KISITLI OPTİMİZASYON

EŞİTLİK KISITLI ÇOK DEĞİŞKENLİ OPTİMİZASYON

Eşitlik kısıtları içeren çok değişkenli bir optimizasyon problemi

$$Min\, f = f(x)$$

$$g_i(x) = 0 \quad , \quad i = 1,2, \ldots , m$$

olarak tanımlanır. Burada, $x = (\, x_1 \,,\, x_2 \,,\, \ldots \,, x_n\,)$ ve $m \leq n$ dir. $m > n$ durumunda problemin çözümü yoktur.

Bu tür problemlerin çözümü için geliştirilen çeşitli optimizasyon yöntemlerinden en kullanışlı olanı Lagrange Çarpanları yöntemidir.

Lagrange Çarpanları yöntemi:

$$Min\, f = f(x)$$

$$g_i(x) = 0 \quad , \quad i = 1,2, \ldots , m$$

problemi verilmiş olsun. Bu bir optimizasyon problemidir.

Her bir $g_i(x)$ kısıtı için λ_i Lagrange çarpanı kullanılarak

$$L(X,\lambda) = f(\, x_1 \,,\, x_2 \,, \ldots \,, x_n\,) - \sum_{i=1}^{m} \lambda_i [g_i(\, x_1 \,,\, x_2 \,, \ldots \,, x_n\,)]$$

fonksiyonu oluşturulur. Bu fonksiyona Lagrange fonksiyonu adı verilir.

Lagrange fonksiyonu yardımıyla kısıtlı problem, kısıtsız bir optimizasyon problemine dönüştürülür.

$f(x)$ in, $g_i(x) = 0$ kısıtları altında optimize edilmesi; $L(X,\lambda)$ Lagrange fonksiyonunun optimize edilmesine eşdeğerdir. Burada $f(x)$ ve $g_i(x)$ fonksiyonlarının sürekli ve türevlenebilir olduğu varsayılır.

$L(X,\lambda)$ fonksiyonu için kısmi türevler alınıp sıfıra eşitlenirse,

$$\frac{\partial L}{\partial x_j} = \frac{\partial f}{\partial x_j} - \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} = 0 \quad ; \quad j = 1\,, 2\,, \ldots \,, n$$

$$\frac{\partial L}{\partial \lambda_i} = 0$$

denklemleri elde edilir. Bu denklem sisteminin çözümüyle $(x_0\,, \lambda_0)$ ile gösterilen kritik(durağan) nokta bulunur.

Bulunan durağan(kritik) noktanın yerel minimum yada yerel maksimum olduğunu anlayabilmek için H_B sınırlı Hessian matrisi,

$$H_B = \begin{bmatrix} 0_{mxm} & P \\ P^T & Q \end{bmatrix}$$

biçiminde oluşturulur. Burada;

0_{mxm} : mxm boyutlu sıfır matrisi,

$$P = \begin{bmatrix} \nabla g_1(x) \\ \nabla g_2(x) \\ \vdots \\ \nabla g_m(x) \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1(x)}{\partial x_1} & \frac{\partial g_1(x)}{\partial x_2} & \cdots & \frac{\partial g_1(x)}{\partial x_n} \\ \frac{\partial g_2(x)}{\partial x_1} & \frac{\partial g_2(x)}{\partial x_2} & \cdots & \frac{\partial g_2(x)}{\partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial g_m(x)}{\partial x_1} & \frac{\partial g_m(x)}{\partial x_2} & \cdots & \frac{\partial g_m(x)}{\partial x_n} \end{bmatrix}$$ (kısıtların birinci kısmi türevleri),

$$Q = \left[\frac{\partial^2 L(X,\lambda)}{\partial x_i \partial x_j}\right] \quad ; \quad j = 1, 2, \dots, n \;\; ve \;\; i = 1, 2, \dots, m$$

anlamındadır.

$L(X,\lambda)$ Lagrange fonksiyonunun $(x_0\,, \lambda_0)$ durağan noktasındaki H_B sınırlı Hessian matrisi için,

1) H_B sınırlı Hessian matrisinin $(2m+1)\,'inci$ esas minörü ile başlayan son $(n-m)$ tane esas minörünün işareti, $(-1)^{m+1}$ ile başlayarak sırasıyla değişiyorsa bulunan durağan nokta yerel maksimum noktadır.
2) H_B matrisinin $(2m+1)\,'inci$ esas minöründen başlayarak son $(n-m)$ tane esas minörünün işaretleri $(-1)^m$ ile aynıysa, bulunan durağan nokta yerel minimum noktadır.

Yukarıdaki şartları sağlayan noktalar kesinlikle optimal noktalardır. Ancak bu şartları sağlamayan fakat aslında optimal olan bir durağan nokta olabilir. Bu sebeple optimal(ekstremum) nokta için gerekli ve yeterli şartlar aşağıdaki gibi verilebilir:

A matrisi,

$$A = \begin{bmatrix} 0_{mxm} & P \\ P^T & Q - \mu I \end{bmatrix}$$

olarak $(x_0\,, \lambda_0)$ noktasında oluşturulur. Burada P ile Q daha önce tanımlamış olduğumuz matrislerdir. I : birim matris, μ ise bilinmeyen parametredir.

Eğer, $|A| = 0$ polinomunun $(n - m)$ tane kökünün hepsi

1) Pozitif ise x_0 noktası yerel minimum noktadır.
2) Negatif ise x_0 noktası yerel maksimum noktadır.

Örnek: Aşağıdaki problemin çözümünü Lagrange çarpanları yöntemiyle bulunuz.

$$Min\ f(x) = x_1^2 + x_2^2 + x_3^2 + x_4^2$$

$$g_1(x):\ x_1 + 2x_2 + 3x_3 + 5x_4 - 10 = 0$$

$$g_2(x):\ x_1 + 2x_2 + 5x_3 + 6x_4 - 15 = 0$$

Çözüm: Lagrange fonksiyonu aşağıdaki gibi oluşturulur.

$$L(X, \lambda) = x_1^2 + x_2^2 + x_3^2 + x_4^2 - \lambda_1(x_1 + 2x_2 + 3x_3 + 5x_4 - 10) - \lambda_2(x_1 + 2x_2 + 5x_3 + 6x_4 - 15)$$

Lagrange fonksiyonundaki bilinmeyenlere göre ayrı ayrı kısmi türevler alınır.

$$\frac{\partial L}{\partial x_1} = 2x_1 - \lambda_1 - \lambda_2 = 0$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - 2\lambda_1 - 2\lambda_2 = 0$$

$$\frac{\partial L}{\partial x_3} = 2x_3 - 3\lambda_1 - 5\lambda_2 = 0$$

$$\frac{\partial L}{\partial x_4} = 2x_4 - 5\lambda_1 - 6\lambda_2 = 0$$

$$\frac{\partial L}{\partial \lambda_1} = -(x_1 + 2x_2 + 3x_3 + 5x_4 - 10) = 0$$

$$\frac{\partial L}{\partial \lambda_2} = -(x_1 + 2x_2 + 5x_3 + 6x_4 - 15) = 0$$

Elde edilen denklem sisteminin ortak çözümünden $(x_0\,,\lambda_0)$ durağan noktası,

$$(x_0\,,\lambda_0) = (\,x_{10}\,,x_{20}\,,x_{30}\,,x_{40}\,;\,\lambda_{10}\,,\lambda_{20}\,) = (\frac{-5}{74}\,,\ \frac{-10}{74}\,,\ \frac{155}{74}\,,\ \frac{60}{74}\,;\ \frac{-90}{37}\,,\ \frac{85}{37})$$

olarak bulunur. Bu noktanın durumunu araştırmak için H_B matrisi oluşturulur.

$$P = \begin{bmatrix} \nabla g_1(x) \\ \nabla g_2(x) \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 5 \\ 1 & 2 & 5 & 6 \end{bmatrix}$$

$$Q = \left[\frac{\partial^2 L(X,\lambda)}{\partial x_i \partial x_j}\right] = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

olup,

$$H_B = \begin{bmatrix} 0_{mxm} & P \\ P^T & Q \end{bmatrix} = \left[\begin{array}{cc|cccc} 0 & 0 & 1 & 2 & 3 & 5 \\ 0 & 0 & 1 & 2 & 5 & 6 \\ \hline 1 & 1 & 2 & 0 & 0 & 0 \\ 2 & 2 & 0 & 2 & 0 & 0 \\ 3 & 5 & 0 & 0 & 2 & 0 \\ 5 & 6 & 0 & 0 & 0 & 2 \end{array}\right]$$ elde edilir.

H_B sınırlı Hessian matrisinin $(2m+1)\,'inci$ esas minörü [2x2+1 = 5'inci esas minörü] ile başlayan son $(n-m)$ tane [4-2 = 2 tane] esas minörünün işaretine bakılacak. Yani 5'inci ve 6'ıncı esas minörlerin işaretine bakılacak.

5'inci esas minör: $\begin{vmatrix} 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 & 5 \\ 1 & 1 & 2 & 0 & 0 \\ 2 & 2 & 0 & 2 & 0 \\ 3 & 5 & 0 & 0 & 2 \end{vmatrix} = 40$ (pozitif)

6'ıncı esas minör: $\begin{vmatrix} 0 & 0 & 1 & 2 & 3 & 5 \\ 0 & 0 & 1 & 2 & 5 & 6 \\ 1 & 1 & 2 & 0 & 0 & 0 \\ 2 & 2 & 0 & 2 & 0 & 0 \\ 3 & 5 & 0 & 0 & 2 & 0 \\ 5 & 6 & 0 & 0 & 0 & 2 \end{vmatrix} = 296$ (pozitif)

5'inci ve 6'ıncı esas minörlerin işaretleri, $(-1)^m = (-1)^2 = 1 \quad (pozitif)$ ile aynıdır yani 2) durumuna uymaktadır. O halde $x_0 = (\frac{-5}{74}, \frac{-10}{74}, \frac{155}{74}, \frac{60}{74})$ noktası, yerel minimum noktadır.

Örnek: Aşağıdaki problemi Lagrange çarpanları yöntemiyle çözünüz.

$$Min\, f(x) = x_1^2 + x_2^2 + x_3^2$$

$$g_1(x):\ 4x_1 + x_2^2 + 2x_3 - 14 = 0$$

Çözüm:

$$L(X, \lambda) = x_1^2 + x_2^2 + x_3^2 - \lambda(4x_1 + x_2^2 + 2x_3 - 14)$$

$$\frac{\partial L}{\partial x_1} = 2x_1 - 4\lambda = 0$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - 2\lambda x_2 = 0$$

$$\frac{\partial L}{\partial x_3} = 2x_3 - 2\lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = -(4x_1 + x_2^2 + 2x_3 - 14) = 0$$

Elde edilen denklem sisteminin ortak çözümünden (x_0, λ_0) durağan noktası olarak;

$$(x_0, \lambda_0)_1 = (2,\ 2,\ 1,\ 1)$$

$$(x_0, \lambda_0)_2 = (2, -2,\ 1,\ 1)$$

$$(x_0, \lambda_0)_3 = (2.8,\ 0,\ 1.4,\ 1.4)$$

noktaları bulunur.

$$P = [\nabla g_1(x)] = [4 \quad 2x_2 \quad 2]$$

$$Q = \left[\frac{\partial^2 L(X,\lambda)}{\partial x_i \partial x_j}\right] = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2-2\lambda & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

olup sınırlı Hessian matrisi,

$$H_B = \begin{bmatrix} 0_{mxm} & P \\ P^T & Q \end{bmatrix} = \left[\begin{array}{c|ccc} 0 & 4 & 2x_2 & 2 \\ \hline 4 & 2 & 0 & 0 \\ 2x_2 & 0 & 2-2\lambda & 0 \\ 2 & 0 & 0 & 2 \end{array}\right]$$ olarak elde edilir

$$(2m+1) = 2x1 + 1 = 3$$

$$(n-m) = 3 - 1 = 2$$

3. esas minörle başlayan son 2 tane esas minöre, yani 3. ve 4. esas minörlere işaret bakımından bakılacak. (Her bir durağan nokta için ayrı ayrı)

$(\boldsymbol{x_0}, \boldsymbol{\lambda_0})_{\boldsymbol{1}} = (\,\mathbf{2}\,,\ \mathbf{2}\,,\ \mathbf{1}\,,\ \mathbf{1}\,)$ **<u>için</u>:** $\quad H_B = \begin{bmatrix} 0 & 4 & 4 & 2 \\ 4 & 2 & 0 & 0 \\ 4 & 0 & 0 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$ olur.

3'üncü esas minör: $\begin{vmatrix} 0 & 4 & 4 \\ 4 & 2 & 0 \\ 4 & 0 & 0 \end{vmatrix} = -32$ (negatif)

4'üncü esas minör: $\begin{vmatrix} 0 & 4 & 4 & 2 \\ 4 & 2 & 0 & 0 \\ 4 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{vmatrix} = -64$ (negatif)

3'üncü ve 4'üncü esas minörlerin işaretleri, $(-1)^m = (-1)^1 = -1 \quad (negatif)$ ile aynıdır yani 2) durumuna uymaktadır. O halde $x_0 = (\,2\,,\ 2\,,\ 1\,)$ noktası, yerel minimum noktadır.

$(\boldsymbol{x_0}\,,\boldsymbol{\lambda_0})_2 = (\,\mathbf{2}\,,-\,\mathbf{2}\,,\;\;\mathbf{1}\,,\;\;\mathbf{1}\,)$ **<u>için</u>:** $H_B = \begin{bmatrix} 0 & 4 & -4 & 2 \\ 4 & 2 & 0 & 0 \\ -4 & 0 & 0 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$ olur.

3'üncü esas minör: $\begin{vmatrix} 0 & 4 & -4 \\ 4 & 2 & 0 \\ -4 & 0 & 0 \end{vmatrix} = -32$ (negatif)

4'üncü esas minör: $\begin{vmatrix} 0 & 4 & -4 & 2 \\ 4 & 2 & 0 & 0 \\ -4 & 0 & 0 & 0 \\ 2 & 0 & 0 & 2 \end{vmatrix} = -64$ (negatif)

3'üncü ve 4'üncü esas minörlerin işaretleri, $(-1)^m = (-1)^1 = -1 \quad (negatif)$ ile aynıdır yani 2) durumuna uymaktadır. O halde $x_0 = (\,2\,,\;-2\,,\;\;1\,)$ noktası, yerel minimum noktadır.

$(\boldsymbol{x_0}\,,\boldsymbol{\lambda_0})_3 = (\,\mathbf{2.8}\,,\;\;\mathbf{0}\,,\;\;\mathbf{1.4}\,,\;\;\mathbf{1.4}\,)$ **<u>için</u>:** $H_B = \begin{bmatrix} 0 & 4 & 0 & 2 \\ 4 & 2 & 0 & 0 \\ 0 & 0 & -0.8 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$ olur.

3'üncü esas minör: $\begin{vmatrix} 0 & 4 & 0 \\ 4 & 2 & 0 \\ 0 & 0 & -0.8 \end{vmatrix} = 12.8$ (pozitif)

4'üncü esas minör: $\begin{vmatrix} 0 & 4 & 0 & 2 \\ 4 & 2 & 0 & 0 \\ 0 & 0 & -0.8 & 0 \\ 2 & 0 & 0 & 2 \end{vmatrix} = 32$ (pozitif)

3'üncü ve 4'üncü esas minörlerin işaretleri, her iki duruma da uymuyor. O halde $x_0 = (\,2.8\,,\;\;0\,,\;\;1.4\,)$ noktası, ne yerel minimum nede yerel maksimum noktadır.

Acaba bu nokta kesinlikle uç nokta değil midir? Bunu öğrenmek için A matrisine bakılmalıdır.

$$A = \begin{bmatrix} 0_{mxm} & P \\ P^T & Q-\mu I \end{bmatrix} = \begin{bmatrix} 0 & 4 & 2x_2 & 2 \\ 4 & 2-\mu & 0 & 0 \\ 2x_2 & 0 & 2-2\lambda-\mu & 0 \\ 2 & 0 & 0 & 2-\mu \end{bmatrix}$$

$|A| = 0$ 'dan; $5\mu^2 - 6\mu - 8 = 0$ polinomu elde edilir.

Bunun kökleri ise $\mu_1 = 2 \;\; ve \;\; \mu_2 = -0.8$ olarak bulunur.

Köklerin hepsi ne pozitif nede negatif olduğundan bu nokta kesinlikle optimal nokta (maksimum yada minimum) değildir.

Diğer iki kökün durumuna A matrisi ile tekrar bakacak olursak (tekrar bakmaya gerek yok ama uygulamayı görmek için):

$(\boldsymbol{x_0}, \boldsymbol{\lambda_0})_1 = (\mathbf{2}, \; \mathbf{2}, \; \mathbf{1}, \; \mathbf{1})$ **için:**

$|A| = 0$ 'dan; $9\mu^2 - 26\mu + 16 = 0$ polinomu elde edilir.

Bunun kökleri ise $\mu_1 = 2 \;\; ve \;\; \mu_2 = 8/9$ olarak bulunur. Köklerin hepsi pozitif olduğundan bu nokta yerel minimum noktadır.

$(\boldsymbol{x_0}, \boldsymbol{\lambda_0})_1 = (\mathbf{2}, -\mathbf{2}, \; \mathbf{1}, \; \mathbf{1})$ **için:**

$|A| = 0$ 'dan; $9\mu^2 - 26\mu + 16 = 0$ polinomu elde edilir.

Bunun kökleri ise $\mu_1 = 2 \;\; ve \;\; \mu_2 = 8/9$ olarak bulunur. Köklerin hepsi pozitif olduğundan bu nokta yerel minimum noktadır.

ÖDEV 1: Aşağıdaki problemin çözümünü (Lagrange çarpanları yöntemiyle) bulunuz.

$$Min\ f(x) = x_1^2 + 2x_2^2 + 10x_3^2$$

$$Kısıtlar:$$

$$x_1 + x_2^2 + x_3 = 5$$

$$x_1 + 5x_2 + x_3 = 7$$

ÖDEV 2: Aşağıdaki problemin çözümünü (Lagrange çarpanları yöntemiyle) bulunuz.

$$Max\ f(x) = 2x_1 + x_2 + 10$$

$$Kısıtlar:$$

$$x_1 + 2x_2^2 = 3$$

ÖDEV 3: Aşağıdaki problemin çözümünü (Lagrange çarpanları yöntemiyle) bulunuz.

$$Min\ f(x) = x_1^2 + x_2^2 + x_3^2$$

$Kısıtlar$:

$$x_1 + x_2 + 3x_3 = 2$$

$$5x_1 + 2x_2 + x_3 = 5$$