DÜMF Mühendislik Dergisi

web: http://dergipark.gov.tr/dumf

Değişim noktası kestirimi için CUSUM algoritmasının başarım değerlendirmesi

Memduh KÖSE
Ahi Evran Üniversitesi Bilgisayar Bilimleri Araştırma ve Uygulama Merkezi, Kırşehir
ORCID: 0000-0002-4935-4542

Selçuk TAŞÇIOĞLU*
Ankara Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, Ankara
stasci@eng.ankara.edu.tr ORCID:0000-0001-9064-2960, Tel: (312) 203 33 00 (1808)

Ziya TELATAR
Ankara Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, Ankara
Ziya.Telatar@ankara.edu.tr, ORCID: 0000-0002-1146-2337

Geliş: 12.12.2017, Kabul Tarihi: 03.01.2018

Öz

Bu çalışmada, parametrik CUSUM (cumulative sum) algoritmasının değişim noktası kestirim başarımı analiz edilmiştir. Veriye ait dağılımın bilindiği durumda kullanılabilen bu parametrik yöntem, çalışma kapsamında Gauss rasgele sürecinin ortalamasındaki ani bir değişimin yerinin kestirilmesi amacıyla kullanılmıştır. Yöntemde kullanılan log-olabilirlik oranına dayalı eşitlikler Gauss gürültü modeli için türetilmiş ve yöntemin temel çalışma ilkesi örneklerle anlatılmıştır. Sinyal parametrelerinin bilinmediği durumlarda kullanılan kısmi en iyi (suboptimal) çözümler ele alınarak, pratikte yöntemin başarımının en iyi (optimal) çözüme kıyasla ne ölçüde azalacağı incelenmiştir. Çalışma kapsamında ele alınan CUSUM algoritmasının türevleri için değişim noktası kestirim başarımları benzetimler yoluyla karşılaştırmalı olarak sunulmuştur. Kısmi en iyi yaklaşımlarda sinyal parametreleri için yapılan kestirimlerin ve bu parametreler hakkındaki önsel bilginin (a priori information) doğruluğunun, değişim noktası algılama problemi zorlaştıkça daha da önem kazandığı gösterilmiştir. Ayrıca, CUSUM algoritmasının özyinelemeli yapısı sayesinde hızlı hesaplama süresine sahip olduğu, farklı uzunluktaki sinyaller için gösterilmiştir.

Anahtar Kelimeler: *Değişim noktası kestirimi, basamak değişimi, CUSUM algoritması, parametrik yöntemler, önsel bilgi, log-olabilirlik oranı.*

* Yazışmaların yapılacağı yazar

DOI:

Giriş

Değişim noktası kestirimi; haberleşme, biyomedikal ve konuşma sinyallerinin işlenmesinden sismik veri analizine, istatistiksel süreç kontrolünden finansal veri analizine kadar pek çok farklı alanda yaygın olarak kullanılan bir yöntemdir. Bu kestirim probleminin çözümü için birçok istatistiksel sinyal işleme tekniği geliştirilmiştir. Veriye ait dağılımın bilindiği ve bilinmediği durumlarda kullanılan yöntemler, parametrik ve parametrik olmayan yöntemler olarak sınıflandırılmıştır. Ölçümlere ait dağılım fonksiyonun bilinmesi oldukça güçlü bir bilgi sunduğundan, genel olarak parametrik değişim noktası kestirim yöntemleriyle, zor problemlerde bile oldukça başarılı sonuçlar elde edilebilmektedir. Ancak bu bilgiye sahip olmak her zaman mümkün olmaz ve dolayısıyla parametrik olmayan yöntemlere başvurulur. Her iki yaklaşım hakkında detaylı bilgi için Brodsky ve Darkhovsky (1993), Basseville ve Nikiforov (1993), Ruanaidh ve Fitzgerald (1996) tarafından ortaya konan eserlere başvurulabilir.

Değişim noktası analizinin yaygın kullanıldığı uygulamalardan biri haberleşme sistemleridir. Örneğin, Taşcıoğlu ve Üreten (2009), Taşcıoğlu vd. (2010) çalışmalarında bilişsel radyo ağlarında çoklu değişim noktası analizine dayalı bir spektrum algılama tekniği önermişlerdir. Başka bir uygulama örneği, kablosuz vericilerin kimlik tespit sistemleridir. Bu sistemlerde kimlik tespiti amacıyla kullanılan vericilere ait geçici rejim işaretlerinin başlangıç noktaları, parametrik bir değişim noktası kestirimi yaklaşımıyla bulunmuştur (Üreten ve Serinken, 1999; Üreten ve Serinken, 2005; Köse vd. 2011; Köse vd., 2015).

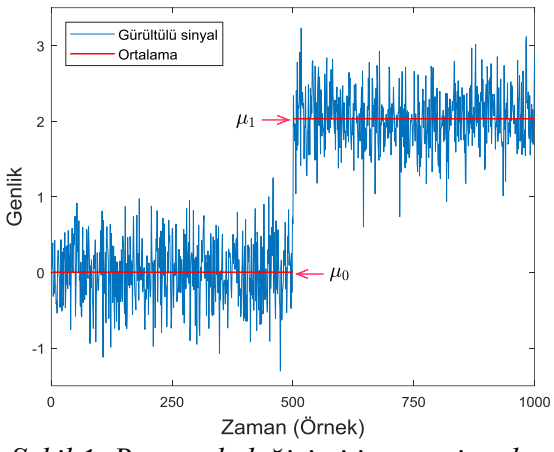
CUSUM algoritması, değişim noktası analizinde yaygın olarak kullanılan yöntemlerden biridir (Basseville ve Nikiforov, 1993; Brodsky ve Darkhovsky, 1993; Montgomery, 2013). Bu çalışmada, parametrik CUSUM algoritmasının değişim noktası kestiriminde kullanımı ele alınmış ve kestirim başarımı çeşitli benzetimlerle analiz edilmiştir.

Problem tanımı

Bağımsız ve aynı dağılıma sahip örneklerden oluşan kesikli rasgele bir $\boldsymbol{x}$ sinyali ele alınsın. Dağılım parametreleri $\boldsymbol{\theta}$ olmak üzere, sinyal örnekleri için olasılık yoğunluk fonksiyonu $f\left(x_i;\boldsymbol{\theta}\right)$ olarak verilsin. Sinyalin m'nci elamanından sonra $\boldsymbol{\theta}$ vektöründeki en az bir parametrede değişim olduğunda sinyal örnekleri için dağılım fonksiyonu

$$x_i \sim \begin{cases} f\left(x_i;\boldsymbol{\theta}_0\right), & i \le m \\ f\left(x_i;\boldsymbol{\theta}_1\right), & i > m \end{cases} \tag{1}$$

biçiminde parçalı yapıda gösterilebilir (Basseville ve Nikiforov, 1993). Bu çalışmada, genellikle değişim noktası analizi problemlerinde ilgilenilen Gauss rasgele süreci için ortalama parametresindeki değişim problemi ele alınacaktır. Bu nedenle $\boldsymbol{\theta}=(\mu,\sigma^2)$ parametre vektöründeki ortalama parametresi μ'nün bir noktada değiştiği, varyans parametresi σ^2'nin ise analiz penceresi boyunca sabit kaldığı kabul edilecektir. Basamak değişimi (step change) olarak da adlandırılan bu yapı için örnek bir sinyal Şekil 1'de verilmiştir. Ortalama parametresinin değişimden önceki ve sonraki değerleri sırasıyla μ_0 ve μ_1 ile gösterilecektir. Bu modelde ilgilenilen problem, değişim noktası m değerinin kestirilmesidir.



Şekil 1. Basamak değişimi içeren sinyal ve değişimden önce ve sonraki ortalama değerler

CUSUM algoritması

İsmini birikimli toplam (cumulative sum) kelimelerinin kısaltmasından alan bu algoritma ilk olarak Page tarafından 1954 yılında önerilmiştir. Algoritmanın, analiz edilecek verinin dağılımı hakkında bilginin bulunduğu ve bulunmadığı durumlar için kullanılabilecek parametrik ve parametrik olmayan türevleri vardır (Brodsky ve Darkhovsky, 1993). Algoritmanın temel çalışma ilkesi ve bazı türevleri hakkındaki bilgiler, Granjon tarafından 2012 yılında yapılan çalışmada anlaşılır bir dille sunulmuştur.

Bu çalışmada, sinyal örneklerinin Gauss dağılımlı olduğu durum ele alınarak, değişim noktası analizinde parametrik CUSUM algoritması uygulanmıştır. Bu algoritma dağılımı bilinen gözlemler için her anda, değişimden önceki ve sonraki parametre değerleriyle elde edilen olabilirlik fonksiyonu (likelihood function) değerlerinin oranı ifadesine dayanır. *i*'nci andaki gözlem değeri x_i için *anlık log-olabilirlik oranı*

$$s_i = \ln\left(\frac{f_{\theta_1}(x_i)}{f_{\theta_0}(x_i)}\right) \tag{2}$$

biçiminde verilir (Basseville ve Nikiforov, 1993, Granjon, 2012). Bu değerlerin birikimli toplamı *k*'ncı anda

$$S_k = \sum_{i=1}^{k} s_i \tag{3}$$

olarak hesaplanır. *N* anına kadarki birikimli toplam değerlerinden en küçük olanın indisi, değişim noktasının bir kestirimidir ve aşağıdaki eşitlikle verilir (Basseville ve Nikiforov, 1993):

$$\hat{m} = \arg\min_{1\le k\le N} S_k \tag{4}$$

CUSUM algoritması için Eşitlik (4)'teki *N* değeri sezim (detection) aşamasında belirlenen eşik değerinin ilk aşıldığı an esas alınarak belirlenir. Bu çalışmanın amacı CUSUM algoritmasının değişim noktası kestirim başarımını analiz etmek olduğu için benzetim yoluyla üretilen sinyaller sezim aşamasından geçip, kestirim aşamasına iletilen sinyaller olarak ele alınmaktadır. Sezim aşamasındaki eşik değerin belirlenmesi konusunda detaylı bilgi için Basseville ve Nikiforov (1993) tarafından yazılan esere başvurulabilir.

Eşitlik (3)'te verilen toplama işlemi, *k* anında hesaplanan log-olabilirlik oranının bir önceki andaki birikimli toplam değerine eklenmesiyle özyinelemeli (recursive) olarak

$$S_k = S_{k-1} + s_k \tag{5}$$

biçiminde ifade edilebilir. Böylece *k* anında birikimli toplam değeri hesaplanırken, her seferinde *k*-1 tane gözleme ait log-olabilirlik oranı değerinin baştan itibaren toplanmasına gerek kalmaz. Bu da algoritmanın hızlı sonuç vermesini sağlar.

Çalışmada ele alınacak olan Gauss gürültü modeli için Eşitlik (2) ile verilen log-olabilirlik oranı (Basseville ve Nikiforov, 1993)

$$s_i = \frac{\mu_1 - \mu_0}{\sigma^2}\left(x_i - \frac{\mu_1 + \mu_0}{2}\right) \tag{6}$$

biçimine dönüşür. Bu eşitliğin, Gauss olasılık yoğunluk fonksiyonu kullanılarak türetimi Ek-1'de sunulmuştur.

Kısmi en iyi CUSUM algoritması

Bir önceki bölümde anlatılan algoritma en iyi (optimal) CUSUM algoritması olarak adlandırılmaktadır. Bu algoritmada μ_0, μ_1 ve σ^2 parametrelerinin bilindiği kabul edilmektedir. Ancak bu kabul, pratik uygulamalar için çok gerçekçi değildir. Bu parametreler için kestirim ve önsel bilginin (a priori information) kullanılması yoluyla CUSUM algoritmasının kısmi en iyi (suboptimal) türevleri oluşturulmuştur. μ_0 ve σ^2 parametreleri, mevcut gözlem değerlerinden en çok olabilirlik

kestiricisi (maximum likelihood estimator) ile tahmin edilebilir. μ_1 parametresinin belirlenmesi amacıyla önerilen yaklaşımlardan bir tanesi *değişim büyüklüğü* (*change magnitude*) için olabilecek en küçük değerin seçilmesi, bir diğeri ise değişim büyüklüğü için en olası değerin seçilmesidir (Basseville ve Nikiforov, 1993). Bu yaklaşımlarda δ değişim büyüklüğü olmak üzere μ_1 parametresi

$$\mu_1 = \mu_0 + \delta \tag{7}$$

biçiminde ifade edilir. Değişim büyüklüğünün belirlenmesindeki hassasiyet, sinyal hakkındaki önsel bilgi doğruluğuna bağlıdır. Bu yaklaşımlar, doğaları gereği kısmi en iyi birer çözüm üretecektir.

Eşitlik (7) ile verilen μ_1 değeri Eşitlik (6)'da yerine konulursa anlık log-olabilirlik oranı

$$\begin{aligned} s_i &= \frac{\mu_0 + \delta - \mu_0}{\sigma^2}\left(x_i - \frac{\mu_0 + \delta + \mu_0}{2} \right) \\ &= \frac{\delta}{\sigma^2}\left(x_i - \mu_0 - \frac{\delta}{2} \right) \end{aligned} \tag{8}$$

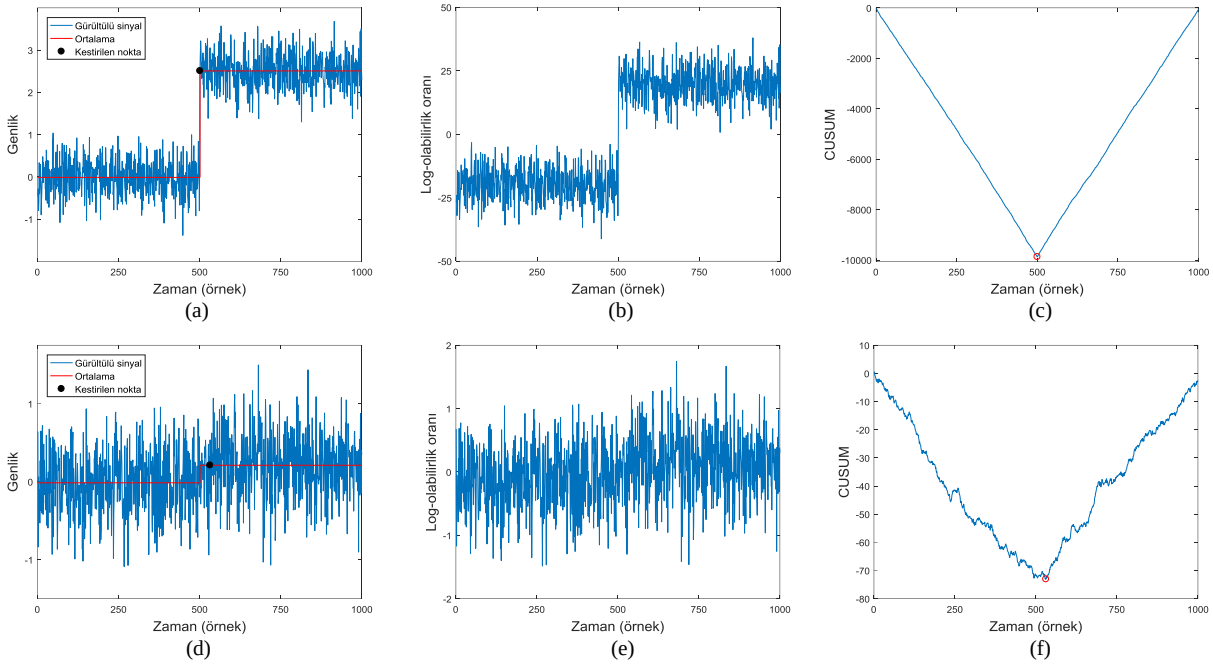
olarak elde edilir. Çalışma kapsamında kısmi en iyi algoritma için kullanılacak olan bu eşitlikte, μ_0 ve σ^2 en çok olabilirlik kestirim değerleri, δ ise önsel bilgiye dayalı en olası değer olarak alınmıştır.

Ayrıca, δ parametresi hakkındaki önsel bilgi doğruluğunun toplam algoritma başarımına etkisini incelemek için önsel bilgi olarak girilen bu parametre değerine belli ölçüde rasgele sapma verilerek testler yapılmıştır. Bu sapma miktarının nasıl modellendiği konusundaki bilgi, test sonuçları ile birlikte benzetim sonuçlarının sunulduğu bölümde verilecektir.

Değişim büyüklüğünün birikimli toplam değerlerine etkisi

Şekil 2 (a) ve (d)'de değişim noktası içeren iki farklı sinyal örneği gösterilmiştir. Bu sinyal örneklerine karşılık gelen anlık log-olabilirlik oranı değerleri ve bu değerlerin birikimli toplamı sırasıyla (b), (e) ve (c), (f) grafikleri ile verilmiştir. Şekil 2 (c), (f)'den görüldüğü gibi birikimli toplam değerleri değişimden önce negatif yöne eğilimli, değişimden sonra pozitif yöne eğilimlidir. Bunun nedeni Eşitlik (2) ile verilen log-olabilirlik oranı değerleri incelenerek açıklanabilir. Değişime kadarki örnekler, θ_0 parametreli dağılımdan geldiği için $f_{\theta_0}(x_i)$ ile verilen olabilirlik değeri, $f_{\theta_1}(x_i)$ ile verilen olabilirlik değerinden büyük olur. Bu durumda logaritma fonksiyonun argümanı 1'den küçük olacağı için logaritma değerleri negatif çıkar. Değişimden sonra ise tam tersi durum olacak ve logaritma değerleri pozitif çıkmaya başlayacaktır. Bu değerler, Şekil 2 (b)'de gösterilmiştir. Dolayısıyla, log-olabilirlik oranı değerlerinin birikimli toplamı alınarak elde edilen eğrinin yön değiştirdiği nokta, sinyalde aranılan değişim noktasına karşılık gelir (Şekil 2-c). Bu nokta, CUSUM eğrisinin en küçük değeri aldığı noktaya karşılık geldiği için problem en küçük değerin yerini bulma işlemine döner.

Şekil 2 (d)'de gösterildiği gibi değişim büyüklüğünün azalması durumunda, değişimden önce ve sonraki ortalama değer parametreleri (μ_0 ve μ_1) birbirine çok yakın olduğunda, Şekil 2 (e)'de gösterildiği gibi olabilirlik fonksiyonu değerleri birbirine yakın çıkar. Dolayısıyla bu değerlerin oranının logaritması da Şekil 2 (e)'de görüldüğü gibi, değişim noktasından önce ve sonra sıfıra yakın pozitif ve negatif değerler alır. Bununla birlikte, μ_0 ve μ_1 parametreleri arasında küçük de olsa fark olduğundan, log-olabilirlik oranı değişimden önceki örneklerin çoğunluğunda negatif ve değişimden sonraki örneklerin çoğunluğunda ise pozitif çıkmaktadır. Bu durum, Şekil 2 (f)'de verilen log-olabilirlik oranlarının birikimli toplamından görülmektedir. Birikimli toplam eğrisi Şekil 2 (c)'deki duruma kıyasla Şekil 2 (f)'de daha salınımlı bir yapıya dönüştüğü için en küçük değerin ortaya çıktığı nokta, farklı gürültü gerçekleştirimleri için daha çok sapma gösterecektir. Şekil 2 (d)'de verilen gürültülü sinyal örneği için değişim noktası Şekil



Şekil 2. Basamak değişimi içeren iki farklı sinyal (a, d) ile karşılık gelen log-olabilirlik oranı (b, e) ve birikimli toplam değerleri (c, f)

2 (f)'de gösterildiği gibi 30 örnek değerinde bir hata ile kestirilmiştir. Farklı gürültü gerçekleştirimleri için hata değerlerinin dağılımı başarım değerlendirmesi bölümünde histogramlar yardımıyla görselleştirilecektir.

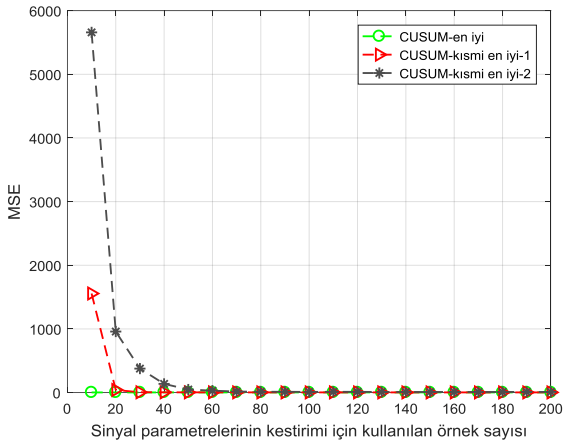
Başarım değerlendirmesi

Parametre kestiriminde kullanılan örnek sayısının değişim noktası kestirimine etkisi

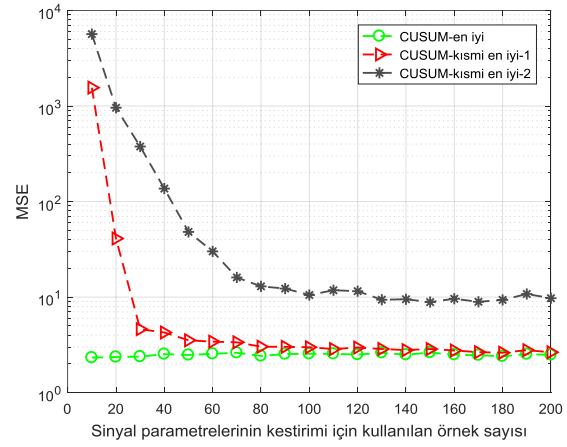
Bu çalışma kapsamında tüm parametre değerlerinin bilindiği en iyi CUSUM algoritmasının yanı sıra kısmi en iyi CUSUM algoritmasının başarımı da incelenmiştir. Bu amaçla kısmi en iyi-1 ve kısmi en iyi-2 isimleri verilen iki algoritma oluşturulmuştur. Bunlardan birincisinde δ parametresi hakkındaki önsel bilginin kesin doğru olduğu, ikincisinde önsel bilgide hata olduğu varsayılmıştır. Bu hata, en çok olması beklenen değişim büyüklüğü değeri δ' olmak üzere, $[\delta'-0.5\delta', \delta'+0.5\delta']$ aralığında bir biçimli dağılıma sahip bir rasgele değişken olarak modellenmiştir.

Her iki kısmi en iyi CUSUM algoritmasında da μ_0 ve σ^2 parametreleri gürültü segmentindeki belli sayıda örnek kullanılarak en çok olabilirlik tahmin edicisi ile kestirilmiştir. Bu parametrelerin kestirimi için kullanılan örnek sayısının değişim noktası kestirim başarımına etkisi benzetimlerle test edilmiştir. Her örnek sayısında, 1000 farklı gürültü gerçekleştirimi için kestirim yapılmış ve ortalama karesel hata (mean square error, MSE) değerleri hesaplatılmıştır. Şekil 3'te ortalama karesel hata değerleri doğrusal ve logaritmik ölçekte gösterilmektedir. Bu şekilden görüldüğü gibi μ_0 ve σ^2 parametrelerinin kestirimi için kullanılan örnek sayısı, kısmi en iyi 1 ve 2 algoritmaları için sırasıyla 50 ve 100'ün üzerine çıktığında MSE değerleri açısından hemen hemen doyuma ulaşıldığı söylenebilir. Bu nedenle, sonuçları bir sonraki bölümde sunulacak olan hata histogramlarının karşılaştırılması için yapılan testlerde, her iki kısmi en iyi CUSUM algoritması için de sinyal parametrelerinin kestiriminde kullanılan örnek sayısı 100 olarak alınmıştır.

δ parametresi hakkındaki önsel bilginin yetersiz ya da hatalı olması durumunda algoritma başarımının azalacağı Şekil 3'teki kırmızı ve siyah eğriler karşılaştırıldığında görülmektedir.



(a)



(b)

Şekil 3. Sinyal parametrelerinin kestiriminde kullanılan örnek sayısının değişim noktası kestirim başarımına etkisi. Ortalama karesel hata değerleri (a) doğrusal ve (b) logaritmik ölçekte verilmiştir.

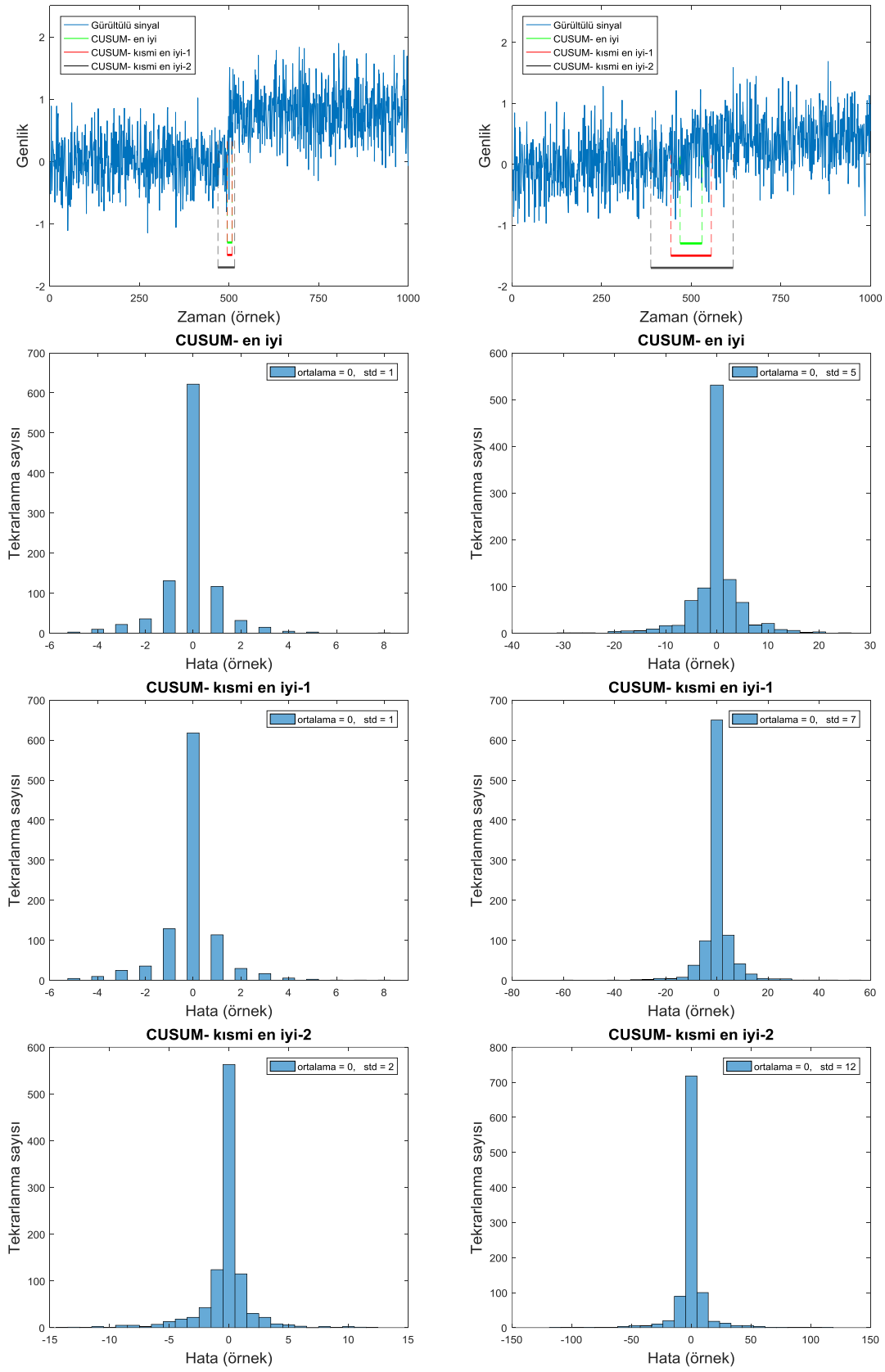
Önsel bilgi doğruluğu arttıkça başarım, kırmızı eğriye yakınsayacak, azaldıkça ise siyah eğride olduğu gibi bozulacaktır. Ayrıca bu şekilde en iyi CUSUM algoritmasının başarım değerleri de karşılaştırma yapılabilmesi amacıyla sunulmuştur. En iyi CUSUM algoritmasında μ_0 ve σ^2 parametrelerinin değerlerinin bilindiği kabul edildiği için değişen örnek sayısı ile başarım yaklaşık aynı kalmaktadır. Bu algoritmanın başarımını gösteren yeşil renkli eğride logaritmik ölçekte gözlenen küçük değişimler, her örnek değeri için rasgele üretilen gürültülerden kaynaklanmaktadır.

CUSUM algoritması ile değişim noktası kestirim başarımının değerlendirilmesi

CUSUM algoritmasının farklı türevleri için değişim noktası kestirim başarımını test etmek amacıyla kolay ve zor problem olarak tanımlanabilecek iki farklı değişim büyüklüğü değeri için testler yapılmıştır. Şekil 4'te sol tarafta verilen sinyal için değişim büyüklüğü 0.8, sağdaki sinyal için 0.4 alınmıştır. Varyansın karekökü olan standart sapma (σ) parametresi her iki durum için 0.4 değerinde sabit alınmıştır. Bu sinyaller için CUSUM algoritmasının farklı türevleri ile elde edilen değişim noktası kestirim hatası histogramları aynı şekilde gösterilmektedir. Bu örnek benzetim sinyalleri incelendiğinde, değişim büyüklüğü azaldıkça değişim noktası algılama probleminin zorlaşacağı görsel olarak fark edilmektedir.

Değişim büyüklüğünün göreli olarak daha fazla olduğu soldaki sinyal için her üç algoritmanın da değişim noktasını benzer bir başarımla kestirdiği histogram grafiklerinden görülmektedir. Değişim büyüklüğü azaltıldığında ise algoritmaların başarımları arasında fark oluşmaya başlamış ve en iyi CUSUM algoritması kısmi en iyi CUSUM algoritmalarına kıyasla daha yüksek başarım göstermiştir. Bu durumda en iyi, kısmi en iyi-1, kısmi en iyi-2 algoritmaları için ortalama hata değerleri sıfır iken standart sapma değerleri sırasıyla yaklaşık 5, 7 ve 12 örnektir. Ayrıca sinyaller üzerinde her üç algoritma için kestirim sonuçlarının değer aralığı yatay çizgilerle gösterilmiştir. Bu çizgilerin uzunluğu histogram sonuçlarında gözlendiği gibi değişim büyüklüğü arttığı durumda birbirine yakınken, değişim büyüklüğü azaldığında birbirinden farklıdır. Sağdaki sinyal örneği için kestirim değer aralıklarının küçükten büyüğe doğru sırasıyla en iyi, kısmi en iyi-1 ve kısmi en iyi-2 algoritmalarıyla elde edildiği ilgili yatay çizgilerden görülmektedir.

Bu sonuçlar göstermektedir ki değişim noktası algılama problemi zorlaştıkça sinyal parametrelerinin kestirim ve önsel bilgi doğruluğunun önemi artmaktadır.



Şekil 4. Basamak değişimi içeren iki farklı sinyal için CUSUM algoritmasının farklı türevleri ile elde edilen değişim noktası kestirimi hata histogramları

Tablo 1. CUSUM algoritmasının ortalama hesaplama süresinin veri uzunluğu ile değişimi

YÖNTEM	HESAPLAMA SÜRESİ			
	$N=10^2$	$N=10^3$	$N=10^4$	$N=10^5$
CUSUM	10 μs	22 μs	161 μs	1700 μs

Hesaplama Süresi

Pratikte, kestirim doğruluğunun yanı sıra sistemin hesaplama karmaşıklığı ve kolay gerçekleştirilebilmesi gibi gereksinimler de dikkate alınmak zorundadır. Günümüzde mobil elektronik aygıtların yaygın kullanımı nedeniyle düşük güç tüketimi önemli bir gereksinim halini almıştır. Bu tür aygıtlarda hesaplama karmaşıklığının çok fazla olduğu algoritmalar yerine, karmaşıklığı daha düşük algoritmalar tercih edilmektedir. CUSUM algoritmasının Eşitlik (5) ile verildiği gibi özyinelemeli yapıda gerçekleştirilebiliyor olması hesaplama karmaşıklığını önemli ölçüde azaltmaktadır. Bu yönüyle CUSUM, özellikle günümüzde yaygınlaşan gömülü sistem uygulamaları için kolay uygulanabilen ve hızlı sonuçlar üretebilen bir algoritma olarak tercih edilebilir.

Tablo 1'de farklı veri uzunluğu değerleri için CUSUM algoritmasının hesaplama süreleri verilmiştir. Bu süre değerleri, algoritmaların çok sayıda tekrar çalıştırılması sonucu ortalama alınarak elde edilmiştir. Benzetimler, Intel i7 işlemcili bir bilgisayarda Matlab ortamında gerçekleştirilmiştir. Veri uzunluğunun 10^2, 10^3, 10^4 ve 10^5 örnek alındığı durumlar için elde edilen hesaplama süreleri incelendiğinde algoritmanın 10^4 örnek uzunluğuna kadar mikro saniyeler mertebesinde hızlı sonuç verdiği görülmektedir.

Sonuç

Bu çalışmada, CUSUM algoritmasının değişim noktası kestirim başarımı analiz edilmiştir. Algoritmanın, gözle tespit edilmesi zor olan problemlerde bile, yüksek başarıma sahip olduğu bazı örneklerle gösterilmiştir. Bu başarının altında yatan temel neden, gözlemlere ait gerçek dağılım fonksiyonun bilindiği durumun ele alınmasıdır. Pratikte, kabul edilen dağılım fonksiyonundan sapmalar olması durumunda algoritma başarımının bir miktar azalacağı göz önünde bulundurulmalıdır. Ayrıca, sinyal parametrelerinin tam olarak bilinmediği gerçek uygulamalarda da algoritma başarımının bir miktar azaldığı çalışma kapsamında gösterilmiştir. Bununla birlikte sinyal parametreleri için yapılan kestirimlerin ve önsel bilginin doğruluğu arttığında, kabul edilebilir seviyede bir başarımın elde edilebileceği benzetimlerle ortaya konulmuştur.

Çalışmada, Gauss gürültü modeli için analiz yapılmıştır. Basitlik ve düşük hesaplama karmaşıklığı gibi önemli avantajlara sahip bu algoritmanın başarımının farklı gürültü modelleri için de test edilmesi önemli bir çalışma konusudur. Örneğin, bazı haberleşme kanallarında ortaya çıkan darbesel gürültü (impulsive noise) ortamlarında algoritmanın başarılı olduğunun gösterilmesinin çevrimiçi haberleşme uygulamaları için önemli olacağı değerlendirilmektedir.

Ek-1

Gauss rasgele değişkeni için olasılık yoğunluk fonksiyonu

$$f\left(x_i;\mu_j,\sigma^2\right)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{\left(x_i-\mu_j\right)^2}{2\sigma^2}} \tag{9}$$

eşitlik (2)'de yerine konularak, gerekli sadeleştirme işlemleri yapılırsa Gauss gürültü

modeli için Eşitlik (6) ile verilen log-olabilirlik oranı ifadesi elde edilebilir:

$$
\begin{aligned}
s_i &= \ln\left(\frac{\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x_i-\mu_1)^2}{2\sigma^2}}}{\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x_i-\mu_0)^2}{2\sigma^2}}}\right) \\
&= \ln\left(e^{\frac{-(x_i-\mu_1)^2+(x_i-\mu_0)^2}{2\sigma^2}}\right) \\
&= \frac{-(x_i-\mu_1)^2+(x_i-\mu_0)^2}{2\sigma^2} \\
&= \frac{2x_i(\mu_1-\mu_0)+\mu_0^2-\mu_1^2}{2\sigma^2} \\
&= \frac{\mu_1-\mu_0}{\sigma^2}\left(x_i-\frac{\mu_1+\mu_0}{2}\right)
\end{aligned}
$$

Kaynaklar

Basseville, M., Nikiforov, I.V., (1993). *Detection of Abrupt Changes: Theory and Application*, 528, Prentice-Hall, NJ, USA.

Brodsky, B.E., Darkhovsky, B.S., (1993). *Nonparametric Methods in Change-point Problems*, 209, Kluwer Academic Publishers, MA, USA.

Granjon, P., *eds*. (2012). The CUSUM algorithm - a small review, Technical report, Gipsa-Lab, Grenoble, France.

Köse, M., Taşcıoğlu, S., Telatar, Z., (2011). Bayesian change point analysis using the phase distribution of complex signals, *Communications Faculty of Sciences University of Ankara Series A2-A3*, **53**, 1, 15-23.

Köse, M., Taşcıoğlu, S., Telatar, Z., (2015). The effect of transient detection errors on RF fingerprint classification performance, *Proceedings*, 14th International Conference on Circuits, System, Electronics, Control & Signal Processing (CSECS2015), 89-93, Konya.

Montgomery, D.C., (2013). *Introduction to Statistical Quality Control*, 754, John Wiley & Sons Inc., NJ, USA.

Page, E.S., (1954). Continuous inspection schemes. *Biometrika*, **41**, 1, 100-115.

Ruanaidh, J.J.K. O, Fitzgerald, W.J., (1996). *Numerical Bayesian Methods Applied to Signal Processing*, 244, Springer-Verlag, NY, USA.

Taşcıoğlu, S., Üreten, O., (2009). Bayesian wideband spectrum segmentation for cognitive radios, *Proceedings*, 18th International Conference on Computer Communications and Networks, 1-6, San Francisco.

Taşcıoğlu, S., Üreten, O., Telatar, Z., (2010). Impact of noise power uncertainty on the performance of wideband spectrum segmentation, *Radioengineering*, **19**, 4, 561-566.

Üreten, O., Serinken, N., (1999). Detection of radio transmitter turn-on transients, *Electronic Letters*, **35**, 23, 1996-1997.

Üreten, O., Serinken, N., (2005). Bayesian detection of Wi-Fi transmitter RF fingerprints, *Electronic Letters*, **41**, 6, 373-374.

Performance evaluation of CUSUM algorithm for changepoint estimation

Extended abstract

Change point estimation is a widely used method in many different applications from communication, biomedical and speech signal processing to seismic data analysis, from statistical process control to financial data analysis. Many statistical signal processing techniques have been developed for solving this estimation problem. The methods used in cases where the distribution is known and unknown are classified as parametric and nonparametric methods, respectively. Knowing the distribution of measurements provides a strong knowledge, which can be exploited in parametric changepoint estimation methods for difficult changepoint estimation problems.

The CUSUM algorithm is one of the methods used for change point analysis. This algorithm is based on the likelihood ratio statistic, where the likelihood values are obtained with the parameters before and after the change. In this algorithm, changepoint detection problem is reduced to finding the location of the minimum value of cumulative sum of log-likelihood ratio values.

In this work, the change point estimation performance of the parametric CUSUM algorithm has been analyzed. This parametric method has been employed to estimate the abrupt change in the mean value of the Gaussian random process. In the literature, this problem is also known as the step change detection. Equations of CUSUM algorithm based on the log-likelihood ratio has been derived for the Gaussian noise model and the basic principle of the method has been explained with examples.

For the practical cases where the signal parameters are unknown, suboptimal solutions are considered. In this study, the performance of the suboptimal CUSUM algorithm as well as the optimal CUSUM algorithm has also been investigated. For this purpose, two different suboptimal algorithms have been considered. The difference between these suboptimal algorithms is the degree of accuracy of knowledge about the change magnitude parameter. For the first algorithm, it was assumed that a priori information about the change magnitude was perfect, and for the second one it was imperfect. For both of these algorithms, mean and variance parameters of the Gaussian model were estimated by a maximum likelihood estimator using different number of samples in the noise segment. The effect of the number of samples used for this aim on the performance of the changepoint estimation was analyzed and the number of samples were determined for the sufficient performance through simulations. It has been demonstrated that as the degree of accuracy of knowledge about the change magnitude parameter decreases, the changepoint estimation performance decreases.

The change point estimation performance has been presented through simulations and compared for the CUSUM algorithm versions considered in this study. It has been demonstrated that estimation and a priori information accuracy for the signal parameters become more important as the change point detection problem becomes more difficult.

In practice, besides the accuracy of estimation, the requirements such as computational complexity and ease of implementation must also be considered. Today, due to the widespread use of mobile electronic devices, low power consumption has become an important requirement. In such devices, algorithms with lower computational complexity are preferred. The fact that the CUSUM algorithm can be implemented in a recursive structure significantly reduces computational complexity. This advantage of the method has been demonstrated through simulations for the signals with different lengths. Due to the significant advantages such as simplicity and having low computational complexity, the CUSUM algorithm may be preferred for embedded systems applications that are becoming widespread.

In this study, the Gaussian noise model is considered. Testing the performance of the CUSUM algorithm for different noise models is considered as an important future work. For example, evaluating the performance of the algorithm for impulsive noise environments which is encountered in some communication channels may be important for online communication applications.

Keywords: *Change point estimation, step change, CUSUM algorithm, parametric methods, a priori information, log-likelihood ratio.*