regresyon modeli kullanılmıştır. Y ölçü uzayı, A tasarım uzayı (katsayılar matrisi), X bilinmeyenler ve ε ise rasgele hatalar vektörüdür. A ve X değerlerinin bilindiği varsayılarak ilk önce hatasız Y değerleri hesaplanmış daha sonra hatasız Y değerlerine rasgele hatalar eklenerek iyi ölçüler elde edilmiştir. Robust yöntemlerin kaba hatalardaki başarısını görmek içinde rasgele bir ölçünün kaba hatalı olması sağlanmıştır. Yapılan uygulamalarda temel amaç parametrelerin gerçek değerleri olan $X_{gerçek}$ değerlerine EKKY ile robust yöntemlerin hangisinin daha fazla yaklaştığını görmektir. Yapılan bütün uygulamalarda robust yöntemler ile gerçek değerlere daha fazla yaklaşıldığı görülmüştür.

Uygulama 1:

Aşağıda sırasıyla lineer bir regresyon modeline ilişkin katsayılar matrisi, ölçü vektörü ve bilinmeyenlerin gerçek değerleri veriliyor. Verilenleri Danimarka yöntemine göre değerlendirelim. 2.ölçü kaba hatalıdır. Bu ölçünün gerçek değeri (hatasız değeri) parantez içinde verilmiştir. Başlangıçta ağırlık matrisi birim matris alınmıştır $(P = I)$. İlk başta EKKY ile çözüm yapıldığında 2. ölçünün düzeltmesi maksimum çıkmıştır, bu nedenle bu ölçünün ağırlığı Danimarka yöntemiyle iteratif olarak Tablo 1'deki gibi elde edilmiştir. Tablo 2' de ise her bir iterasyonda yeni ağırlıklara göre hesaplanan bilinmeyenler verilmiştir Tablo 2'den görüleceği gibi Danimarka yöntemiyle 5 iterasyon sonucu gerçek değere EKKY'den daha fazla yaklaşılmıştır.

$$A = \begin{bmatrix} 1 & 2.9 \\ 1 & 3.6 \\ 1 & 3.9 \\ 1 & 3.7 \\ 1 & 4.4 \\ 1 & 4.1 \\ 1 & 4.7 \\ 1 & 3.7 \\ 1 & 4.5 \\ 1 & 5.1 \\ 1 & 5.7 \\ 1 & 5.5 \end{bmatrix} \quad ; \quad Y = \begin{bmatrix} 4.1086 \\ 5.4765\,(\mathbf{4.82}) \\ 4.0122 \\ 4.4794 \\ 5.5176 \\ 10.8848 \\ 4.8206 \\ 5.8712 \\ 5.9112 \\ 5.9749 \\ 8.1457 \\ 7.3366 \end{bmatrix} \quad ; \quad X_{gerçek} = \begin{bmatrix} 0.5 \\ 1.2 \end{bmatrix}$$

Tablo 1. Şüpheli Ölçünün Ağırlık Değişimi.
Table 1*. Weight Change of the doubtful measurement.*

İterasyon	1	2	3	4	5
Şüpheli Ölçü	2	2	2	2	2
Ağırlık	0.25	0.0214	0.0236	0.0219	0.0219

Tablo 2. Danimarka Yöntemiyle Kestirilen Bilinmeyen Parametreler.
***Table 2.** Unknown parameters which estimated by Danish Method.*

İterasyon	Bilinmeyen Parametreler
1.İterasyon	$X = [0.8254 \quad 1.1569]$
2.İterasyon	$X = [0.3870 \quad 1.2411]$
3.İterasyon	$X = [0.3914 \quad 1.2403]$
4.İterasyon	$X = [0.3879 \quad 1.2410]$
5.İterasyon	$X = [0.3879 \quad 1.2410]$
EKKY	$X = [2.0516 \quad 0.9212]$

Uygulama 2: Aşağıda verilen değerlerin Berberan Yöntemiyle değerlendirilmesiyle aşağıdaki sonuçlar elde edilmiştir. Sınır değerler olarak sırasıyla $\alpha = 0.1, 0.01, 0.001$ anlamlılık düzeylerine karşılık gelen 1.645, 2.58 ve 3.29 değerleri kullanılmıştır. Başlangıçta ağırlık matrisi birim matris olarak alınmıştır. Kaba hatalı olan 2. ölçü ve diğer ölçülerin ağırlıkları Tablo 3'de, Tablo 4'de ise yeni ağırlıklara göre hesaplanan bilinmeyenler verilmiştir. 3.iterasyon sonucu hesaplanan bilinmeyen değerler gerçek değerlere En Küçük Kareler Yöntemine göre daha fazla yaklaşmıştır. Kaba hatalı olan 2. ölçünün gerçek değeri Y vektöründe parantez içinde verilmiştir.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} \; ; \; Y = \begin{bmatrix} 1,4674 \\ 9.1503(\mathbf{3.30}) \\ 4.8253 \\ 6.3877 \\ 6.3535 \\ 10.091 \\ 11.489 \\ 11.662 \end{bmatrix} \; ; \; X_{gerçek} = \begin{bmatrix} 0.5 \\ 1.4 \end{bmatrix}$$

Uygulama 3: Aşağıda verilenleri Huber' in M kestirimine göre değerlendirelim. Bu örnekte önsel standart sapma değeri $\sigma_0 = 0.01$ metredir. EKKY ile bulunan düzeltmelerden hesaplanan standart sapma ise 0.03 metredir. Görüldüğü gibi uyuşumsuz ölçüler varyansın (dolayısıyla standart sapmanın) artmasına neden olmuştur. (23) formülüne göre hesaplanan robust standart sapma ise $\sigma_{rob} = 0.01$ metredir. Yani bu şekilde hesaplanan standart sapma uyuşumsuz ölçülerden daha az etkilenir. 2. ölçünün gerçek değeri 4.82 dir. Bu değere standart sapmanın 11 katı büyüklükte bir kaba hata eklenmiştir. Bu örnekte düzeltmelerin standartlaştırılmasında robust standart sapmadan yararlanılmıştır. c sınır değeri ise 1.5 olarak seçilmiştir. Başlangıçta ağırlık matrisi birim matris olarak alınmıştır. Huber'in ağırlık fonksiyonuna göre elde edilen ağırlıklarla EKKY iteratif olarak uygulandığında Tablo 6'dan de görüleceği gibi 4. iterasyon sonucunda gerçek değerlere EKKY'den daha fazla yaklaşılmıştır.

Tablo 3.Ölçü ağırlıklarının değişimi.
***Table 3.** Change of measurement weights.*

Ölçü	1	2	3	4	5	6	7	8
1.İterasyon	0.8756	0.7053	4.6628	10.131	3.7198	7.8322	3.5787	13.728
2.İterasyon	5.7591	0.8103	31.35	58.272	5.3165	7.2363	4.7219	5.7591
3.İterasyon	5.7591	0.7456	31.35	58.272	5.541	7.8796	4.7219	5.7591

Tablo 4. Berberan yöntemi ile kestirilen bilinmeyen parametreler.
***Table 4.**Unknown parameters estimated by Berberan Method.*

İterasyon Sayısı	Bilinmeyen Parametreler
1.İterasyon	$X = [1.1726 \quad 1.3441]$
2.İterasyon	$X = [0.4496 \quad 1.4762]$
3.İterasyon	$X = [0.4252 \quad 1.4818]$
EKKY	$X = [2.3843 \quad 1.1764]$

Tablo 5. Huber'in ağırlık fonksiyonuna göre iteratif çözümde ölçü ağırlıklarının değişimi.
Table 5. *Change of measurement weights in iterative solution according to Huber's weight function.*

Ölçü	1	2	3	4	5	6	7	8	9	10	11	12
1.İterasyon	0.7168	0.2331	1	1	1	1	1	1	1	1	1	1
2.İterasyon	1	0.2082	1	1	1	1	1	1	1	1	1	1
3.İterasyon	1	0.2051	1	1	1	1	1	1	1	1	1	1
4.İterasyon	1	0.2052	1	1	1	1	1	1	1	1	1	1

$$A=\begin{bmatrix}1 & 2.9\\1 & 3.6\\1 & 3.9\\1 & 3.7\\1 & 4.4\\1 & 4.1\\1 & 4.7\\1 & 3.7\\1 & 4.5\\1 & 5.1\\1 & 5.7\\1 & 5.5\end{bmatrix} \quad ; \quad Y=\begin{bmatrix}3.9701\\4.9300\\5.1829\\4.9548\\5.7914\\5.4132\\6.1271\\4.9393\\5.8967\\6.6116\\7.345\\7.1149\end{bmatrix} \quad ; \quad X_{gerçk}=\begin{bmatrix}0.5\\1.2\end{bmatrix}$$

Tablo 6. Huber'in M kestirimi yöntemiyle kestirilen bilinmeyen parametreler.
Table 6. *Unknown parameters which estimated by Huber's M estimation method.*

İterasyon	Kestirilen Parametreler
1	$X=[0.5076 \quad 1.1990]$
2	$X=[0.5020 \quad 1.2001]$
3	$X=[0.5019 \quad 1.2002]$
4	$X=[0.5019 \quad 1.2002]$
EKKY	$X=[0.5437 \quad 1.1931]$

Uygulama 4: BIBER Kestiricisinin kaldıraç noktalarındaki başarısını görmek için lineer bir regresyon modelinde uygulama yapılmıştır. Bunun için kısmi redundansı en küçük olan ölçüye kaba hata eklenmiştir.

$$A=\begin{bmatrix}1 & 3.575\\1 & 7.153\\1 & 13.58\\1 & 34.18(4.15)\\1 & 3.141\\1 & 6.83\\1 & 5.43\\1 & 13.48\\1 & 10.5\end{bmatrix} \quad Y=\begin{bmatrix}5.2887\\9.6814\\19.575\\9.2324\\4.2342\\10.657\\8.6966\\19.353\\15.364\end{bmatrix} \quad X_{gerçek}=\begin{bmatrix}0.5\\1.4\end{bmatrix}$$

Tablo 7. Kısmi Redundans Sayıları.
Table 7. *Partial Redundancy Numbers.*

Ölçü	1	2	3	4	5	6	7	8	9
r_i	0.82	0.87	0.88	**0.15**	0.81	0.87	0.85	0.88	0.89

Tablo 8. BIBER Kestiricisi ile iteratif çözümde Baarda test istatistiği en büyük olan ölçü (şüpheli ölçü), Ağırlıklar ve Bilinmeyen Parametreler.
Table 8. *Measurement(doubtful) which have maximum Baarda test statistic in iterative solution by BIBER estimator.*

	Şüpheli Ölçü	Ağırlık	Bilinmeyen Parametreler
1. İterasyon	4. Ölçü	0.0096	$X=[1.0603 \quad 1.3231]$
2. İterasyon	4.Ölçü	0.0015	$X=[0.5533 \quad 1.3917]$
EKKY	4.Ölçü	1	$X=[10.1228 \quad 0.0970]$

Simülasyon çalışmasında yukarıda verilen değerler kullanılmıştır. Bunlar sırasıyla tasarım matrisi, ölçü matrisi ve bilinmeyenlerin gerçek değeridir. Tasarım matrisinde 4 numaralı noktanın gerçek değeri 4.15 iken 34.18 yapılmıştır. Yani bu nokta bir kaldıraç noktasıdır. BIBER kestiricisinde verilen k formülünde c değeri 1.5 olarak seçilmiştir. BIBER kestiricisinin kaldıraç noktalarındaki kaba hatalı ölçülere karşı başarısını görmek için 4. ölçüye yani kısmi redundans sayısı en küçük olan ölçüye kaba hata eklenmiştir. Tablo 7'den görüleceği gibi 4 numaralı nokta kaldıraç noktası olduğu için kısmi redundansı diğerlerine göre oldukça küçüktür.

BIBER kestiricisi uygulandığı zaman kaldıraç noktasına karşılık gelen ölçünün Baarda test istatistiği maksimum çıkmıştır. (29) formülüyle hesaplanan k değeri ise 0.057 dir. (32) formülüyle de yeniden ağırlıklandırma yapılmıştır. BIBER Kestiricisi ile daha 2. iterasyonda iyi bir sonuç elde edilmiştir.

Uygulama 5: Verilen bir veri setinde lineer bağımlılığa ilave olarak uyuşumsuz ölçülerinde olması problemi incelenmiştir. Kaba hatalı olan 1. ölçünün gerçek değeri Y vektöründe parantez içinde verilmiştir. Bunun için basit bir lineer regresyon modelinde robust – ridge kestiricisi uygulanmıştır. Normal matrisin kondisyonunu arttırmada kullanılan k değeri (35) numaralı formüle göre elde edilmiştir. $Y = AX + \varepsilon$ şeklinde lineer bir regresyon modelini ele alalım. Bu modele ait katsayılar (tasarım) matrisi, ölçüler matrisi ve bilinmeyenlerin gerçek değeri sırasıyla aşağıda verilmiştir.

$$A = \begin{bmatrix} 1 & 1.9 \\ 1 & 2.1 \\ 1 & 2 \\ 1 & 2 \\ 1 & 1.8 \end{bmatrix} \quad ; \quad Y = \begin{bmatrix} 5.7051(\mathbf{2.78}) \\ 2.1872 \\ 2.9627 \\ 3.0438 \\ 2.0868 \end{bmatrix} \quad ; \quad X_{gerçek} = \begin{bmatrix} 0.5 \\ 1.2 \end{bmatrix}$$

İlk önce Huber' in M Kestirimi uygulanmıştır. Huber' in M Kestirimine göre elde edilen ağırlıklar ve bilinmeyenlerin kestirilen değerleri sırasıyla Tablo 9 ve Tablo 10'da verilmiştir. Düzeltmelerin standartlaştırılmasında (23) numaralı formüle göre hesaplanan robust standart sapma kullanılmıştır.

Table 9. Huber'in ağırlık fonksiyonuna göre ölçü ağırlıkları.

Table 9. *Measurement weights according to Huber's weight function.*

Ölçü	1	2	3	4	5
1.İterasyon	0.496	1	1	1	0.775
2.İterasyon	0.448	1	1	1	0.984
3.İterasyon	0.423	1	1	1	1
4.İterasyon	0.418	1	1	1	1

Table 10. İteratif ağırlıklandırmalı EKKY'ne göre bilinmeyenlerin kestirilen değerleri.

Table 10. Estimated values of unknowns according to iterative rewighted least squares algorithm.

İterasyon	EKKY	1	2	3	4
X vektörü	$[8.0866 \quad -2.4947]$	$[6.5167 \quad -1.8005]$	$[4.4280 \quad -0.7820]$	$[4.1239 \quad -0.6371]$	$[4.0884 \quad -0.6205]$

Tablo 11. Hoerl, Kennard ve Baldwin' e göre hesaplanan k değerleri.

Table 11. *k values which calculated according to Hoerl, Kennard and Baldwin.*

İterasyon	1	2	3	4	5
k	0.017	0.305	0.728	0.760	0.762

Tablo 12. Robust – Ridge Kestiricisine göre bilinmeyen parametrelerin kestirilen değerleri.

Tablo 12. *Estimated values of unknown parameters according to Robust – Ridge estimator.*

İterasyon	Kestirilen Parametreler
1	$X = [1.9412 \quad 0.4715]$
2	$X = [0.7019 \quad 1.0853]$
3	$X = [0.6241 \quad 1.1005]$
4	$X = [0.6213 \quad 1.1001]$
5	$X = [0.6211 \quad 1.1001]$

Verilenlerde A matrisine bakılırsa bu matrisin sütunları arasında yakın bir lineer bağımlılık vardır. Bu nedenle Huber yöntemiyle 1. adımda 1. ölçü uyuşumsuz olarak belirlenmiştir. İteratif ağırlıklandırmalı EKKY çözümünde bu ölçünün ağırlığı düşürülerek etkisi azaltılmıştır. Gerekirse bu ölçü tekrar yapılabilir. Bu problemde uyuşumsuz ölçü probleminden başka lineer bağımlılık problemi de olduğu için salt Huber' in M kestirimi yeterli olmaz. Bunun için 2. adımda (34) ve (35) numaralı formüllere göre robust – ridge kestiricisi uygulanmıştır. İteratif çözümde elde edilen k değerleri ve bilinmeyen parametrelerin kestirilen değerleri sırasıyla Tablo 11 ve Tablo 12'de verilmiştir.

Normal matrisin kondisyonunu arttırmada kullanılan k değeri ilk iterasyonda EKKY ile bulunan çözüme göre hesaplanmıştır. Formüldeki standart sapma (23) numaralı formüle göre elde edilen robust standart sapma olarak ele alınmıştır. 5. iterasyon da elde edilen sonuçlar bilinmeyenlerin gerçek değerleriyle karşılaştırılırsa elde edilen sonucun hem kaba hata hem de lineer bağımlılık problemine rağmen oldukça iyi olduğu söylenebilir.

SONUÇLAR

- Kaba hatalı ölçüler EKKY ile bulunan bütün sonuçları ve aynı zamanda yayma etkisi nedeniylede test büyüklüklerini bozarlar. Böylece EKKY ile bulunan sonuçların doğruluğunun bir garantisi olmadığı gibi bu yönteme dayalı olarak uyuşumsuz ölçü belirlemenin güvenilirliği de azdır.
- Robust yöntemlerin en büyük avantajı uyuşumsuz ölçülerin etkisini yerelleştirmeleri yani yaymamalarıdır. Dolayısıyla kaba hatalı ölçülerin daha doğru bir şekilde belirlenmesi mümkün olmaktadır.
- Uyuşumsuz ölçü belirleme başarısı; kaba hatanın büyüklüğüne, sayısına, cinsine, konumuna, serbestlik derecesine, ağın geometrisine, kısmi redundans sayılarına ve stokastik yapıya bağlı olarak değişmektedir.
- Uyuşumsuz ölçü belirlemenin başarısı önemli ölçüde stokastik modelin doğru kurulmasına bağlıdır. Bilindiği gibi dengelemede ölçülerin alacağı düzeltmelerin büyüklükleri ağırlıklarıyla ilişkilidir. Uyuşumsuz ölçü araştırmasında temel veri ise ölçülerin düzeltmeleridir.
- Klasik uyuşumsuz ölçü testleri içerisinde en iyi yöntem Baarda yöntemidir. Yani varyansın bilinmesi durumunda uyuşumsuz ölçü belirleme başarısı artmaktadır.
- Robust kestirimde ölçülerin uyuşumsuz olup olmadığına karar vermede önemli bir etken olan c parametresi için her durumda geçerli bir değer verilemez. İstatistikçiler deneysel çalışmalara göre 1.5, 2 gibi değerler önermişlerdir. c parametresinin A katsayılar matrisine ve α yanılma olasılığına bağlı olarak her ölçü için ayrı ayrı hesaplanması da mümkündür. c parametresi olarak Baarda yöntemindeki sınır değerler alınabilir. Ancak bu durumda standartlaştırılmış düzeltmeler olarak Baarda test büyüklükleri alınmalıdır.
- Normal dağılımın iki parametresinden varyans, ortalamaya göre uyuşumsuz ölçülerden daha fazla etkilenmektedir. Bu nedenle uyuşumsuz ölçü araştırmasında POPE yönteminde olduğu gibi EKKY ile dengelemeyle bulunan varyansı kullanmak veya Robust yöntemlerde düzeltmelerin standartlaştırılmasında bu varyansı kullanmak sakıncalıdır. Bunun için varyans biliniyorsa varyansı bilinmiyorsa robust standart sapmayı kullanmak gerekir.
- BIBER kestiricisi özellikle kaldıraç noktalarındaki kaba hatalı ölçüleri belirlemede de başarılıdır. Kaldıraç noktalarına karşı başarılı bir diğer yöntem genelleştirilmiş M kestirimidir. Lineer bağımlılık ve uyuşumsuz ölçü problemlerinin bir arada olması durumunda robust – ridge kestiricisi kullanılabilir.

KAYNAKLAR

Aduol F.W.O., 1994, Robust Geodetic Parameter Estimation Through Iterative Weighting, Survey Review, 32: 359 – 367

Ata M., 1999, Statik Deformasyon Analizinde Robust Kestirim Yöntemlerinin Kullanılması Üzerine Bir Araştırma, Doktora Tezi, YTÜ Fen Bilimleri Enstitüsü, İstanbul

Berber M., 1997, Kenar Ağlarında Uyuşumsuz Ölçülerin Klasik Uyuşumsuz Ölçü Testleri ve M – Kestirimi ile Belirlenmesi ve Karşılaştırılması, Yüksek Lisans Tezi, YTÜ Fen Bilimleri Enstitüsü, İstanbul

Berberan A., 1995, Multiple Outlier Detection, Survey Review, 33:255

Csapo G., Kis M. and Völgyesi L., 2003, Different Adjustment Methods for the Hungarian Part of the Unified European Gravity Network, XXIII General Assembly of the International Union of Geodesy and Geophysics, June 30 – July 11, 2003, Sapporo, Japan

Demirel H., 2005, Dengeleme Hesabı, YTÜ, İnşaat Fakültesi, Sayı:İN.JFM-2005.001

Dilaver A., Konak H., Çepni, M.S., 1998, Jeodezik ağlarda uyuşumsuz ölçülerin yerelleştirilmesinde kullanılan yöntemlerin davranışları, H. K. M.O. dergisi, sayı:84, Ankara

Erenoğlu R.C., 2003, Jeodezik Ağlarda Uyuşumsuz Ölçülerin Robust Yöntemlerle ve Uyuşumsuz Ölçü Testleriyle Belirlenmesi ve Birbirleriyle Karşılaştırılması, Yüksek Lisans Tezi, YTÜ Fen Bilimleri Enstitüsü, İstanbul

Gross J., 2003, Linear Regression, Springer Verlag

Hampel, F., Ronchetti, E., Rousseeuw, P., Stahel, W. 1986, Robust Statistics The Approach based on influence functions. John Wiley and Sons, New York

Hekimoğlu Ş., 1998, Application of Equiredundancy Desing to M – Estimation, Journal of Surveying Engineering , ASCE, 124(3):103 – 124

Hekimoğlu Ş. and Berber M., 2003, Effectiveness of Robust Methods in Heterogeneous Linear Models, Journal of Geodesy, 76:706 – 713

Hekimoğlu Ş., 2005, Do Robust Methods Identify Outliers More Reliably Than Conventionel Test for Outliers, Zeitschrift für Vermessungwesen, 2005/3, 174 – 180

Hekimoğlu Ş., 2006, Kaba Hataların Belirlenmesindeki Sorunlar, Harita Dergisi:S.135:80 – 93:01/2006

Huber P., 1981, Robust Statistics, John Wiley & Sons Inc. New York

Kuang S.H., 1996, Geodetic Network Analysis and Optimal Design, Ann Arbor Pres, INC Chelsea Michigan

Leick A., 1995, GPS Satellite Surveying, John Wiley & Sons, New York

Rousseeuw P. and Leroy A.M., 1987 Robust Regression and Outlier Detection. John Wiley & Sons Inc. New York.

Seemkooei, A.A., 2001, Strategy for Designing Geodetic Network with High Reliability and Geometrical Strength, Journal of Surveying Engineering, Vol. 127, No.3

Yang Y., Song L. and Xu T., 2002, Robust Estimator for Correlated Observations Based on Bifactor Equivalent Weights, Journal of Geodesy, 76:353 – 358

Wetherill, G.B., 1986, Regression Analysis with Applications, Champan & Hall, London UK.

Wicki F., 2001, Robust Estimator for the Adjustment of Geodetic Networks, First International Symposium on Robust Statistics and Fuzzy Techniques in Geodesy and GIS March 12 – 16 Zurich, Switzerland

Wilcox R.R., 2001, Fundementals of Modern Statistical Methods, Springer Verlag, New York