

NEWTON OPTİMİZASYONA DAYALI ENBÜYÜK OLABİLİRLİK BULANIKLIK VE GÖRÜNTÜ TANIMA

Faruk SARI

Tübitak MAM Bilişim Teknolojileri Araştırma Enstitüsü
P.K. 21, 41470, Gebze, Kocaeli
e-posta: faruk@btae.mam.gov.tr

Anahtar Sözcükler: Görüntü ve Bulanıklık Tanıma, Newton Optimizasyon, “Fisher Scoring” Yöntemi.

ABSTRACT

In a previous paper we have formulated the blur identification problem as a ML problem by using a Newton type optimization method based on “The Method of Scoring” [1]. We also have provided a comparative analysis of EM and Newton type optimization methods.

In this paper the same structure is extended to the blur and image parameter identification. Although in the previous study we have chosen the convergence rate experimentally, in this study we analyze the convergence of the proposed method and provide new results about the selection of convergence rate. The main advantages of the proposed method are that it needs a few iterations to converge and in cases where the second derivative of the log-likelihood function is not available it gives a possibility of solution based on a Newton-type algorithm requiring only the first derivatives. The performance of the method is shown using simulated data.

1. GİRİŞ

Görüntü onarım problemi bulanık ve gürültülü halinden özgün görüntünün kestirilmesi şeklinde tanımlanabilir. Görüntü onarım problemi genelde kötü-koşullanmış (ill-conditioned) bir problemdir. Verilen data üzerindeki küçük oynamalar çözümde büyük değişimlere sebep olur [2]. Bulanıklık transfer fonksiyonun doğrudan tersi genelde yüksek frekanslı bileşenlerde büyük genliğe sahip olmakta, böylelikle bu frekanslarda gürültü aşırı kuvvetlendirilmektedir. Görüntü onarımının ilk adımı bozulmanın bilinmediği durumlarda bulanıklığın tanınması işlemidir.

Bulanıklık tanıma konusundaki ilk çalışmalar, sadece birim daire üzerindeki sıfırları olan nokta yayılım fonksiyonu (PSF)’in tanınması üzerine yoğunlaşmıştır. Görüntü ve bulanıklık tanıma problemine ML yaklaşımı

Lagendijk [3], Lay ve Katsaggelos [4] tarafından önerilmiştir. Buradaki temel fikir, gözlenen görüntüdeki en olabilir parametre değerlerini bulmaktır.

Lagendijk [5], bulanıklık tanıma problemini zorlanmış bir ML problemi olarak ele aldı ve dürümsel bayır temelli bir yordam uygulayarak doğrusal olmayan optimizasyon problemi olarak çözdü. Bayır iniş (Gradient Descent) optimizasyon teknikleri sistemin performansında önemli rol oynayan bir yakınsama parametresi içerirler. Bayır iniş algoritmasının yakınsaması için yakınsama parametresinin uygun bir değerinde seçilmesi gerekir.

Son zamanlarda olasılıksal modeller konusundaki çalışmalar ile, Beklenti Enbüyükleme (Expectation Maximization-EM) yöntemi optimizasyon için bayır temelli yöntemlere bir alternatif olarak kullanılmaktadır. Özgün ML problemi doğrusal olmayan bir optimizasyon problemi olmasına rağmen, EM algoritması lineer eşitlikleri içerir. Bu doğrusallığa rağmen, EM algoritması her durumda başarı ile kullanılabilen bir çözüm değildir. Birçok değişik örnekte hem E adımı hem de M adımı kontrol edilemez yapıda olabilmektedir. Böyle bir ikilem içinde kaldığında çoğu istatistikçi hemen “scoring” gibi diğer algoritmalara dönmektedirler. Radner and Walker [6] geniş bir şekilde makalelerinde tartışmışlar ve super-doğrusal (quasi-Newton, Scoring yöntemi) ve ikinci derece (Newton) yöntemler genelde EM’e tercih edilmesi gerektiği sonucuna varmışlardır.

2. PROBLEMİN TANIMI

Özgün görüntü, $f(i,j)$ şeklinde gösterilir ve 2-B özbağlanımlı (AR) model olarak,

$$f = Af + v \quad (1)$$

şeklinde verilmektedir. Burada A , AR görüntü model katsayılarından oluşan matris, v ise modelleme gürültüsüdür.

Görüntü biçimleme süreci tipik olarak birinci tip integral denklemi olarak modellenir. Matrisel formda,

$$g = Df + w \quad (2)$$

şeklinde verilir. Burada f bilinmeyen gerçek görüntüyü, g, f 'in gürültülü ve bulanık halı, D 'de bulanıklığı temsil eden büyük ölçüde kötü koşullanmış bulanıklık matrisidir ve w modelleme gürültüsüdür.

(1) ve (2) ifadelerinde gürültü terimlerinin her ikisinin de sıfır ortalamalı, sırasıyla $Q_v = \sigma_v^2 I$ ($\sigma_v^2 > 0$) ve $Q_w = \sigma_w^2 I$ ($\sigma_w^2 > 0$) ortak değişinti matrislerine sahip tektürel Gauss dağılımlı beyaz gürültü süreci olduğu ve sırasıyla f ve v ile ilişkisiz olduğu kabul edilir.

ML kestirici, gözlenen bulanık gürültülü gözlemin en benzer sonucu olan θ 'nın değişik değerlerini bulur. Bilinmeyen parametreler θ vektörü ile gösterilir ve parametre vektörü θ 'nın ML kestirimi,

$$\hat{\theta}_{ML} = \arg \left\{ \max_{\theta \in \Theta} L(\theta) \right\} = \arg \left\{ \max_{\theta \in \Theta} \log p(g; \theta) \right\} \quad (3)$$

olarak tanımlanır.

Burada $L(\theta)$, θ 'nın (log)-olabilirlik işlevi, $p(g; \theta)$, verilen θ gözlem görüntüsünün önsel olasılık yoğunluk işlevi ve Θ ise, θ parametrelerinin değişim bölgesini gösterir.

ML görüntü tanıma problemi,

$$\hat{\theta}_{ML} = \arg \left\{ \min_{\theta \in \Theta} L(\theta) \right\} = \arg \min \left\{ \log(\det|P|) + g^T P^{-1} g \right\} \quad (4)$$

şeklinde ifade edilmiş ve ortak değişinti matrisi P ,

$$P = E \left\{ D(I - A)^{-1} v + w \right\} (D(I - A)^{-1} v + w)^T \quad (5)$$

$$P = \sigma_v^2 D(I - A)^{-1} (I - A)^{-T} D^T + \sigma_w^2 I$$

olarak tanımlanmıştır [5].

Olabilirlik işlevi $L(\theta)$, D ve A matrislerinin blok-dönen (block-circulant) yapısı yüzünden frekans bölgesinde etkin ve kolayca gerçekleştirilebilir [7].

3. NEWTON OPTİMİZASYON VE "FISHER SCORING" YÖNTEMİ

θ , $L(\theta)$ 'nin içinde karmaşık bir biçimde bulunduğundan tek bir basit kapalı-form oluşturmak imkansızdır. Bu yüzden, $L(\theta)$ 'yi en küçük yapmak için, sayısal optimizasyon yöntemleri kullanmak gerekmektedir. Genelde bu yöntemler eğer başlangıç tahmini gerçek minimuma yakın ise ML kestirimini üretecek aksi durumda yakınsama gerçekleşmeyecek ya da sadece lokal minimuma yakınsayacaktır.

Bayır iniş yöntemleri, özellikle değişken sayısının çok fazla olduğu durumlarda önemli bir şekilde etkisiz hale gelmektedirler. Bu zorluğu aşmanın en bilinen yolu Hessian matrisinin hesaplandığı Newton yöntemini kullanmaktır [8].

Newton yöntemi,

$$\hat{\theta}^{i+1} = \hat{\theta}^i - H \left(\hat{\theta}^i \right)^{-1} g \left(\hat{\theta}^i \right) \quad (6)$$

olarak yazılabilir. Burada $H(\theta)$, $L(\theta)$ 'nin Hessian matrisi ve $g(\theta)$, $L(\theta)$ 'nin gradyan vektörüdür. Eğer Hessian pozitif tanımlı ise, extremum bir (yerel) minimumdur.

Görüntü tanıma problemi için, $L(\theta)$ θ 'nin elemanlarından birine göre kısmi türevi alınarak,

$$\frac{\partial L(\theta)}{\partial \theta_i} = \frac{\partial}{\partial \theta_i} [\log \det|P|] + \frac{\partial}{\partial \theta_i} [g^T P^{-1} g] \quad (7)$$

şeklinde verilir. Ayrık Fourier Dönüşüm (AFD) düzleminde,

$$\frac{\partial L(\theta)}{\partial \theta_i} = \sum_{u,v} \left\{ \left[\frac{1}{P(u,v)} - \frac{|G(u,v)|^2}{P(u,v)^2} \right] \frac{\partial P(u,v)}{\partial \theta_i} \right\} \quad (8)$$

şeklinde verilir. Burada $\frac{\partial P(u,v)}{\partial \theta_i}$ türevi bulanıklık ve görüntü katsayılarına göre yazılırsa,

$$\frac{\partial P}{\partial d(m,n)} = \sigma_v^2 (I - A)^{-1} (I - A)^{-T} \left\{ E_{m,n} D^T + E_{m,n}^T D \right\} \quad (9)$$

$$\frac{\partial P}{\partial a(k,l)} = \sigma_v^2 D(I - A)^{-1} (I - A)^{-T} D^T \left\{ E_{k,l} (I - A)^{-1} + E_{k,l}^T (I - A)^{-T} \right\} \quad (10)$$

ifadeleri elde edilir. Burada $E_{m,n}$, m, n elemanda değeri 1 diğer yerlerde 0 olan ilkel matrisi göstermektedir.

A ve D matrislerinin blok-dönen yapısı sayesinde,

$$\frac{\partial P(u,v)}{\partial d(m,n)} = \sigma_v^2 \frac{1}{|1-A(u,v)|^2} \left\{ D(-u,-v) \exp \left[-2\pi j \left(\frac{um}{M} + \frac{vn}{N} \right) \right] + D(u,v) \exp \left[2\pi j \left(\frac{um}{M} + \frac{vn}{N} \right) \right] \right\} \quad (11)$$

$$\frac{\partial P(u,v)}{\partial d(k,l)} = \sigma_v^2 \frac{|D(u,v)|^2}{|1-A(u,v)|^2} \left\{ \frac{1}{1-A(u,v)} \exp \left[-2\pi j \left(\frac{uk}{M} + \frac{vl}{N} \right) \right] + \frac{1}{1-A(u,v)} \exp \left[2\pi j \left(\frac{uk}{M} + \frac{vl}{N} \right) \right] \right\} \quad (12)$$

şeklinde yazılabilir.

Şimdi Hessian matrisini oluşturmaya çalışalım. Birçok problemde, amaç fonksiyonu ve gradyanı hesaplanabilmekte ancak Hessian'ı hesaplanamaktadır. Bu tür problemlerde sıklıkla "Fisher Scoring" yöntemi kullanılmaktadır [9].

"Scoring" yöntemi, $\left. \frac{\partial^2 L(\theta)}{\partial \theta^2} \right|_{\theta=\theta_i} \approx -I(\theta_i)$ kabulünü yapar

burada $I(\theta)$, "Fisher Information" matrisidir. Şimdi yukarıdaki yaklaşıklığı kullanarak değiştirilmiş Newton optimizasyon yöntemini görüntü ve bulanıklık tanıma için oluşturalım. $I(\theta)$ matrisi,

$$[I(\theta)]_{i,j} = -E \left[\frac{\partial^2 L(\theta)}{\partial \theta_i \partial \theta_j} \right] = E \left[\frac{\partial L(\theta)}{\partial \theta_i} \frac{\partial L(\theta)}{\partial \theta_j} \right]_{\substack{i=1,2,\dots,p \\ j=1,2,\dots,p}} \quad (13)$$

olarak tanımlanır.

(13) nolu ifadenin hesaplanmasında kullanılan gerekli eşitlikler Ek A'da verilmiştir. Buna göre,

$$\frac{\partial L(\theta)}{\partial \theta_i} = \text{tr} \left(\frac{\partial P}{\partial \theta_i} P^{-1} \right) - g^T \frac{\partial P^{-1}}{\partial \theta_i} g \quad (14)$$

yazılır ve (13) nolu ifade içine yerleştirildiğinde,

$$[I(\theta)]_{i,j} = \text{tr} \left(P^{-1} \frac{\partial P}{\partial \theta_i} \right) \text{tr} \left(P^{-1} \frac{\partial P}{\partial \theta_j} \right) + \text{tr} \left(P^{-1} \frac{\partial P}{\partial \theta_i} \right) E \left\{ g^T \frac{\partial P^{-1}}{\partial \theta_j} g \right\} + E \left\{ g^T \frac{\partial P^{-1}}{\partial \theta_i} g \right\} \text{tr} \left(P^{-1} \frac{\partial P}{\partial \theta_j} \right) + E \left\{ g^T \frac{\partial P^{-1}}{\partial \theta_i} g g^T \frac{\partial P^{-1}}{\partial \theta_j} g \right\} \quad (15)$$

ifadesi elde edilir. Sonuç olarak,

$$[I(\theta)]_{i,j} = 2 \text{tr} \left(P^{-1} \frac{\partial P}{\partial \theta_i} P^{-1} \frac{\partial P}{\partial \theta_j} \right) \quad (16)$$

ifadesi elde edilir [1]. AFD ile ifade edildiğinde,

$$[I(\theta)]_{(i,j)} = 2 \sum_{u,v} \left\{ \frac{1}{P^2(u,v)} \frac{\partial P(u,v)}{\partial \theta_i} \frac{\partial P(u,v)}{\partial \theta_j} \right\} \quad (17)$$

ifadesine ulaşılır. Ulaşılan sonuç oldukça yalın ve kolay işlem yapılabilir bir ifadedir.

"Scoring" algoritmasını kullanarak yeni Newton algoritması,

$$\hat{\theta}^{i+1} = \hat{\theta}^i + I(\theta)^{-1} \frac{\partial L(\theta)}{\partial \theta} \quad (18)$$

şeklini alır.

Bulanıklık katsayıları için,

$$d^{i+1}(m,n) = d^i(m,n) + I_d^{-1} \frac{\partial L(\theta)}{\partial d(m,n)} \quad (19)$$

Görüntü katsayıları için,

$$a^{i+1}(k,l) = a^i(k,l) + I_a^{-1} \frac{\partial L(\theta)}{\partial a(k,l)} \quad (20)$$

ifadeleri elde edilir. Hesaplama yoğunluğunu azaltmak için bilinmeyen PSF'in gerçel ve simetrik olduğu kabul edilmiştir.

4. YÖNTEMİN YAKINSAMASI

EM algoritması monoton yakınsayan bir algoritma olmakla birlikte aynı zamanda bir yakınsama parametresine ihtiyaç duymaz. Bayır iniş, quasi-Newton ve Newton yöntemleri de dahil olmak üzere gradyan temelli optimizasyon teknikleri yakınsamayı basitçe garanti edemezler [10]. Bayır iniş yöntemi için, adım aralığı β her adımında uygun bir değerde seçilmeli ya da çok küçük bir değere sabitlenmelidir. Ancak bu çok küçük değer kötü koşullanmış matris yüzünden çok çok küçük olacak ve yakınsama çok çok yavaş gerçekleşecektir. Yakınsama parametresi β için uygun değer, $0 < \beta < 2/\lambda_{\max}$ aralığındaki değer olup burada, λ_{\max} log-olabilirlik fonksiyonun Hessian'ın en büyük özdeğeridir.

Eğer bir matrisin özdeğer yayılımı geniş ise, diğer bir deyişle koşul sayısı (condition number) çok büyük ise kötü koşullanmış olarak tanımlanır. Koşul sayısı şu şekilde tanımlanır:

$$\text{cond}(A) = \frac{\lambda_M(A)}{\lambda_m(A)} \quad (21)$$

A'nın koşul sayısı olarak adlandırılır. Burada, $\lambda_M(A)$ ve $\lambda_m(A)$ sırasıyla A'nın en büyük ve en küçük özdeğerleridir. Kötü koşullanmış bir matrisde de bayır iniş algoritması kullanıldığında yakınsama çok yavaş olmaktadır. Koşul sayısı 1'e yaklaştıkça

yakınsama hızlanmakta ve süper-doğrusal yakınsama oranı olarak adlandırılmaktadır.

Newton yöntemi ile sonuca ulaşabilmek için, dürümsel süreç genelde çözüme yakın noktalardan başlama ihtiyacı duyar, aksi takdirde Hessian tanımsız olabilir ve yakınsama gerçekleşmez [11].

“Fisher Scoring” yöntemi tanımsız Hessian matrisi problemini aşabilir ancak quadratik olmayan ve doğrusal olmayan optimizasyon probleminde yöntem aşağıda verilen kurala uyan β adım aralığı değerine ihtiyaç duyar [12].

$$\|I + \beta IH(\theta^{i-1})\| < 1 \quad (22)$$

Bayır iniş ve EM algoritmalarının yakınsaması koşul sayısı ile ölçülen log-olabilirlik fonksiyonun şekline bağlıdır. Koşul sayısı çok büyük ise yineleme yönü zig zag çizer ve yakınsama çok yavaş olur.

“Fisher Scoring” yönteminde kullandığımız kabulü kullanarak yakınsama katsayısı β için uygun değerleri bulmaya çalışalım.

(22) nolu ifade “Fisher Scoring” yaklaşıklığı kullanılarak,

$$\|I - \beta H(\theta^i)H(\theta^{i-1})\| < 1 \quad (23)$$

halini alır. Hesaplama kolaylığı için, $H(\theta^i)H(\theta^{i-1}) = R(\theta)$ kabulünü yapalım.

(22) nolu yakınsama koşulu ifadesini kullanarak,

$$\left(\sum_{i,j} ((R(\theta))^2)_{ij} - \text{tr}((R(\theta))^2) \right) \beta^2 - 2\text{tr}(R(\theta)) + (N-1) + \text{tr}((R(\theta))^2) < 0 \quad (24)$$

şartı elde edilir. Bu ikinci derece ifadeyi sağlayan yakınsama değerleri hesaplanabilir. Gerçek köklerin bulunabilmesi için,

$$\left(\text{tr}(R(\theta)^2) \right) > \left(\sum_{i,j} ((R(\theta))^2)_{ij} - \text{tr}((R(\theta))^2) \right) \left((N-1) + \text{tr}((R(\theta))^2) \right) \quad (25)$$

koşulu sağlanmalıdır. Burada tr trace operatörünü, $(R)^2$ işlemi matrisin eleman eleman karesini, N ise Hessian matrisinin köşegen eleman sayısını göstermektedir.

Birçok durumda yukarıdaki koşulun kolayca sağlanması zordur. Bu koşulun sağlanabilmesi için R matrisinin köşegen elemanlarına uygun pozitif değerlerin eklenmesi gerekir.

İki kere türevi alınabilen bir fonksiyon ancak ve ancak Hessian matrisi tüm değerler için pozitif tanımlı ise kesin dışbükeydir (strictly convex). Kesin dışbükey fonksiyonların tek çözümü olup Newton metodu global olarak yakınsar. Ancak çoğu zaman log-olabilirlik fonksiyonu dışbükey olmamakta ve ML kestirimi çoğu zaman tek çözümlü olmamakta ve çözüm kolayca yerel minimumlara takılabilmektedir. Bunu aşmanın yolu gerçek sonuca yakın değerler ile çözüme başlamaktır.

5. DENEYSEL SONUÇLAR

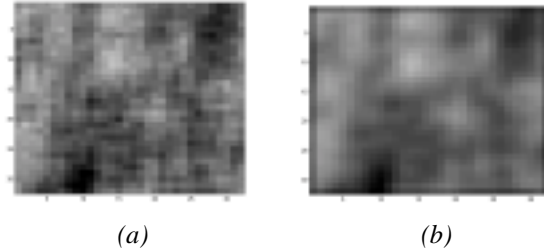
Deneysel sonuçlar (32 x 32) boyutunda suni olarak üretilen görüntüler üzerinde bayır iniş, EM ve Scoring yöntemleri kullanılarak gerçekleştirilmiştir. Suni görüntü, asimetrik yarı düzlem (Non-symmetric Half Plane-NSHP) bir AR modeli kullanılarak üretilmiştir ve bağımsız beyaz gauss gürültüsü bu suni görüntüye eklenmiştir. Kullanılan NSHP katsayıları ve yerleşimi aşağıda verilmiştir.

$$a_1 = -0.449 \quad a_2 = 0.630 \quad a_3 = 0.127 \quad a_4 = 0.671$$

$$a(k,l) = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & & \end{bmatrix}$$

Şekil.1. NSHP görüntü destek bölgesi

Gürültü eklenmiş olan görüntü (3 x 3) boyutunda PSF bulanıklık katsayıları ile bulanıklaştırılmış ve toplamsal beyaz gauss gürültüsü eklenerek bozulmuş gözlem görüntüsü Şekil 2. de görüldüğü gibi elde edilmiştir.



Şekil.2. (a) Suni olarak oluşturulan özgün görüntü (b) Bulanık ve gürültülü hali

Kullanılan (3 X 3) boyutundaki PSF değerleri,

$$d = \begin{bmatrix} d_{-1,-1} & d_{-1,0} & d_{-1,1} \\ d_{0,-1} & d_{0,0} & d_{0,1} \\ d_{1,-1} & d_{1,0} & d_{1,1} \end{bmatrix} = \begin{bmatrix} 0.075 & 0.12 & 0.075 \\ 0.12 & 0.22 & 0.12 \\ 0.075 & 0.12 & 0.075 \end{bmatrix}$$

Tablo 1’de gürültülü ve bulanık suni görüntüler üzerinde bulanıklık tanıma sonuçları herbir algoritma ve iterasyon sayısına göre özetlenmiştir.

	$d_{-1,-1}$	$d_{-1,0}$	$d_{0,-1}$	$d_{0,0}$	Dürüm Sayısı
Bayır İniş	0.0812	0.1142	0.1142	0.2179	93
EM	0.0932	0.1053	0.1053	0.2241	9
Scoring Yöntemi	0.0695	0.1108	0.1108	0.2369	3

Tablo 1. PSF katsayılarının tanınması, SNR=20 dB.

Tablo 2’de ise AR katsayılarına ilişkin sonuçlar verilmiştir.

	a_1	a_2	a_3	a_4	Dürüm Sayısı
Bayır İniş	-0.453	0.598	0.132	0.668	87
EM	-0.439	0.618	0.136	0.680	11
Scoring Yöntemi	-0.431	0.609	0.129	0.675	5

Tablo 2. AR katsayılarının tanınması, SNR=20 dB.

6. SONUÇ

Bu çalışmada, ML bulanıklık ve görüntü tanıma problemi Newton tipi optimizasyon algoritması kullanılarak çözülmüştür. Bayır iniş, EM ve Scoring algoritmaları suni olarak üretilen görüntüler üzerine uygulanmıştır. Bayır iniş algoritmasının uygulanmasında yakınsama parametresinin seçimi yöntemin başarımın doğrudan etkilemektedir. Newton yöntemi sonuca birkaç adımda ulaşırken bayır iniş algoritmasının yakınsaması çok fazla adım gerektirmektedir. Bununla birlikte Newton yöntemi uygulaması zor ve hem Hessian’ın hem de gradyanın hesaplanması gerektiğinden hesap yükü fazladır. Aynı zamanda çoğu zaman log-olabilirlik fonksiyonunun şekli yüzünden birinci türev mevcut ancak ikinci türev mevcut olmamakta ve Hessian hesaplanamamaktadır. Bu durum Scoring yöntemi kullanılarak sadece gradyan değerleri kullanılarak aşılabilmektedir.

Önerilen yöntem için yakınsama parametresinin seçimi ile ilgili analiz yapılmıştır. Uygun yakınsama parametresinin seçimi problemin doğası gereği zor olabilmektedir. Uygun yakınsama parametreleri kullanılarak yapılan optimizasyon sonucunda oldukça hızlı sonuca yakınsayan bir yapı elde edilmektedir.

7. KAYNAKLAR

[1] F. Sarı, “Newton Type Optimization for Maximum Likelihood Blur Identification and Restoration”, *IEEE Int. Conf. On Acoustics, Speech and Signal Processing*, Vol.4, sayfa 3716-3719, 2002.

- [2] H.C. Andrews and B.R. Hunt, *Digital Image Restoration*, Englewood Cliffs, NJ :Printce-Hall, 1977.
- [3] R.L. Lagendijk, J. Biemond, D.E. Boeke, ‘Identification and Restoration of Noisy Blurred Images Using the Expectation-Maximization Algorithm’ *IEEE Tans. Acoust. Speech Sig. Proc.*, ASSP-38(7), sayfa 1180-1191, 1990.
- [4] K.T. Lay, A.K. Katsaggelos, ‘Image Identification and Restoration Based on Expectation-Maximization Algorithm’ *Optical Engineering*, Vol.29, No.5, sayfa 436-445, 1990.
- [5] R.L. Lagendijk, A.K. Katsaggelos, J. Biemond, ‘Iterative Identification and Restoration of Images’, *IEEE Int. Conf. On Acoustics, Speech and Signal Processing*, sayfa 992-995, 1988.
- [6] Radner, R.A., Walker, H.F. “Mixtures densities, maximum likelihood, and the EM algorithm”, *SIAM Review*, Vol. 26, sayfa 195-239, 1984.
- [7] R.C. Gozalez, R.E. Woods, *Digital Image Processing*, Addison-Wesley Pub., 1992.
- [8] P.E. Gill, W. Murry, M.H. Wright, *Practical Optimization*, Academic Press, 1981.
- [9] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Printice Hall, 1993.
- [10] M. Hanke, J.G.Nagy, C.Vogel, “ Quasi-Newton approach to nonnegative image restorations”, *NH Elsevier, Linear Algebra and its Applications*, Vol.316, sayfa 223-226, 2000.
- [11] E. Arian, S. Ta’san, “ Analysis of the Hessian for Aerodynamic Optimization:Inviscid Flow” *NASA CR-198328 ICASE Report*, No.96-28, 1996.
- [12] L. Xu, M.I. Jordan, ‘On Convergence of the EM Algorithm for Gaussian Mixtures’, *Neural Computation*, Vol.8, sayfa 129-151, 1996.

Ek-A

$$\frac{\partial}{\partial \theta_i} [\log \det |P|] = \text{tr} \left(P^{-1} \frac{\partial P}{\partial \theta_i} \right) \quad (\text{A1})$$

$$\frac{\partial P^{-1}}{\partial \theta_i} = -P^{-1} \frac{\partial P}{\partial \theta_i} P^{-1} \quad (\text{A2})$$

$$E(y^T z) = \text{tr} \left\{ E \left(zy^T \right) \right\} \quad (\text{A3})$$

$$E(gg^T) = P \quad (\text{A4})$$

$$E(y^T A y y^T B y) = \text{tr}(AC) \text{tr}(BC) + 2 \text{tr}(ACBC) \quad (\text{A5})$$

Burada $C = E(yy^T)$ ve A ve B simetrik matrislerdir.