Optimizasyon Teknikleri

Ders Notu – 6

DOĞRUSAL OLMAYAN (NONLINEER) OPTİMİZASYON

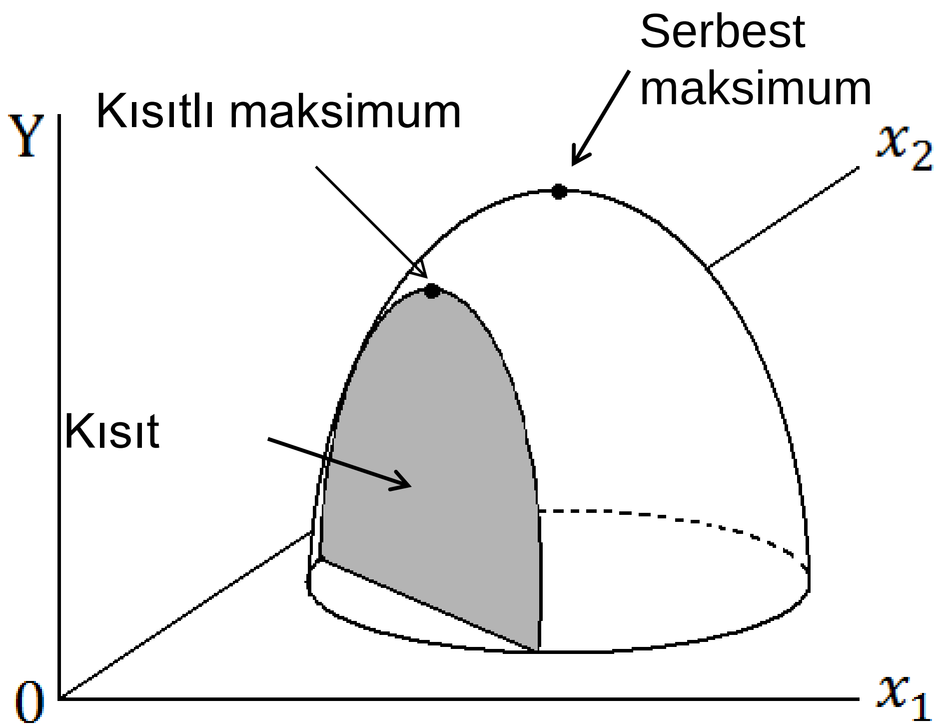
Prof. Dr. Bilal ALATAŞ

İÇERİK

- **KISITLI OPTİMİZASYON PROBLEMİ**
  - **DÖNÜŞTÜRME METOTLARI**
    - **PENALTY METOTLARI**
      - **EXTERIOR PENALTY METODU**
      - **INTERIOR PENALTY METODU**
      - **EXTENDED-INTERIOR PENALTY METODU**
    - **AUGMENTED LAGRANGIAN METODU**
    - **PENALTY VE AUGMENTED LAGRANGIAN METOTLARI**
  - **DİREK METOTLAR (SEARCH METHODS)**
    - **METHOD OF FEASIBLE DIRECTIONS**
    - **GRADIENT PROJECTION METODU (İz Düşüm Matrisi Hesabı)**

KISITLI OPTIMIZASYON

Grafiksel olarak , serbest optimum ve kısıtlı optimum Arasındaki fark şu şekilde gösterilebilir:



KISITLI OPTIMIZASYON PROBLEMI

- Genel bir kisitli optimizasyon problem ifadesi:

  ✓ ***Minimize***

$$f(x) = f(x_1, x_2, \ldots\ldots, x_n)$$

  ✓ ***kisit***

$$\begin{aligned} g_j(x) &\le 0 \qquad j = 1, \ldots\ldots, n_g \\ h_k(x) &= 0 \qquad k = 1, \ldots\ldots, n_e \end{aligned}$$

Non lineer problemlerin çözümü için iki tip yöntem vardir:

1- <u>Dönüstürme metotlari (transformation methods):</u> Kisitli problemleri kisitsiz problemlere dönüstürerek çözerler.
2- <u>Direk metotlar (direct methods)</u> : Kisitli problemi oldugu gibi kabul ederek "*search direction*" konsepti ile çözerler.

DÖNÜSTÜRME METOTLARI

Dönüstürme metotlari (transformation methods) kisitli problemleri ardisik kisitsiz problemlere dönüstürerek çözerler:

- Penalty metotlari
  - ✓ Exterior penalty metodu
  - ✓ Interior penalty metodu (Barrier metodu)
  - ✓ Extended-interior penalty metodu

- Augmented Lagrangian metodu

Yer Değiştirme Metodu

Maks

$$y = 5x_1x_2$$

$$2x_1 + x_2 = 100$$ Kısıt

$$x_2 = 100 - 2x_1$$

$$y = 5x_1x_2$$

$$y = 5x_1(100 - 2x_1)$$

$$= 500x_1 - 10x_1^2$$

$$\frac{dy}{dx_1} = 500 - 20x_1 = 0$$

$$-20x_1 = -500$$

Kritik değer $\therefore x_1 = \frac{-500}{-20} = 25$

$$\frac{dy}{dx_1} = 500 - 20x_1 = 0$$

$$\frac{d^2y}{dx_1^2} = -20 < 0 \quad \therefore \text{relative max}$$

$$\therefore \text{if } x_1 = 25 \text{ then } 100 = 2(25) + x_2$$

$$\therefore 100 - 50 = 50 = x_2$$

PENALTY METOTLARI

- Kisitli optimizasyon problemi

  ✓ ***Minimize***

  $$f(x) = f(x_1, x_2, \ldots\ldots, x_n)$$

  ✓ ***Kisit***

  $$g_j(x) \le 0 \qquad j = 1, \ldots\ldots, n_g$$
  $$h_k(x) = 0 \qquad k = 1, \ldots\ldots, n_e$$

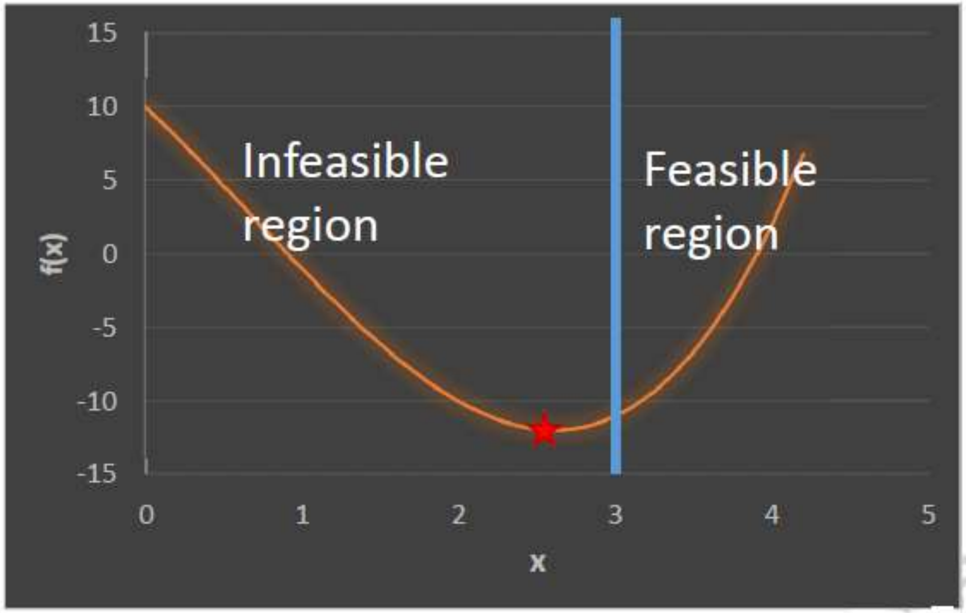
- Esdeger kisitsiz optimizasyon problemi

  ✓ ***Minimize***

  $$\Phi(x, r_p) = f(x) + r_p P(x)$$

  ✓ *rp* penalty katsayisi

  ✓ *P(**x**)* penalty fonksiyonu: kisit ihlalinde devreye giren, kullanilan kisit türüne ve çözüm metoduna göre degisik tipte olan bir fonksiyon).



Minimize

$f(x) = x^3 - 10x - 2x^2 + 10$

Subject to $g(x) = x \geq 3$

Or, $g(x) = x - 3 \geq 0$

$$F(x, R) = f(x) + R\langle g(x)\rangle^2$$

Where,

$\langle g(x)\rangle = 0$ if $x \geq 3$

$\langle g(x)\rangle = g(x)$ otherwise

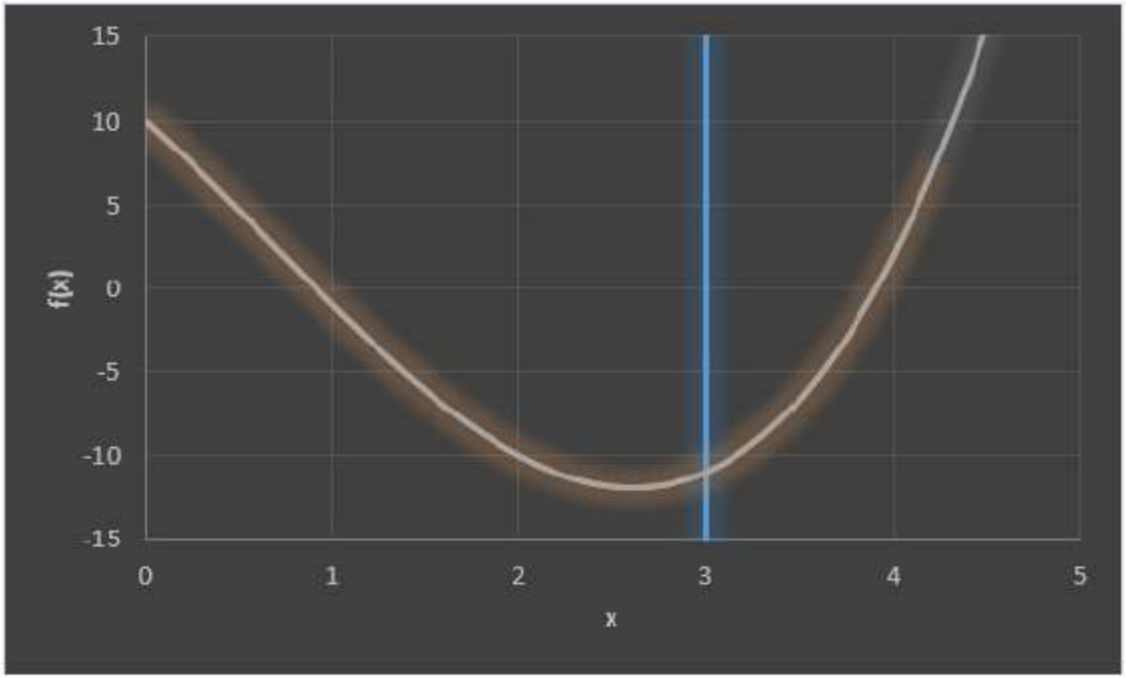
$\langle\ \rangle$ function $min(g, 0)$

$$F(x, R) = (x^3 - 10x - 2x^2 + 10) + R\langle x - 3\rangle^2$$

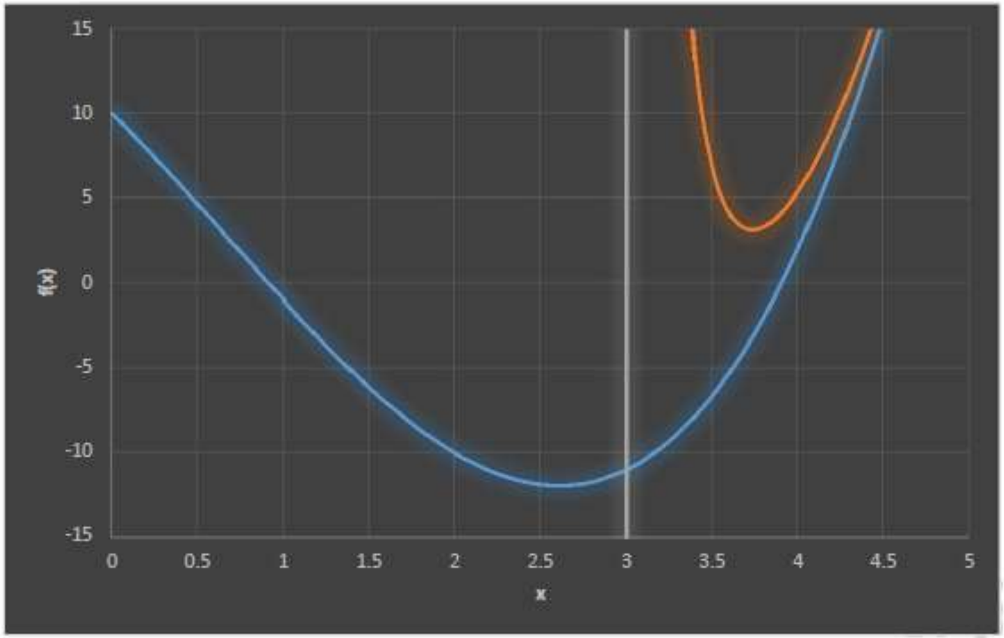
$$F(x, R) = (x^3 - 10x - 2x^2 + 10) + R\big(min(x - 3{,}0)\big)^2$$

Minimize $\quad \mathrm{F}(x,R) = (x^3 - 10x - 2x^2 + 10) + R\big(min(x-3,0)\big)^2$

R 0



R’yi değiştirerek infeasible bölgeden kaçış mümkün olur.



Minimize

$$f(x) = x^3 - 10x - 2x^2 + 10$$

Subject to $\quad g(x) = x \geq 3$

Or, $g(x) = x - 3 \geq 0$

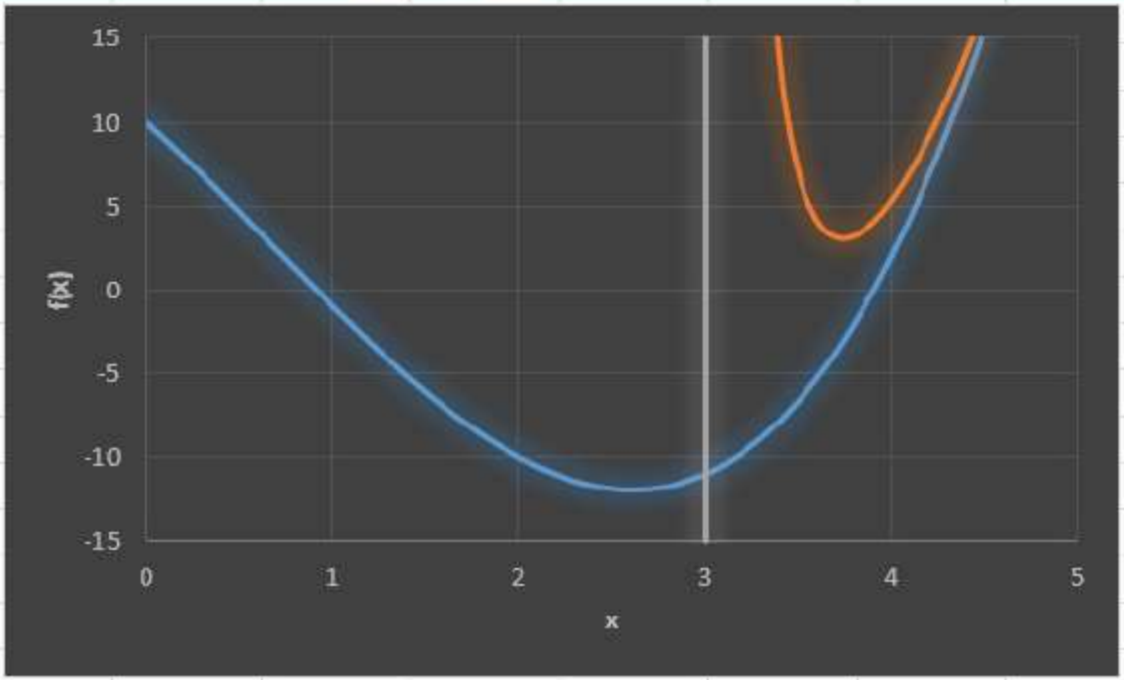
$$F(x, R) = (x^3 - 10x - 2x^2 + 10) + \boxed{R\frac{1}{g(x)}}$$

Bu terim sadece feasible bölgede eklenir

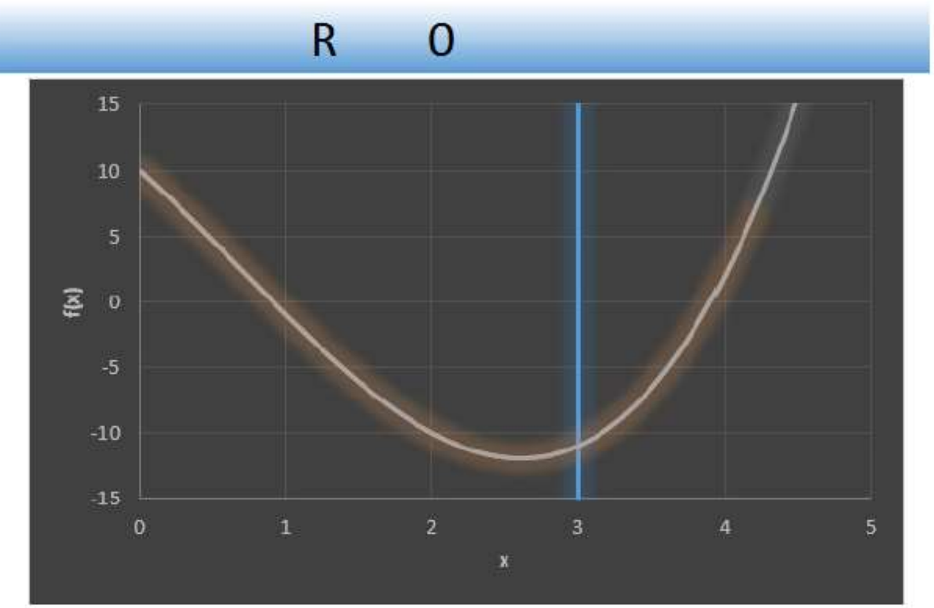
$$F(x, R) = (x^3 - 10x - 2x^2 + 10) + R\frac{1}{(x-3)}$$

Minimize $\mathrm{F}(x, R) = (x^3 - 10x - 2x^2 + 10) + R\frac{1}{(x-3)}$

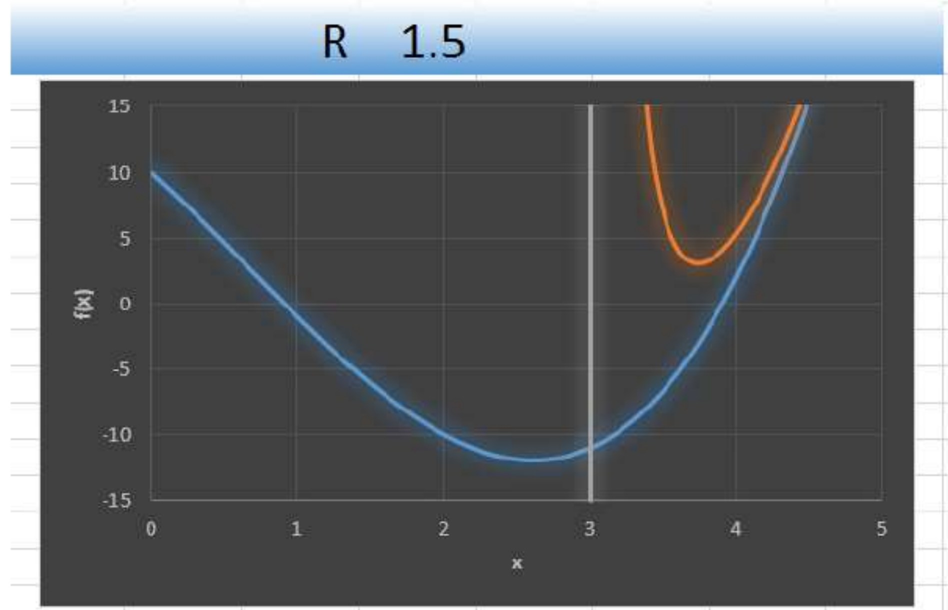
R 1.5



Exterior penalty method



Interior penalty method



Dönüşüm şu şekilde yazılır

$$\mathrm{F}(X,R) = f(X) + \boxed{\Psi\big(g(X), h(X)\big)}$$

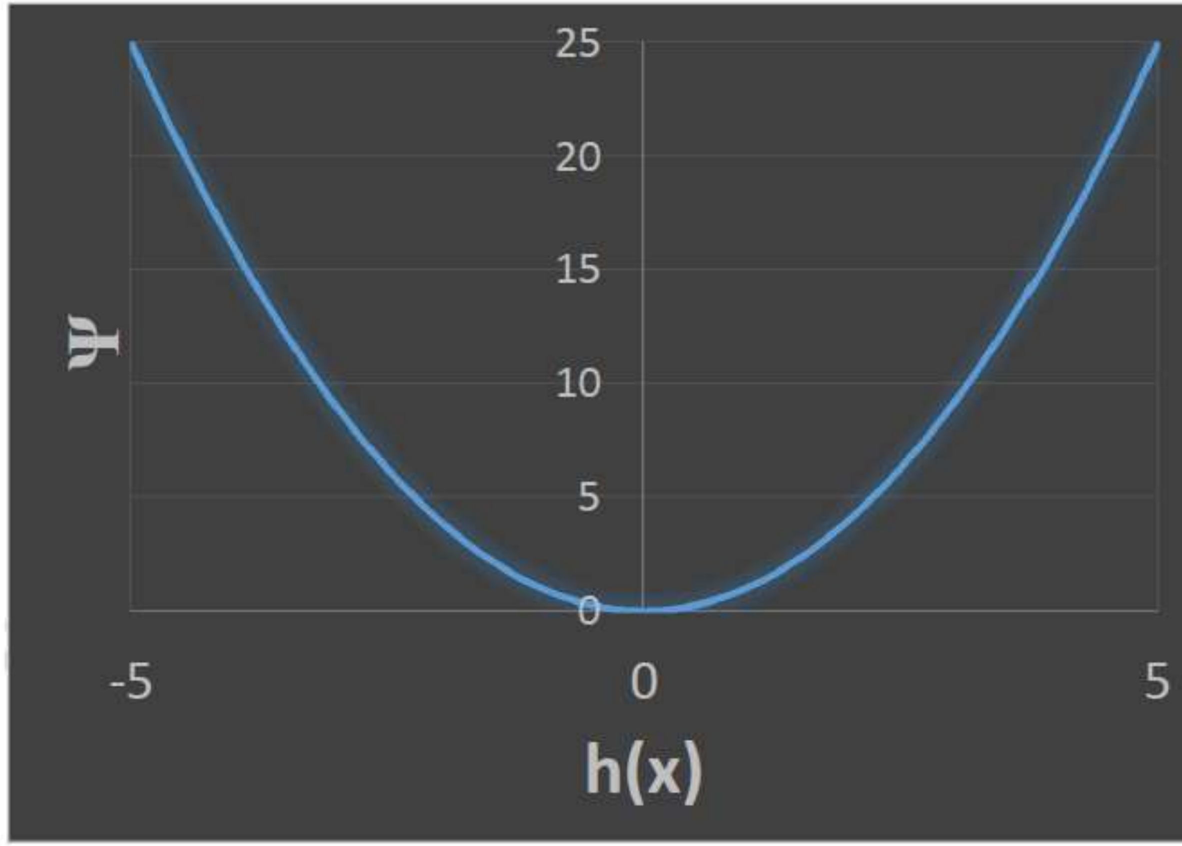
Penaltı terimi

R penaltı
parametreleridir

Penalty terms

Parabolic penalty

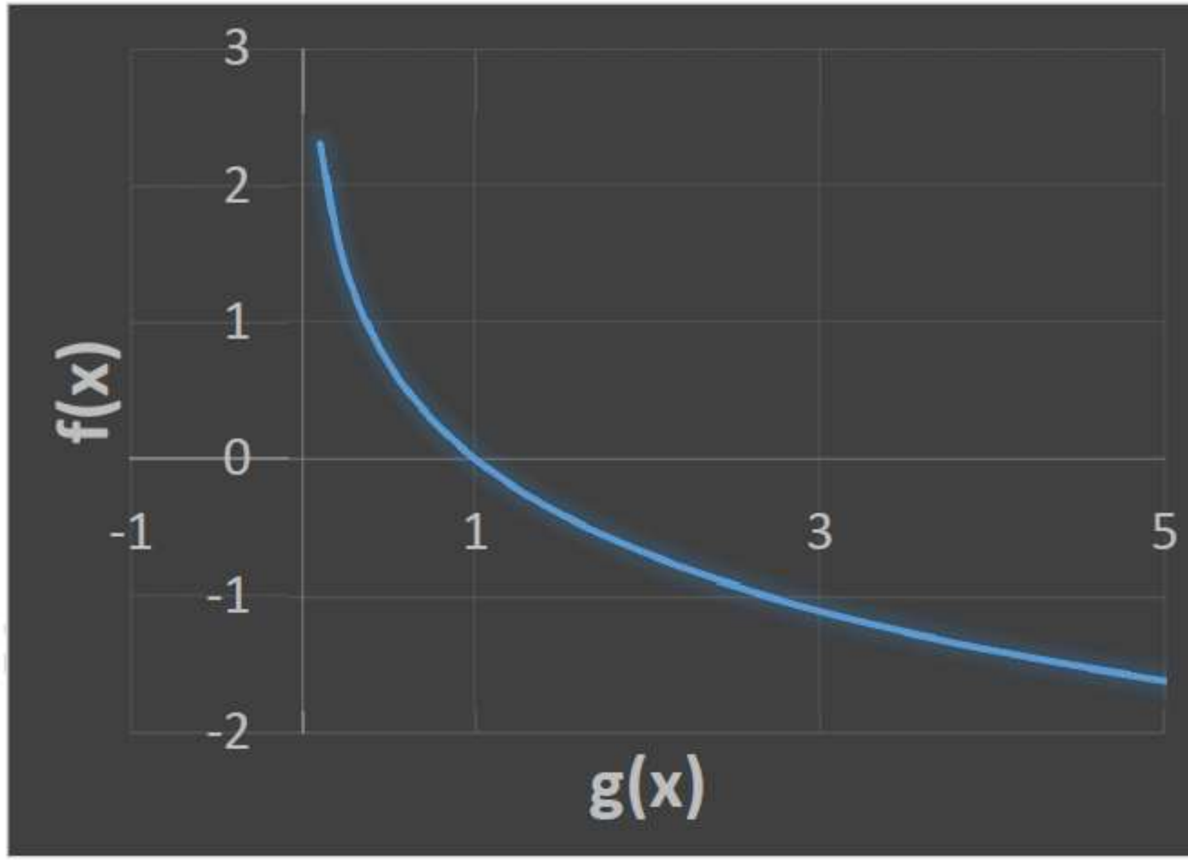
$$\Psi = R[h(x)]$$



Penalty terms

Log penalty

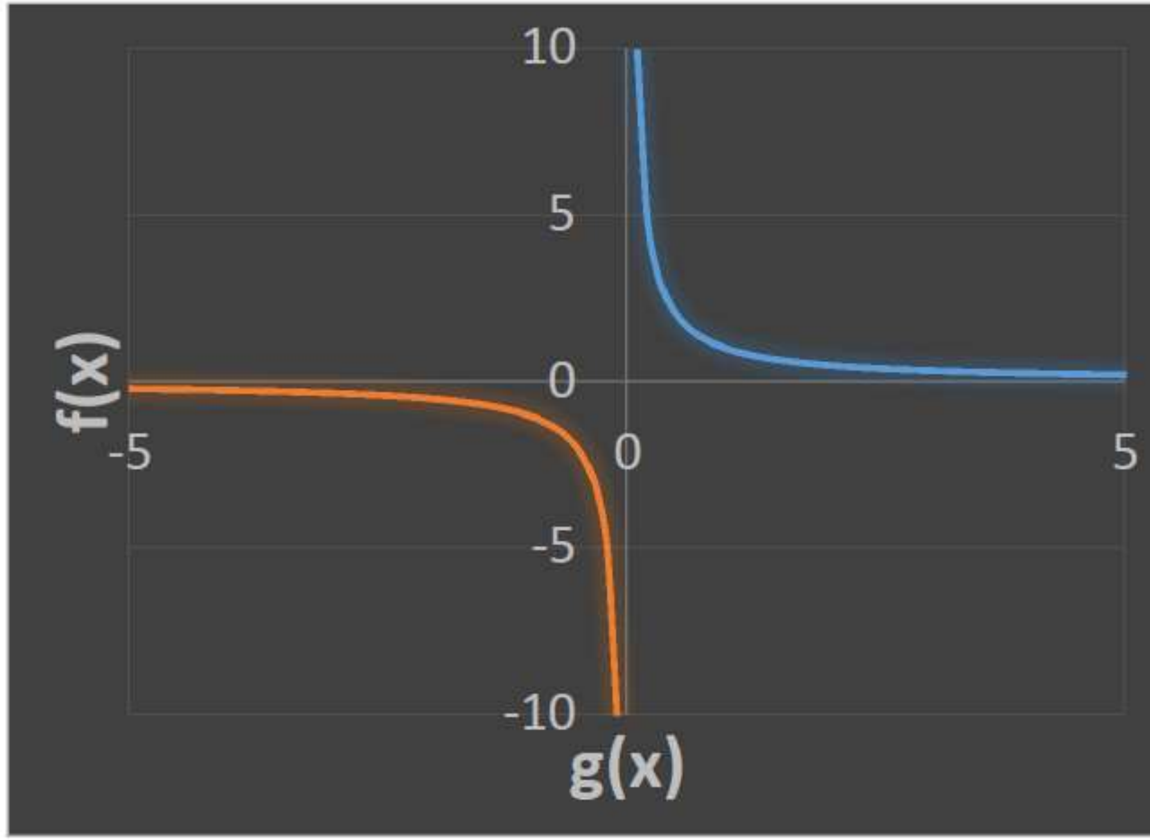
$$\Psi = -R\ln[g(x)]$$



Penalty terms

Inverse penalty

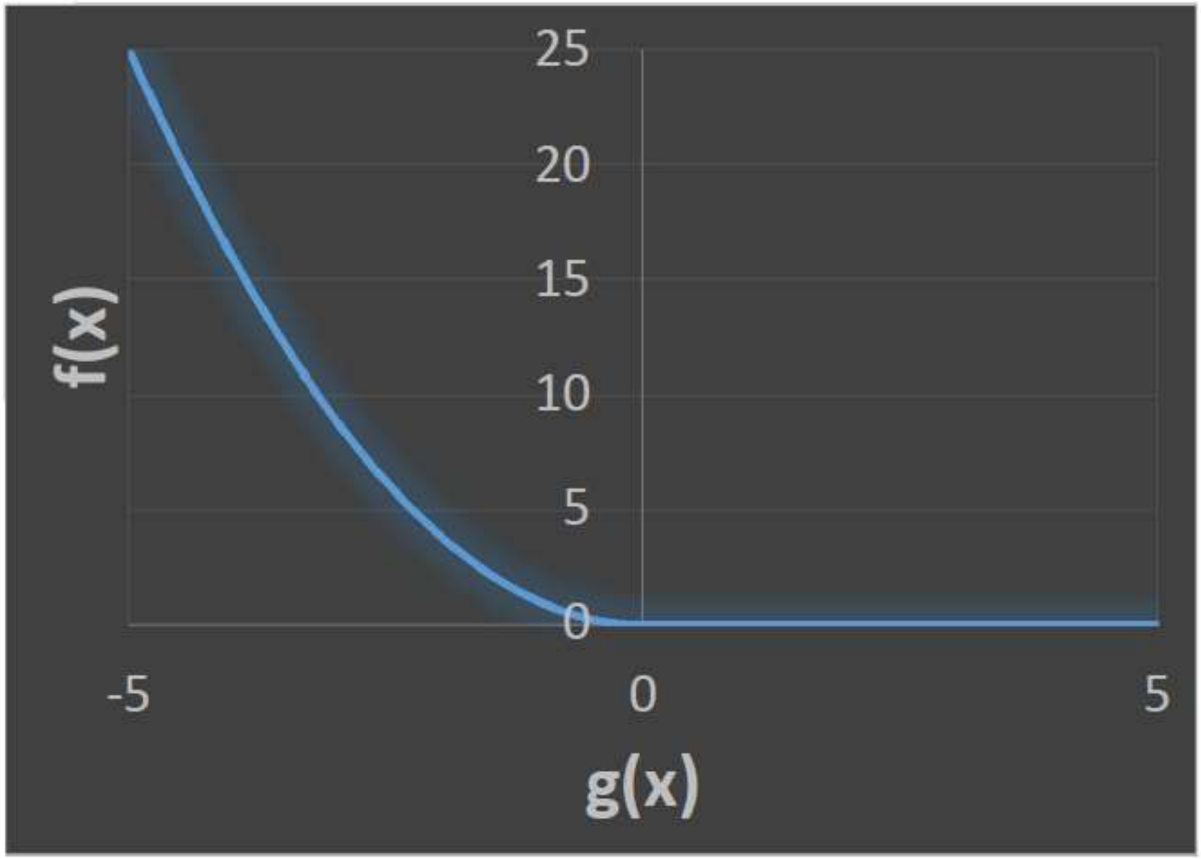
$$\Psi = -R\left[\frac{1}{g(x)}\right]$$



Penalty terms

Bracket operator

$$\Psi = -R\langle g(x)\rangle$$



EXTERIOR PENALTY METODU

- Penalty Fonksiyonu

$$P(x) = \sum_{k=1}^{n_e} [h_k(x)]^2 + \sum_{j=1}^{n_g} \{max[0, g_j(x)]\}^2$$

- Bütün kisitlar saglanirsa, ***P(x) = 0***.

- Esdeger optimizasyon probleminin feasible bir noktadan baslayan çözümü hemen bir infeasible noktaya yakinsar (gerçek kisitlar direk kullanilmadigi için).

- Penalty katsayisi artirilirsa, infeasible bir noktadan baslayan esdeger problemin çözümü gittikçe gerçek çözüme yaklasir.

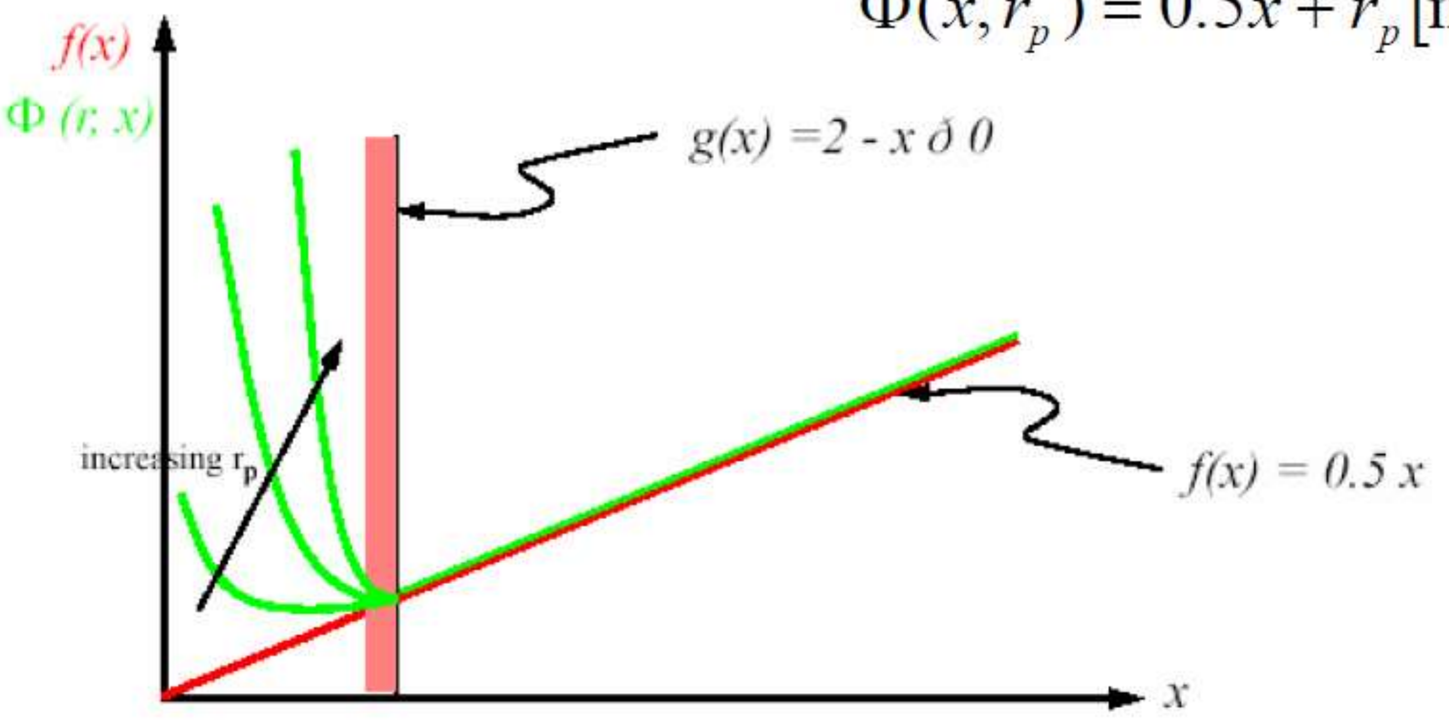
- Penalty katsayisi çok artirilirsa, esdeger problemin çözümü *"ill-conditioning"* durumlarinin olusmasi nedeniyle zorlasir.

EXTERIOR PENALTY METODU

Örnek 1: $f(x)= 0.5x$
$g(x)=x-2<=0$

optimizasyon probleminin exterior penalty metodu ile çözümü:

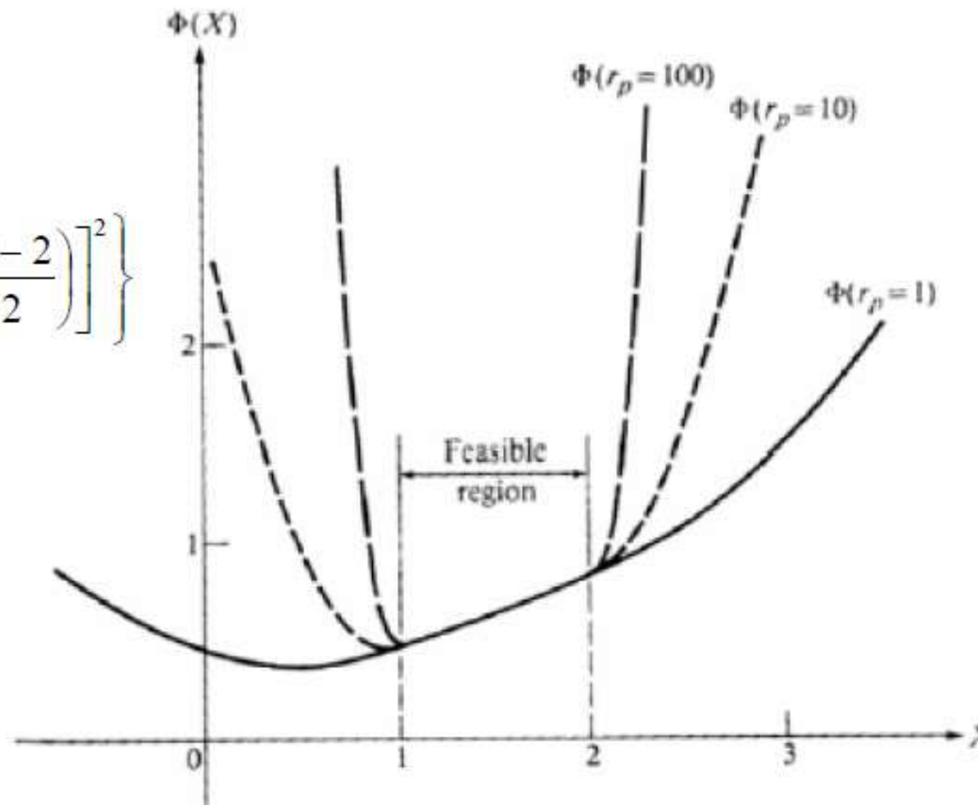
$$\Phi(x, r_p) = 0.5x + r_p\left[\max(0, x-2)\right]^2$$



EXTERIOR PENALTY METODU

Örnek 2: f(x)= (1/20)*(x+2)^2
g1(x)=(1-x)/2<=0
g2(x)=(x-2)/2<=0
optimizasyon probleminin exterior penalty metodu ile çözümü:

$$\Phi(x, r_p) = \frac{(x+2)^2}{20} + r_p\left\{\left[\max\left(0, \frac{1-x}{2}\right)\right]^2 + \left[\max\left(0, \frac{x-2}{2}\right)\right]^2\right\}$$



Örnek 3

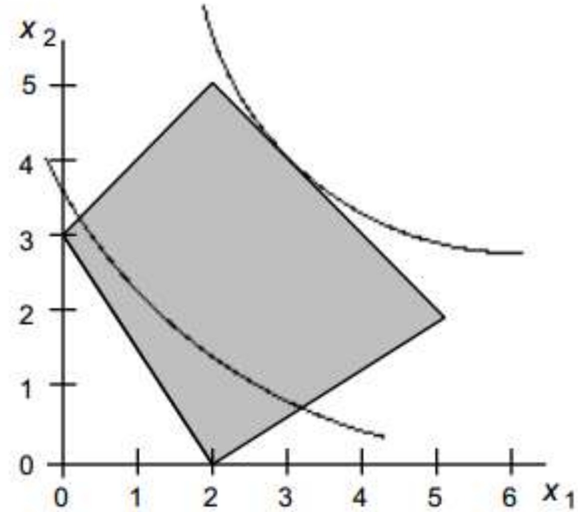
Minimize $f(\mathbf{x}) = (x_1 - 6)^2 + (x_2 - 7)^2$

subject to $g_1(\mathbf{x}) = -3x_1 - 2x_2 + 6 \le 0$

$g_2(\mathbf{x}) = -x_1 + x_2 - 3 \le 0$

$g_3(\mathbf{x}) = x_1 + x_2 - 7 \le 0$

$g_4(\mathbf{x}) = \frac{2}{3}x_1 - x_2 - \frac{4}{3} \le 0$



$$\theta(c,\mathbf{x}) = (x_1 - 6)^2 + (x_2 - 7)^2 + c\big((\max\{0, -3x_1 - 2x_2 + 6\})^2 + (\max\{0, -x_1 + x_2 - 3\})^2 + (\max\{0, x_1 + x_2 - 7\})^2 + (\max\{0, \tfrac{2}{3}x_1 - x_2 - \tfrac{4}{3}\})^2\big).$$

İhlal

$$\theta(c,\mathbf{x}) = (x_1 - 6)^2 + (x_2 - 7)^2 + c(\max\{0, x_1 + x_2 - 7\})^2.$$

$$\nabla_{\mathbf{x}}\theta(c,\mathbf{x}) = \begin{pmatrix} 2x_1 - 12 + 2c(x_1 + x_2 - 7) \\ 2x_2 - 14 + 2c(x_1 + x_2 - 7) \end{pmatrix}.$$

Gradyan

$$x_1^*(c) = \frac{6(1+c)}{1+2c} \text{ and } x_2^*(c) = 7 - \frac{6c}{1+2c}$$

limit of $x_1^*(c)$ and $x_2^*(c)$ as $c \to \infty$,

$x_1^* = 3$ and $x_2^* = 4$,

EXTERIOR PENALTY METODU

<u>Degerlendirme</u>:

- Penalty uygulanmasi, kisit ihlalinden sonra baslar.

- Feasible bölge içindeki amaç fonksiyonu bundan etkilenmez.

- Penalty fonksiyonu amaç fonksiyonun tanimli oldugu her yerde tanimlidir. Optimumu bulmak için, feasible bir çözümden baslama sarti yoktur.

- Penalty fonksiyonun çözümü her zaman biraz feasible bölge disina düser. Amaç fonksiyonu, feasible domain disinda tanimli olmayabilir.

- Penalty katsayisini artirmak, çözümü gerçek problemin çözümüne yaklasatirirken nümerik çözümü de zorlastirir.

- Penalty metodu ile, esitlik veya esitsizlik seklindeki problemler çözülebilir.

INTERIOR PENALTY METODU

Penalty Fonksiyonu:

- Farkli türleri vardir fakat,
- fonksiyonu yaygin kullanilir.

$$\boxed{P(x) = \sum_{j=1}^{n_g} \frac{-1}{g_j(x)}}$$

penalty

- Esdeger amaç fonksiyonu sadece feasible bölge içinde tanimlidir. Interior penalty metodu ile çözüm için, çözümün feasible bir noktadan baslatilmasi gerekir.

- Penalty katsayisinin gittikçe küçülen degerlerinin kullanilmasiyla gerçek çözüme yaklasilir.

- Esdeger amaç fonksiyonu kisit sinirlarinda tanimsizdir. Feasible bir çözüm elde etmek için 1-D arastirmalarinda (1-D search) dikkatli olmalidir.

- Esitlik seklindeki kisitlar, Exterior penalty metodu ile benzer sekilde dikkate alinir.

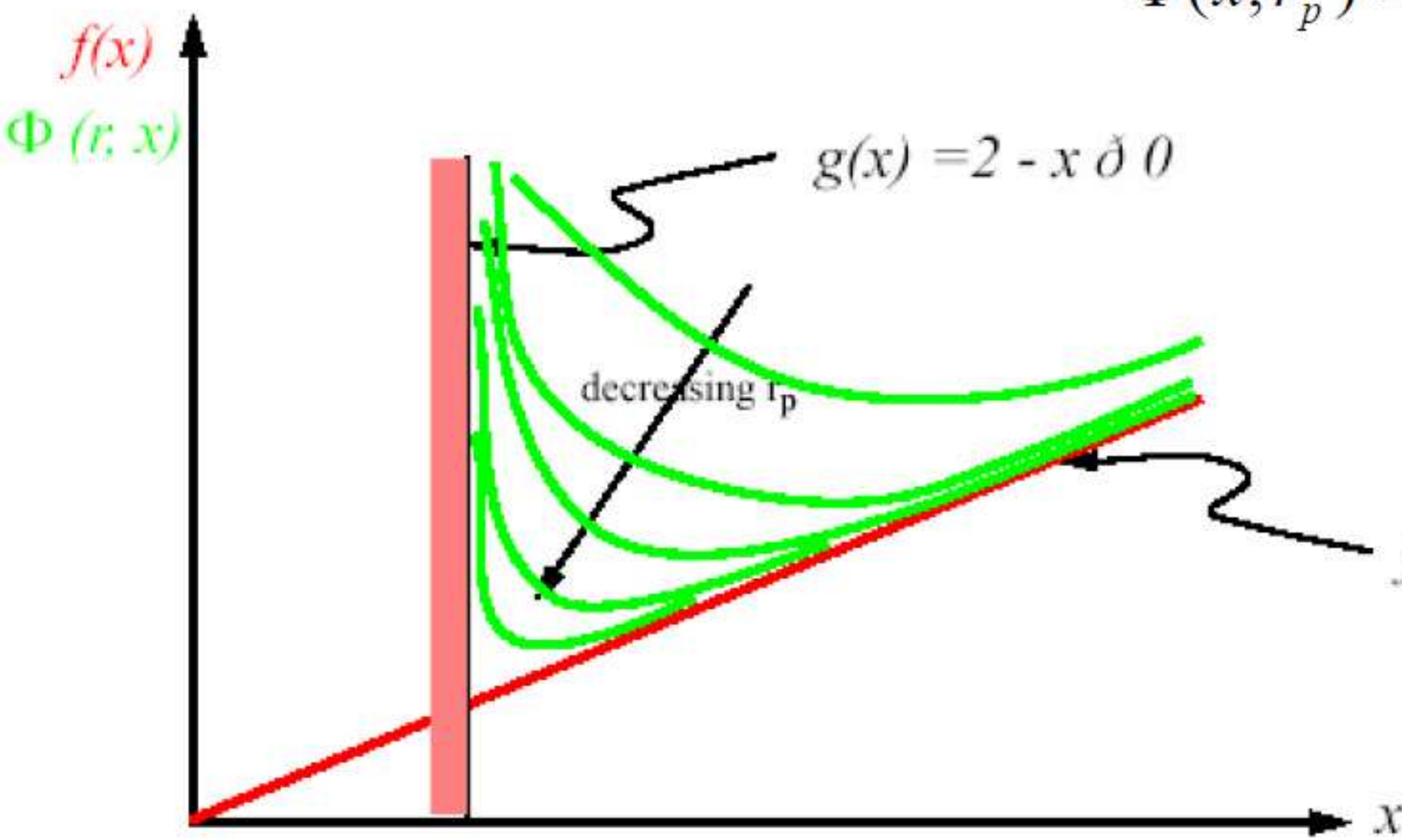
INTERIOR PENALTY METODU

Örnek 1: $f(x)= 0.5x$

$g(x)=x-2<=0$

optimizasyon probleminin interior penalty metodu ile çözümü:

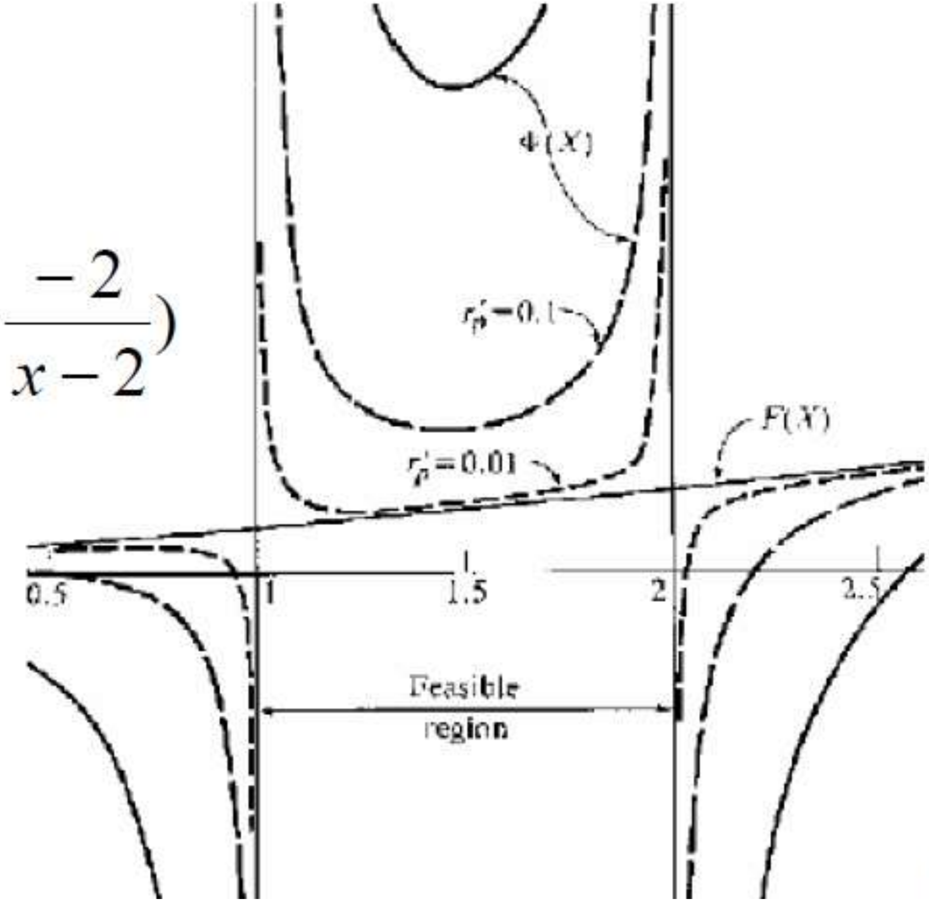
$$\Phi(x, r_p^{'}) = 0.5x + r_p^{'}\left(\frac{-1}{x-2}\right)$$



INTERIOR PENALTY METODU

<u>Örnek 2</u>: f(x)= (1/20)*(x+2)^2
g1(x)=(1-x)/2<=0
g2(x)=(x-2)/2<=0
optimizasyon probleminin interior penalty metodu ile çözümü:

$$\Phi(x,r_p^{'}) = \frac{(x+2)^2}{20} + r_p^{'}\left(\frac{-2}{1-x} + \frac{-2}{x-2}\right)$$



Örnek 3

Minimize $\{f(\mathbf{x}) = 2x_1^2 + 9x_2 \text{ subject to } x_1 + x_2 \geq 4\}$.

$$B(r,\mathbf{x}) = 2x_1^2 + 9x_2 + r\left(\frac{-1}{-x_1 - x_2 + 4}\right).$$

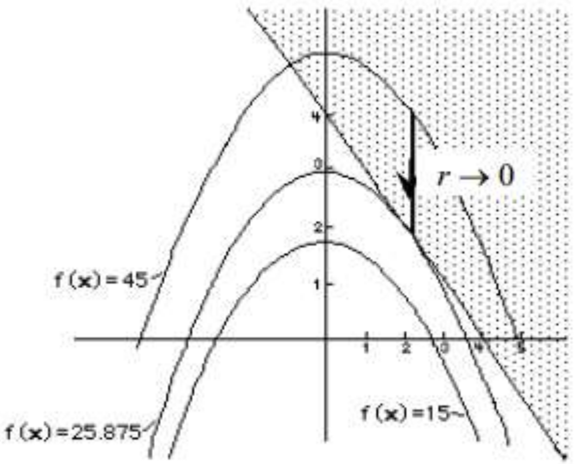
$$\nabla_{\mathbf{x}} B(r,\mathbf{x}) = \begin{pmatrix} 4x_1 - r(-x_1 - x_2 + 4)^{-2} \\ 9 - r(-x_1 - x_2 + 4)^{-2} \end{pmatrix}.$$

Gradyan

$x_1(r) = 2.25$ and $x_2(r) = 0.333\sqrt{r} + 1.75$ for all $r > 0$.

Limit

$x_1 = 2.25$ and $x_2 = 1.75$



Örnek 4

Min $2x_1^2 + 3x_2^2$

Sınırlayıcı: $x_1 + 2x_2 = 5$

Penaltı fonksiyonu:

$P(x_1,x_2,r) = 2x_1^2 + 3x_2^2 + 1/r\,[x_1 + 2x_2 - 5]^2$

x_1 ve x_2 ye göre birinci kısmi türevler sıfıra eşitlenirse:

$$\frac{\partial P}{\partial x_1} = 4x_1 + \frac{2}{r}[x_1 + 2x_2 - 5] = 0$$

$$\frac{\partial P}{\partial x_2} = 6x_2 + \frac{4}{r}[x_1 + 2x_2 - 5] = 0$$

x_1 ve x_2 için çözülürse

$$x_1 = \frac{15}{11+6r} \qquad x_2 = \frac{20}{11+6r}$$

r'ye sıfır ver

$x_1 = 15/11$ $\qquad$ $x_2 = 20/11$

INTERIOR PENALTY METODU
(Degerlendirme)

Avantaj:

- Her iterasyondaki çözüm feasible bölge içindedir. Her hangi bir nedenle çözüm iterasyonlari durdurulursa, eldeki çözüm *geçerli* (feasible) bir çözümdür.

Dezavantaj:

- Çözüme feasible bir noktadan baslanmak zorundadir.

- Penalty fonksiyonu kisit sinirinda sonsuzdur kisit disinda ise tanimsizdir. Herhangi bir nedenle çözüm feasible bölge disina tasarsa, geri feasible bölgeye dönemez.

- Penalty etkisinii artirmak çözümü gerçege yaklastirir, fakat ayni zamanda zorlastirir.

- Sadece esitsizlik tipindeki problemler çözülebilir.

EXTENDED-INTERIOR PENALTY METODU

- Penalty fonksiyonu kisit ihlalinin mertebesine göre degisiklik gösterir (g0 esik deger olarak kullanilir). Kuadratik penalty fonksiyonu.

$$P_j(x) = \frac{-1}{g_j(x)} \qquad IF \quad g_j(x) \le g_o$$

$$P_j(x) = \frac{-1}{g_o}\left\{\left[\frac{g_j(x)}{g_o}\right]^2 - 3\left[\frac{g_j(x)}{g_o}\right] + 3\right\} \qquad IF \quad g_j(x) > g_o$$

- Kisit sinirlarinda tanimsizlik durumu yoktur.
- Feasible veya infeasible baslangiç noktasi ile çözüme baslanabilir.
- Metod sürekli feasible bölgede çalisir. O yüzden elde edilen sonuçlar hep feasible olur.

EXTENDED-INTERIOR PENALTY METODU

- Interior metottan Exterior metoda geçis go degeri ile kontrol edilir.

$$g_o = -C(r_p)^a \qquad Where \qquad \frac{1}{3} \le a \le \frac{1}{2} \qquad and \qquad C = 0.2$$

- Diger Extended-Interior Penalty Fonksiyonlari
  - ✓ Lineer Extended Penalty Fonksiyonu

$$P_j(x) = \frac{-1}{g_j(x)} \qquad IF \quad g_j(x) \le g_o$$

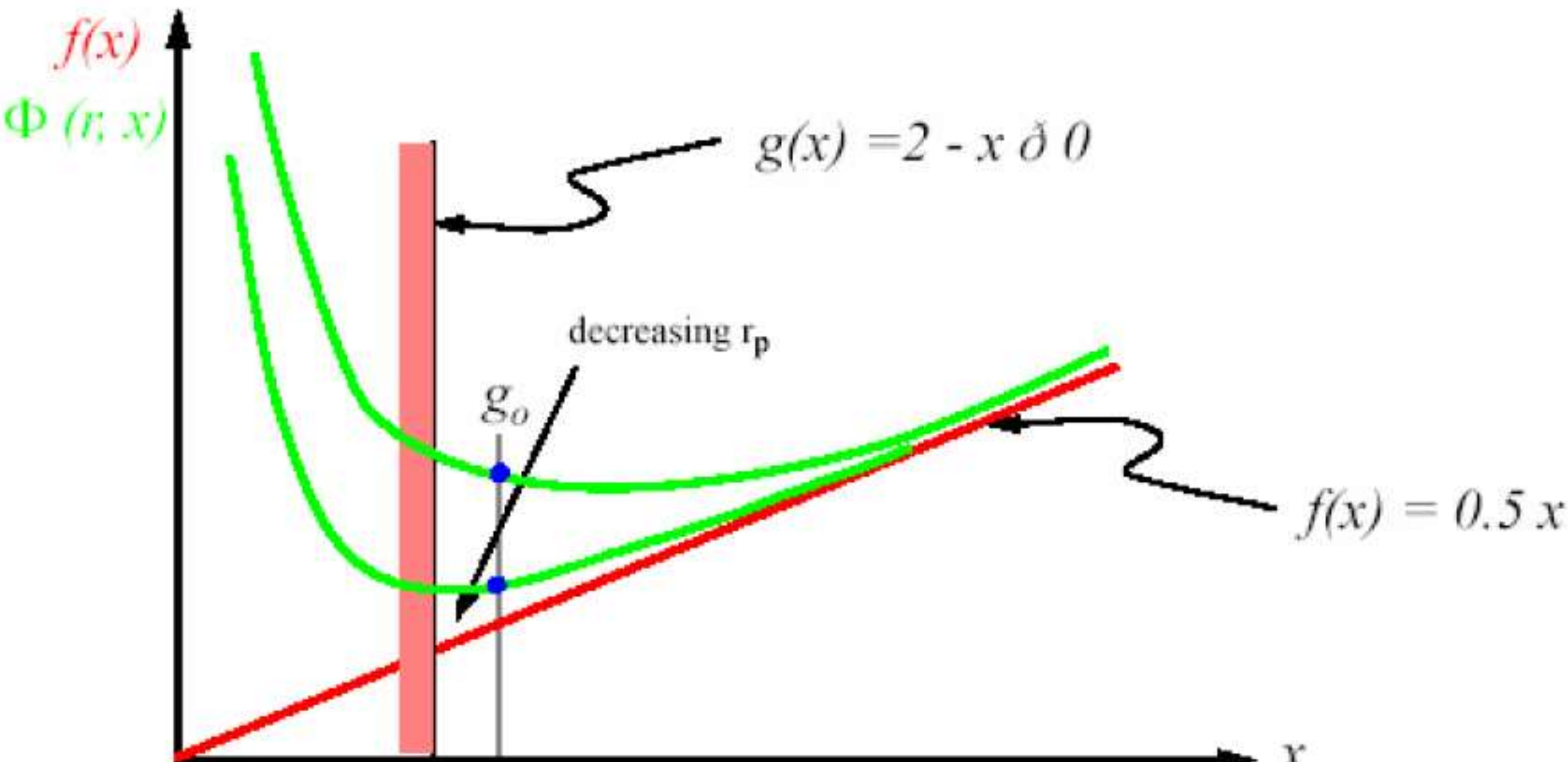
$$P_j(x) = -\frac{2g_o - g_j(x)}{g_o^2} \qquad IF \quad g_j(x) > g_o$$

  - ✓ Degisken penalty fonksiyonu

EXTENDED-INTERIOR PENALTY METODU

<u>Örnek</u>: f(x)= 0.5x
g(x)=x-2<=0

optimizasyon probleminin extended-interior metodu ile çözümü çözümü:



AUGMENTED LAGRANGIAN METODU

- Lagrangian Metodu (Sadece esitlik türündeki kisitlarla minimizasyon)

  - ✓ *Minimize*

$$f(x) = f(x_1, x_2, \ldots\ldots, x_n)$$

  - ✓ *kisit*

$$h_k(x) = 0 \qquad k = 1, \ldots\ldots, n_e$$

- Kuhn-Tucker Gerek Sartlari

  - ✓ *Dönüm noktasi*

$$L(x, \lambda) = f(x) + \sum_{k=1}^{n_e} \lambda_k h_k(x)$$

  - ✓ *kisit*

$$h_k(x) = 0 \qquad k = 1, \ldots\ldots, n_e$$

AUGMENTED LAGRANGIAN METODU

- Augmented Lagrangian (Esitlik türündeki kisitlar durumu)

$$\Phi(x, \lambda, r_p) = f(x) + \sum_{k=1}^{n_e} \{\lambda_k h_k(x) + r_p[h_k(x)]^2\}$$

✓ Optimal *Lamda_k* degeri bilinmez.
✓ Eger *Lamda_k* =0 ise metod Exterior Penalty metoduna dönüsür.
✓ Eger optimal *Lamda_k* degeri bilinirse, herhangi bir penalty katsayisi degeri kullanarak yapilan optimizasyon hemen gerçek sonucu verir.
✓ Optimal *Lamda_k* degeri (tahmini bir baslangiç degeri kullanarak) iteratif olarak hesaplanir.

$$\lambda_k^{p+1} = \lambda_k^p + 2r_p h_k(x^p) \qquad k = 1, \ldots\ldots, n_e$$

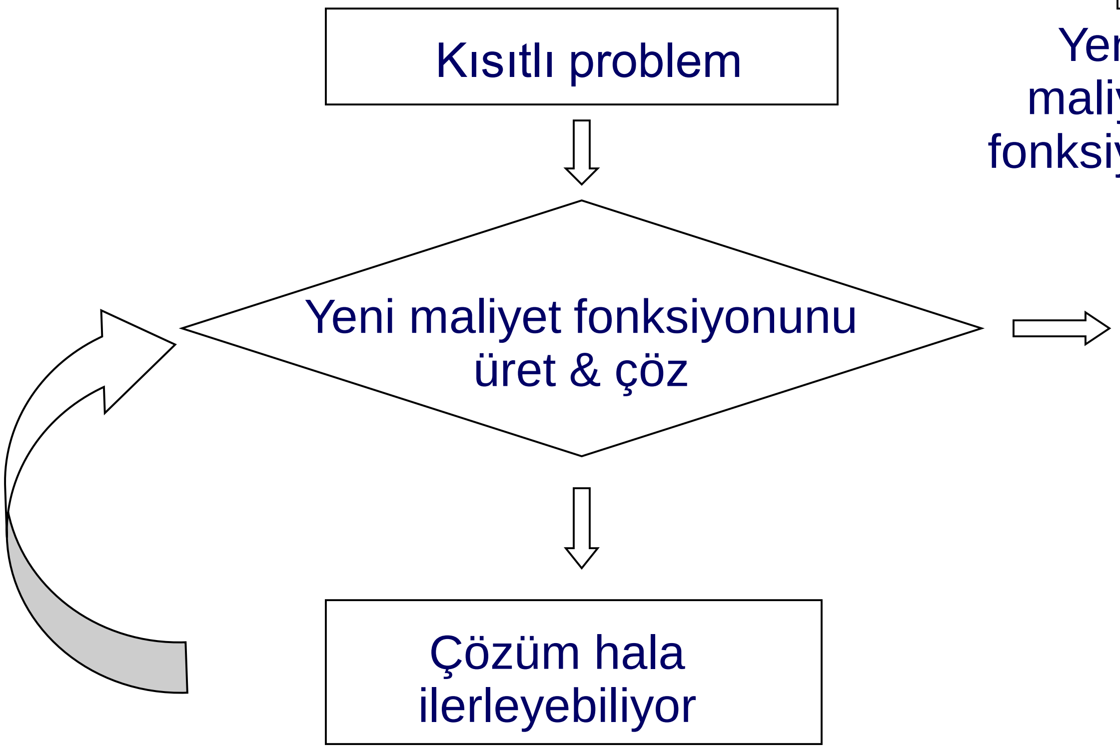
Augmented Lagrange Metodu

$$\ell(x,y) = f\left(x,y\right) + \lambda g\left(x,y\right)$$

Yeni maliyet fonksiyonu

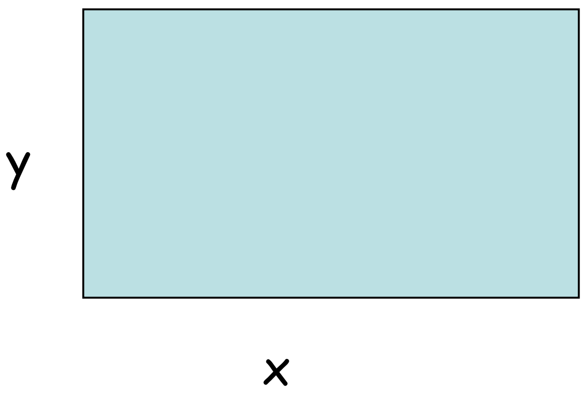
Maliyet fonksiyonu

Penaltı



Basit bir örnek

Problem: 50 cm’lik çitim var ve bunla maksimum alanı kaplamak istiyorum



f(x,y) = xy Amaç

g(x,y) = 2x + 2y = 50 Sınırlayıcı

Bul

$$f_x(x,y) = y;\ \ f_y(x,y) = x$$

$$g_x(x,y) = 2;\ g_y(x,y) = 2$$

$$\begin{cases} Set\ \ f_x = \lambda g_x \rightarrow y = 2\lambda \\ Set\ \ f_y = \lambda g_y \rightarrow x = 2\lambda \\ \qquad 2x + 2y = 50 \end{cases}$$

Eş zamanlı çöz

$$4\lambda + 4\lambda = 50$$

$$8\lambda = 50$$

$$\lambda = 6.25$$

$$x = 12.5$$

$$y = 12.5$$

$$\max_{x,y} \; f(x, y)$$

$$subject\ to \quad g(\mathrm{x}, \mathrm{y}) \geq 0$$

Lagrange

$$\ell(x, y) = f(x, y) + \lambda g(x, y)$$

Maksimize edilecek fonksiyon

Çarpan

Sınırlayıcı(lar)

Lagrange düzenlendikten sonra, türev alınıp sıfıra eşitlenir

$$\ell(x,y) = f(x,y) + \lambda g(x,y)$$

__**Birinci derece yeter şartlar**__

$$\ell_x(x,y) = f_x(x,y) + \lambda g_x(x,y) = 0$$
$$\ell_y(x,y) = f_y(x,y) + \lambda g_y(x,y) = 0$$

Şimdi, "Çarpan" şartlarımız var…

$$\lambda \geq 0 \qquad g(x,y) \geq 0 \qquad \lambda g(x,y) = 0$$

Lagrange yöntemi

- f(x1,x2,……,xn) fonksiyonunun,
- g1(x1,x2,……,xn) = b1
- g2(x1,x2,……,xn) =b2
- …
- gm(x1,x2,……,xn) =bm

biçimindeki kısıtlayıcıların sınırlayıcı koşulları altında en büyük(veya en küçük) değerine ulaştığı noktanın belirlenmek istendiğini varsayalim. Bu sorunun çözülmesi icin Lagrange yöntemi denilen bir matematiksel yönteme başvurulur.

Bu genel gösterimde sunulan problemin en iyi çözümünün hesaplanmasında kısıtlayıcı fonksiyonların her biri için bir tane olmak üzere kısıtlayıcı şayisi kadar yani m tane yeni değişken tanımlanması ile,

$L(x_1,x_2,\ldots\ldots x_n,\lambda_1,\lambda_2,\ldots\ldots\lambda_m)= f(x_1,x_2,\ldots\ldots x_n) + \sum_{i=1}^{m} \lambda_i[\, b_i- g_i\,(x_1,x_2,\ldots\ldots x_n)]$

olarak ifade edilen yeni bir fonksiyon oluşturulur. Bu fonksiyona Lagrange Fonksiyonu , λ_i'ye Lagrange Çarpanı denir.

Lagrange fonksiyonuna dayanan Lagrange yönteminin esasi,Lagrange fonksiyonu için en büyük veya en küçük gösteren $(x_1^*,x_2^*,\ldots\ldots x_n^*,\lambda_1^*,\lambda_2^*,\ldots\ldots\lambda_m^*)$ noktasının belirlenmesidir. Lagrange fonksiyonu için en iyi çözüm olma özelliğine sahip bu nokta ayni zamanda,esas fonksiyonun da en büyük değerinin ortaya çıktığı $(x_1^*,x_2^*,\ldots\ldots x_n^*,\lambda_1^*,\lambda_2^*,\ldots\ldots\lambda_m^*)$ noktasında,

$\frac{\partial L}{\partial \lambda i} = b_i - g_i(x_1,x_2,\ldots\ldots,x_n) = 0$ olur.

$\frac{\partial L}{\partial \lambda i}$ =$b_i - g_i(x_1,x_2,\ldots\ldots,x_n)$ =0

bağıntısı $(x_1^*,x_2^*,\ldots\ldots.x_n^*)$'in açıklanan problemin uygun çözümü olduğunu göstermektedir. Bu noktanın yalnızca uygun değil ayni zamanda, en iyi çözüm olduğu da gösterilebilir. Bunun için problemin uygun çözüm bölgesinden $(x_1^{'},x_2^{'},\ldots\ldots\ldots\ldots.x_n^{'})$ gibi bir nokta seçelim. $(x_1^*,x_2^*,\ldots\ldots\ldots\ldots.x_n^*,\lambda_1^*,\lambda_2^*,\ldots\ldots\ldots.\lambda_m^*)$ Lagrange fonksiyonunu en büyüklediğinden,herhangi sayılar kümesi $(\lambda_1^{'},\lambda_2^{'},\ldots\ldots\ldots.\lambda_m^{'})$ için,

$$L(x_1^*,x_2^*,\ldots\ldots\ldots\ldots.x_n^*,\lambda_1^*,\lambda_2^*,\ldots\ldots\ldots.\lambda_m^*)\geq L(x_1^{'},x_2^{'},\ldots\ldots\ldots\ldots.x_n^{'},\lambda_1^{'},\lambda_2^{'},\ldots\ldots\ldots.\lambda_m^{'}) \quad (*)$$

yazılabilir. $(x_1^*,x_2^*,\ldots\ldots\ldots\ldots.x_n^*)$ ve $(x_1^{'},x_2^{'},\ldots\ldots\ldots\ldots.x_n^{'})$ uygun olduklarından Lagrange fonksiyonundaki bütün λ'li terimler sıfıra eşit olur. Bu durumda (*), $f(x_1^*,x_2^*,\ldots\ldots\ldots\ldots.x_n^*)\geq f(x_1^{'},x_2^{'},\ldots\ldots\ldots\ldots.x_n^{'})$ sekline dönüşür. Böylece $(x_1^*,x_2^*,\ldots\ldots\ldots\ldots.x_n^*)$'in problemi çözdüğü belirlenmiş olur.Buna göre ,$(x_1^*,x_2^*,\ldots\ldots\ldots\ldots.x_n^*,\lambda_1^*,\lambda_2^*,\ldots\ldots\ldots.\lambda_m^*)$ aşağıdaki kısıtlayıcısı en büyükleme probleminin çözümü ise, $(x_1^*,x_2^*,\ldots\ldots\ldots\ldots.x_n^*)$ kısıtlayıcılı problemin de çözümü olur.

$$\text{enb}L(x_1^*,x_2^*,\ldots\ldots\ldots\ldots.x_n^*,\lambda_1^*,\lambda_2^*,\ldots\ldots\ldots.\lambda_m^*)$$

Daha önce açıklandığı gibi $(x_1^*,x_2^*,\ldots\ldots\ldots\ldots x_n^*,\lambda_1^*,\lambda_2^*,\ldots\ldots\ldots\lambda_m^*)$ noktasının

enb$L(x_1^*,x_2^*,\ldots\ldots\ldots\ldots x_n^*,\lambda_1^*,\lambda_2^*,\ldots\ldots\ldots\lambda_m^*)$ problemini çözebilmesi için $(x_1^*,x_2^*,\ldots\ldots\ldots\ldots x_n^*,\lambda_1^*,\lambda_2^*,\ldots\ldots\ldots\lambda_m^*)$ noktasında,

$$\frac{\partial L}{\partial x1} = \frac{\partial L}{\partial x2} = \cdots = \frac{\partial L}{\partial xn} = \frac{\partial L}{\partial \lambda 1} = \frac{\partial L}{\partial \lambda 2} = \cdots = \frac{\partial L}{\partial \lambda n} = 0$$

olmalıdır.

Bu eşitliği sağlayan herhangi bir $(x_1^*,x_2^*,\ldots\ldots\ldots\ldots x_n^*,\lambda_1^*,\lambda_2^*,\ldots\ldots\ldots\lambda_m^*)$ noktasının sınırlı optimizasyon probleminin en iyi çözümü olabilmesinin koşulları bir sonraki slaytlardaki teoremlerle açıklanmaktadır.

Teorem-3: $f(x_1,x_2,\ldots\ldots,x_n)$ en büyüklenmek istenen bir konkav fonksiyon olsun. $i=1,2,\ldots..m$ için $g_i=(x_1,x_2,\ldots\ldots,x_n)$ doğrusal ise

$$\frac{\partial L}{\partial x1}=\frac{\partial L}{\partial x2}=\cdots=\frac{\partial L}{\partial xn}=\frac{\partial L}{\partial \lambda 1}=\frac{\partial L}{\partial \lambda 2}=\cdots=\frac{\partial L}{\partial \lambda n}=0$$

sağlayan herhangi bir $(x_1^*,x_2^*,\ldots\ldots\ldots\ldots\ldots x_n^*,\lambda_1^*,\lambda_2^*,\ldots\ldots\ldots.\lambda_m^*)$ noktası problemin en iyi çözümü $(x_1^*,x_2^*,\ldots\ldots\ldots\ldots\ldots x_n^*)$'I verir.

Teorem-3´: $f(x_1,x_2,\ldots\ldots,x_n)$ en küçüklenmek istenen bir konveks fonksiyon olsun. $i=1,2,\ldots..m$ icin $g_i=(x_1,x_2,\ldots\ldots,x_n)$ doğrusal ise

$$\frac{\partial L}{\partial x1}=\frac{\partial L}{\partial x2}=\cdots=\frac{\partial L}{\partial xn}=\frac{\partial L}{\partial \lambda 1}=\frac{\partial L}{\partial \lambda 2}=\cdots=\frac{\partial L}{\partial \lambda n}=0$$

sağlayan herhangi bir $(x_1^*,x_2^*,\ldots\ldots\ldots\ldots\ldots x_n^*,\lambda_1^*,\lambda_2^*,\ldots\ldots\ldots.\lambda_m^*)$ noktası problemin en iyi çözümü $(x_1^*,x_2^*,\ldots\ldots\ldots\ldots\ldots x_n^*)$'I verir.

Lagrange yöntemiyle problem çözümüne örnek olması bakımından biri en büyükleme, diğeri en küçükleme amaçlı iki örnek problem çözelim.

Örnek: $f(x,y)=8x+16y-x^2-y^2$ fonksiyonunun $g(x,y)=2x+4y=6$ kısıtlayıcısı altında en büyükleyen uç noktayı bulunuz.

Önce Lagrange fonksiyonunu yazalım.

$L(x,y,\lambda)=8x+16y-x^2-y^2+\lambda(6-2x-4y)$

$=8x+16y-x^2-y^2+6\lambda-2\lambda x-4\lambda y$

Lagrange fonksiyonun x,y ve λ'ya göre kısmi türevlerini alalım ve her birini sıfıra eşitleyelim. Buna göre;

$\frac{\partial L}{\partial x}=8-2x-2\lambda=0$

$\frac{\partial L}{\partial y}=16-2y-4\lambda=0$

$\frac{\partial L}{\partial \lambda}=6-2x-4y=0$

olur. Bu üç eşitlikten x,y ve λ bilinmeyenleri hesaplanmalıdır. Bu hesaplama yapıldığında;

$\lambda=17\backslash 5$; $x=3\backslash 5$; $y=6\backslash 5$ bulunur.Simdi f(x,y)'nin konkav olup olmadığını inceleyelim.Bunun için fonksiyonun Hessian matrisini oluşturalım. Fonksiyonun Hessian matrisi şöyledir;

$H(x,y)=\begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$

H(x,y)'nin 1.asal minörü $-2<0$

H(x,y)'nin 2.asal minörü $4>0$ olduğundan f(x,y) konkavdır. g(x,y) doğrusal olduğundan en iyi çözüm

$x=0.6$; $y=1.2$; $Z_{enb}=15$ olarak belirlenmiş olur.

Örnek: $f(x,y,z)=4x^2+y^2+5z^2$ fonksiyonunu $g(x,y,z)=2x+3y+4z=12$ kısıtlayıcısı altında küçükleyen üç noktayı bulunuz.

$L(x,y,z,\lambda)= 4x^2+y^2+5z^2+\lambda(12-2x-3y-4z)$

$=4x^2+y^2+5z^2+12\lambda-2\lambda x-3\lambda y-4\lambda z$

Kısmi türevleri alalım;

$\frac{\partial L}{\partial x}=8x-2\lambda=0$

$\frac{\partial L}{\partial y}=2y-3\lambda=0$

$\frac{\partial L}{\partial z}=10z-4\lambda=0$

$\frac{\partial L}{\partial \lambda}=12-2x-3y-4z=0$

olur. Bu dört eşitlikten x,y,z ve λ hesaplandığında;

X=5\11 ; y=30\11 ; z=8\11 ; λ=20\11 olarak hesaplanacaktır.

Şimdi de fonksiyonun konveks olup olmadığını inceleyelim.f(x,y,z) fonksiyonunun Hessian matrisi aşağıdaki gibidir:

$H=\begin{bmatrix} 8 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 10 \end{bmatrix}$ buradan;

$H_1(x,y,z)=8>0$

$H_2(x,y,z)=16>0$

$H_3(x,y,z)=160>0$ tüm asal minörleri pozitif olduğundan f(x,y,z) konvekstir. g(x,y,z) doğrusal olduğundan en iyi çözüm (x,y,z)=(5\11,30\11,8\11) ve $Z_{enk}=10.91$ olarak belirlenmiş olur.

AUGMENTED LAGRANGIAN METODU

- Augmented Lagrangian (Esitsizlik türündeki kisitlar durumu)

$$\Phi(\mathbf{x}, \lambda, \mathbf{s}, r_p) = f(\mathbf{x}) + \sum_{k=1}^{n_g} \{\lambda_k[g_k(\mathbf{x}) + s_k^2] + r_p[g_k(\mathbf{x}) + s_k^2]^2\}$$

✓ Fakat bu formülasyonda ng adet (*sk*) ilave degisken kullanilir.

$$\Phi(\mathbf{x}, \lambda, r_p) = f(\mathbf{x}) + \sum_{k=1}^{n_g} \{\lambda_k \psi_k + r_p \psi_k^2\} \qquad \psi_k = Max\left[g_k(\mathbf{x}), \frac{-\lambda_k}{2r_p}\right]$$

✓ Lagrange çarpanlari iteratif olarak asagidaki gibi hesaplanir:

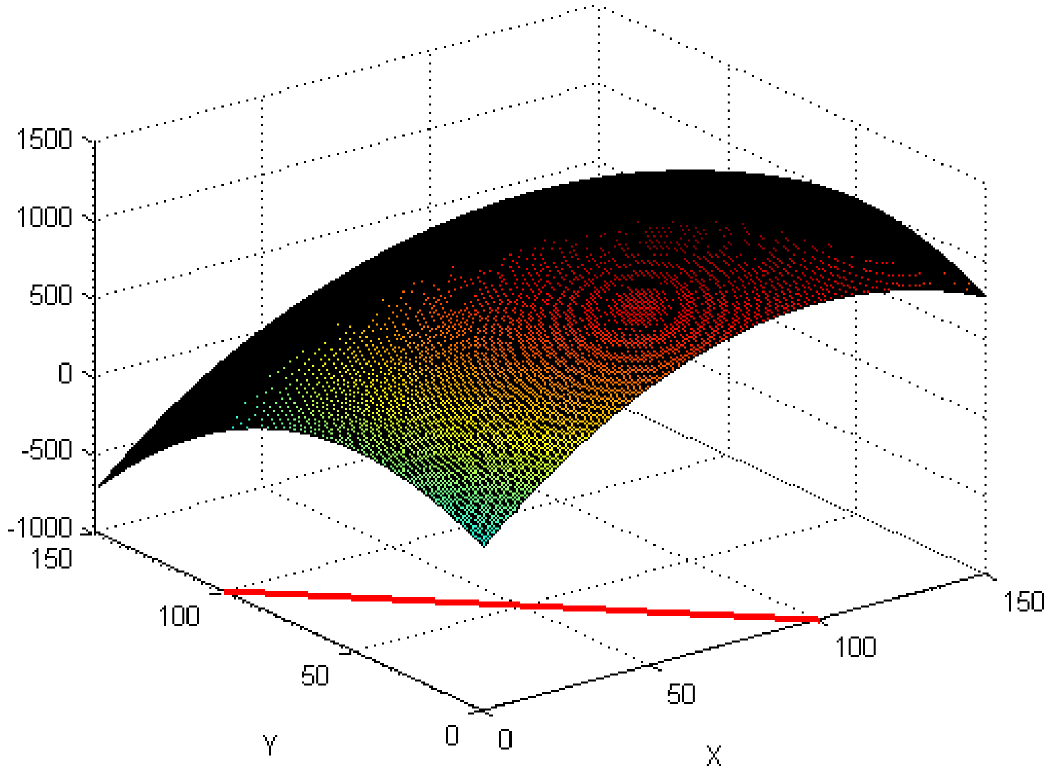
$$\lambda_k^{p+1} = \lambda_k^p + 2r_p\left(Max\left[g_k(\mathbf{x}), \frac{-\lambda_k}{2r_p}\right]\right) \qquad k = 1, \ldots\ldots, n_g$$

Örnek

$$\pi(x, y) = 10x + 20y - .1(x^2 + y^2)$$

$$x + y \leq 100$$

Kısıt



$$\max_{x>0,y>0}\left\{10x+20y-.1(x^2+y^2)\right\}$$

$$subject\ to \quad x+y\le 100$$

Çarpan

İlk adım bir *Lagrange* oluşturmak

$$\ell(x,y)=\underbrace{10x+20y-.1(x^2+y^2)}_{\textbf{Amaç fonksiyonu}}+\lambda\underbrace{(100-x-y)}_{\substack{\textbf{Sınırlayıcı}\\ 100-x-y\ge 0}}$$

$$\ell(x,y) = 10x + 20y - .1(x^2 + y^2) + \lambda(100 - x - y)$$

<u>Birinci derece yeter şartlar</u>

$$\ell_x(x,y) = 10 - .2x - \lambda = 0$$

$$\ell_y(x,y) = 20 - .2y - \lambda = 0$$

"Çarpan" şartları

$$\lambda \geq 0 \qquad 100 - x - y \geq 0 \qquad \lambda(100 - x - y) = 0$$

Bu sadece eşitlikle sağlanacak

$$\ell_x(x,y) = 10 - .2x - \lambda = 0$$

$$\ell_y(x,y) = 20 - .2y - \lambda = 0$$

$$100 - x - y = 0$$

$$\boxed{\lambda = 10 - .2x = 20 - .2y}$$

$$y - x = 50$$

$$y + x = 100$$

$$2y = 150$$

$$\boxed{\begin{aligned} y &= 75 \\ x &= 25 \end{aligned}}$$

Çarpan

$$\lambda = 10 - .2x = 20 - .2y$$

$$y = 75$$

$$x = 25$$

$$\lambda = 5 \geq 0$$

Örnek

$f(x,y)=25-x^2-y^2$ denkleminin $4-x-y=0$ şartını sağlayan minimum değerini bulalım.

$$L(x,y,\alpha)=25-x^2-y^2+\alpha\left(4-x-y\right)$$

$$\frac{\partial L(x,y,\alpha)}{\partial x}=-2x+\alpha=0$$

$$\frac{\partial L(x,y,\alpha)}{\partial y}=-2y+\alpha=0$$

$$\frac{\partial L(x,y,\alpha)}{\partial \alpha}=4-x-y=0$$

Örnek Devam

$f(x,y)=25-x^2-y^2$ denkleminin $4-x-y=0$ şartını sağlayan minimum değerini bulalım.

$$x = 0.5\alpha \qquad \alpha = 4$$

$$y = 0.5\alpha \qquad x = 2$$

$$x + y = 4 \qquad y = 2$$

$f(x=2, y=2) = 25 - 2^2 - 2^2 = 17$ bulunur.

Örnek

Maks

$$y = 5x_1x_2$$

$$2x_1 + x_2 = 100$$

Kısıt

$$y_\lambda = 5x_1x_2 + \lambda(100 - 2x_1 - x_2)$$

$$\left.\begin{aligned} \frac{\partial y_\lambda}{\partial x_1} &= 5x_2 - 2\lambda = 0 \\ \frac{\partial y_\lambda}{\partial x_2} &= 5x_1 - \lambda = 0 \\ \frac{\partial y_\lambda}{\partial \lambda} &= 100 - 2x_1 - x_2 = 0 \end{aligned}\right\} \begin{array}{c} \text{3 unknowns:} \\ x_1, x_2, \lambda \\ \text{3 equations} \end{array}$$

$$5x_2 - 2\lambda = 0$$

$$5x_2 = 2\lambda$$

$$\therefore x_2 = \frac{2\lambda}{5}$$

$$5x_1 - \lambda = 0$$

$$5x_1 = \lambda$$

$$x_1 = \frac{\lambda}{5}$$

$$\therefore 100 - 2\left(\frac{\lambda}{5}\right) - \left(\frac{2\lambda}{5}\right) = 0$$

$$500 = 4\lambda$$

$$100 = \frac{4\lambda}{5}$$

$$\therefore \lambda = 125$$

$$x_1 = \frac{\lambda}{5} = \frac{125}{5} = 25$$

$$x_2 = \frac{2\lambda}{5} = \frac{2(125)}{5} = 50$$

Örnek

$$\max \ x - y^2$$
subject to
$$x^2 + y^2 = 4$$
$$x \geq 0, \ y \geq 0.$$

Form the Lagrangian

L=x-y$^2 - \mu(x^2 + y^2 - 4) + \lambda_1 x + \lambda_2 y.$

The first order conditions become:

$$(1)\ \frac{\partial L}{\partial x} = 1 - 2\mu x + \lambda_1 = 0,$$

$$(2)\ \frac{\partial L}{\partial y} = -2y - 2\mu y + \lambda_2 = 0,$$

$$(3)\ x^2 + y^2 - 4 = 0$$

$(4)\ \lambda_1 x = 0,$ $(5)\ \lambda_2 y = 0,$

$(6)\ \lambda_1 \geq 0,$ $(7)\ \lambda_2 \geq 0,$

$(8)\ x \geq 0,$ $(9)\ y \geq 0.$

55

Örnek Devam

By (1), $1+\lambda_1 = 2\mu x$, since $\lambda_1 \geq 0, 1+\lambda_1 > 0$

$\Rightarrow \mu > 0$ and $x > 0$, by (4), $\lambda_1 = 0$.

by (2), $\lambda_2 = 2y(1+\mu)$, since $1+\mu > 0$

$\Rightarrow$ either both y and λ_2 are zero, or both

are positive, from (5) $\Rightarrow y = 0, \lambda_2 = 0$.

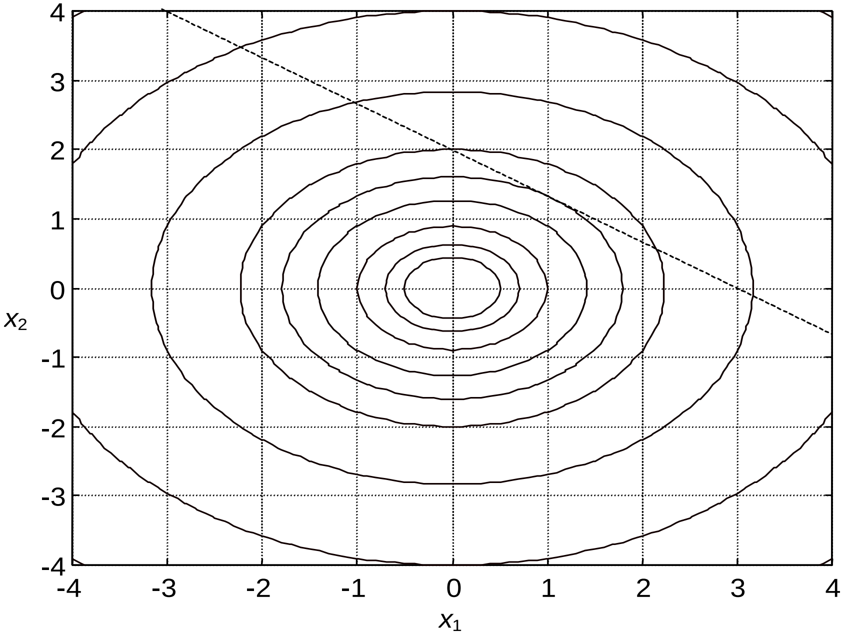
By (3) and (8) $\Rightarrow$ x=2, from (4), $\lambda_1 = 0$,

by (1), $\mu = \dfrac{1}{4}$. So, $(x, y, \mu, \lambda_1, \lambda_2) = (2, 0, \dfrac{1}{4}, 0, 0)$.

Örnek

$$\min_{x_1,x_2} f(\mathbf{x}) = 4x_1^2 + 5x_2^2 \qquad \text{(a)}$$

$$\text{s.t. } 2x_1 + 3x_2 = 6 \qquad \text{(b)}$$



$$L = 4x_1^2 + 5x_2^2 + \lambda(2x_1 + 3x_2 - 6)$$

$$\frac{\partial L}{\partial x_1} = 8x_1 + 2\lambda = 0 \qquad \text{(a)}$$

$$\frac{\partial L}{\partial x_2} = 10x_2 + 3\lambda = 0 \qquad \text{(b)}$$

$$\frac{\partial L}{\partial \lambda} = 2x_1 + 3x_2 - 6 = 0 \qquad \text{(c)}$$

$$x_1 = -\frac{\lambda}{4},\ x_2 = -\frac{3\lambda}{10} \rightarrow -\frac{\lambda}{2} - \frac{9\lambda}{10} - 6 = 0 \rightarrow \lambda = -\frac{30}{7} = -4.281$$

$$\text{Böylece,} \quad x_1 = \frac{15}{14} = \underline{\underline{1.071}},\ x_2 = \frac{90}{70} = \underline{\underline{1.286}}$$

PENALTY VE AUGMENTED LAGRANGIAN METOTLARI

Genel degerlendirme:

- Penalty metotlarin uygulanmasi basittir.
- Penalty katsayilarinin seçimi için genel bir kriter olmadigindan uygun degeri belirlemek zordur (deneme yanilma ile). Her problem için seçim ayridir.
- Augmented Lagrangian metod, penalty metotlarin gelistirilmis halidir. Matematiksel olarak biraz daha karmasiktir, fakat interior ve exterior penalty metotlarin olumsuzluklarinin bir çogunu ortadan kaldirir.
- Augmented Lagrangian metod, optimizasyon probleminin çözümünü penalty katsayisinin (dogru degerinin) seçimine daha az bagimli hale getirir.

DIREK METOTLAR (SEARCH METHODS)

- Problemi dönüstürmeden oldugu gibi (direk) çözmeye çalisirlar. Çözüm metodu olarak kisitsiz optimizasyon problemlerinde anlatilan *arastirma yönü* konseptini (*search directions*) kullanirlar.

- Arastirma yönü hesaplarinda:

✓ hem amaç fonksiyonunu iyilestirecek,

✓ hem de kisitlari ihlal etmeyecek arastirma yönleri kullanilir.

- Kisitli problemlerin çözümünde yaygin kullanilan direk metotlar:

✓ Method of feasible directions

✓ Gradient projection method

METHOD OF FEASIBLE DIRECTIONS

- Kisitli problemlerin çözümünde arastirma yönü konseptini kullanan bir metottur.

  ✓ *Minimize*

  $$f(\mathbf{x}) = f(x_1, x_2, \ldots\ldots, x_n)$$

  ✓ *kisit*

  $$h_k(\mathbf{x}) = 0 \qquad k = 1, \ldots\ldots, n_e$$

  $$g_l(\mathbf{x}) \le 0 \qquad l = 1, \ldots\ldots, n_g$$

- Metot optimum için uygun bir yön seçimine ve bu yön üzerinde ilerleyerek (1-D search) optimumu bulma fikrine dayanir:

$$\mathbf{x} = \mathbf{x}^q + \alpha\, \mathbf{d}^q$$

METHOD OF FEASIBLE DIRECTIONS

Arastirma yönü hesabi:

Arastirma yönü hem amaç fonksiyonunu iyilestirecek, hem de bu yön üzerinde belli bir miktar gidildiginde kisitlari ihlal etmeyecek sekilde seçilir.

- Amaç fonksiyonunu iyilestiren yön kullanilabilir bir yöndür (usable direction):

$$\boxed{\nabla f(x^q) \bullet d^q \leq 0}$$

- Belli bir miktar gidildiginde aktif kisitlari ihlal etmeyen yön ise geçerli bir yöndür (feasible direction):

$$\boxed{\nabla g_a(x^q) \bullet d^q \leq 0}$$

- Her iki istegi saglayan yön ise hem faydali hem geçerli (usable-feasible direction) bir yöndür, kisacasi *optimum bir yöndür.*

METHOD OF FEASIBLE DIRECTIONS

<u>Arastirma yönü hesabi (devam)</u>:

- Konveks problemlerde, kisit ihlalini önlemek için aktif kisit sinirlarindan bir itme faktörü kullanilarak uzak durulur (push-off factor)

$$\nabla g_a(x^q) \bullet d^q + \theta \le 0$$

- Pratikte, amaç fonksiyonunda iyilestirme daha çok olacaksa büyük push-off faktörleri kullanilabilir. ($\nabla f(x^q) \bullet d^q$ ‘nin küçük degerleri için)

$$\nabla g_a(x^q) \bullet d^q - [\nabla f(x^q) \bullet d^q]\theta \le 0$$

- *beta* gibi yeni bir degisken tanimlanirsa, *beta* ’nin maksimizasyonu ile $\nabla f(x^q) \bullet d^q$ ‘nin minimizasyonu empoze edilmis olur.

$$\nabla f(x^q) \bullet d^q + \beta \le 0$$

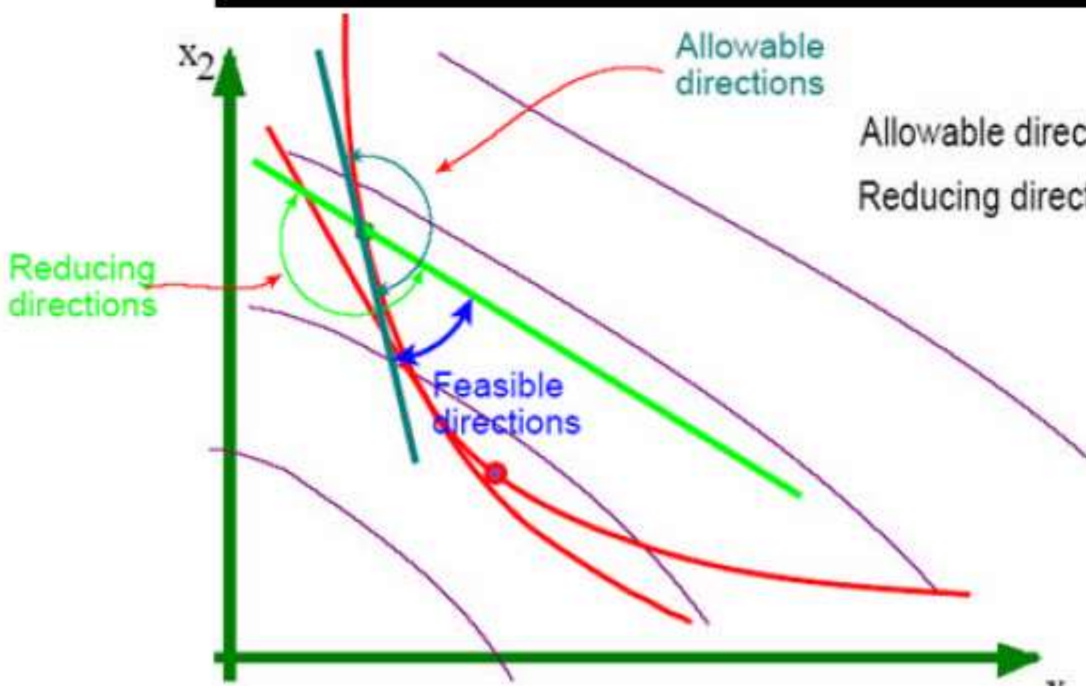
METHOD OF FEASIBLE DIRECTIONS

Arastirma yönü hesabi (devam):

- Maximize $\boxed{\beta}$

- kisit

$$\boxed{\begin{array}{ll} \nabla f(x^q) \bullet d^q + \beta \le 0 & \\ \nabla g_j(x^q) \bullet d^q + \theta_j \beta \le 0 & j \in J_a \end{array}}$$



Allowable direction d: $\nabla g_i \bullet d \le 0$

Reducing direction d: $\nabla g_0 \bullet d \le 0$

METHOD OF FEASIBLE DIRECTIONS

Degerlendirme:

- Çözüm iteratiftir (adim adimdir).

- Sonuç garantilidir (completely robust). Çözüm feasible bir noktadan baslarsa, feasible bölgenin disina asla çikmaz.

- Her yeni çözüm hem feasible'dir, hem de bir öncekinden daha iyidir.

- Arastirma yönü seçiminde türev (gradyent) bilgisi gerekir. Fakat 1-D line search (belirlenen arastirma yönü üzerinde minimumun bulunmasi) için istenen metot kullananilabilir (Golden Section, Polynomial vs.).

- Hatali yön seçimi yakinsamayi yavaslatir, fakat tamamen durdurmaz.

- Metodun programlanmasi biraz karisiktir. Arastirma yönü baska bir optimizasyon probleminin çözümünden elde edilir. Bu da ikinci bir optimizasyon metodu (örnegin Simplex metodu) kullanmayi gerektirir.

GRADIENT PROJECTION METODU

- Gradient projection methodu, kisitsiz problemler için gelistirilen Steepest descent metodundan ilham almistir.

- Arastirma yönü, amaç fonksiyonun negatif gradyenti ’ni aktif kisitlar düzlemine iz düsürerek (projection) elde edilir.

- Iz düsüm matrisi (projection matrix, P) ile feasible yön vektörünün hesabi biraz zahmetlidir (hesap maliyeti fazladir).

GRADIENT PROJECTION METODU
(Iz Düsüm Matrisi Hesabi)

Gradient metodu ile lineer kisitli problemlerin çözümünde çok fazla problem yasanmaz.

Ancak nonlineer kisitli problemlerin çözümü, kisit sinirlarinin egri bir yüzey olusturmasindan dolayi zordur. Egri yüzeyler gradient projection ile bulunan yönde gidilince çözümün hemen infeasible olmasina yol açar. Bu sorunu çözmek için iz düsüm matrisi nonlineer kisitlarda farkli hesaplanir.

Gradient projection metodu ile arastirma yönü (Sk) hesabi: Sk=-Pk*gk
Pk iz düsüm matrisidir ve lineer ve nonlineer kisitlar için farkli hesaplanir.

- Lineer kisit durumunda iz düsüm matrisi (projection matrix) hesabi:

$$\mathbf{P}_k = [\mathbf{I} - \mathbf{A}_q^T(\mathbf{A}_q\mathbf{A}_q^T)^{-1}\mathbf{A}_q]$$

- Nonlineer kisit durumunda iz düsüm matrisi (projection matrix) hesabi:

$$\mathbf{P}_k = [\mathbf{I} - \nabla\mathbf{h}(x_k)^T\,[\nabla\mathbf{h}(x_k)\nabla\mathbf{h}(x_k)^T]^{-1}\,\nabla\mathbf{h}(x_k)]$$

Kaynaklar

- Hasan KURTARAN, “Mühendislik tasarımların optimizasyonu”, 2005
- Edwin K.P. Chong Stanislaw H.Zak, “An Introduction to Optimization”, 2001, John Wiley&Sons
- Jorge Nocedal, Stephan J. Wright, “Numerical Optimization”, 1999, Springer-Verlag
- R. Fletcher, “Practical Methods of Optimization”, 1987, John Wiley&Sons
- David G.Luenberger, “Introduction to Linear and Nonlinear Programming”, 1987, Addison-Wesley
- S.G. Nash, A. Sofer, “Linear and Nonlinear programming”, 1996, McGraw Hill
- Dimitri P. Bertsekas, “Constrained Optimization and Lagrange Multiplier Methods”, 1982, Academic Press
- Abbas Azimli, “Matematiksel Optimizasyon”, 2011, Papatya Yayınevi