KONU 8: DOĞRUSAL OLMAYAN OPTİMİZASYON

Bir matematiksel programlama probleminde, amaç fonksiyonu ile kısıtların bazıları ya da tümü doğrusal olmayan ifadeler ise, probleme, doğrusal olmayan programlama problemi (d.o.p.p.) denir. En iyi çözümü elde edilmek istenilen bir d.o.p.p.

$$\begin{aligned} &\max/\min\ Z = f(\mathbf{X}) \\ &g_i(\mathbf{X}) \le 0 \ , \ i = 1,2,\ldots,t \\ &h_j(\mathbf{X}) = 0 \ , \ j = 1,2,\ldots,m \end{aligned} \tag{8.1}$$

biçiminde tanımlanabilir. Karar değişkenleri vektörü, $\mathbf{X} = [X_1\ X_2 \ldots X_n]'$ olmak üzere, $g_i(\mathbf{X}) \le 0$, $i = 1,2,\ldots,t$ kısıt sistemi ele alınsın. $g_i(\mathbf{X}) = 0$ denklemini sağlayan $\mathbf{X}'$ lerin kümesi, tasarım uzayında bir çok boyutlu yüzey oluşturur. Bu kısıt yüzey, tasarım uzayını, $g_i(\mathbf{X}) < 0$ ve $g_i(\mathbf{X}) > 0$ olmak üzere ikiye böler. $g_i(\mathbf{X}) < 0$ olan bölgedeki noktalar, "uygun/kabul edilebilir" olup, $g_i(\mathbf{X}) > 0$ olan bölgedeki noktalar "uygun olmayandır" yorumu yapılır. $h_j(\mathbf{X}) = 0$, $j = 1,2,\ldots,m$ kısıt yüzeyleri topluluğuna bileşik kısıt yüzeyi denir.

Sınır noktası: Bir ya da daha fazla kısıt yüzeyinde bulunan tasarım noktasıdır.
Serbest nokta: Herhangi bir kısıt yüzeyinde yer almayan noktadır.

Bir d.o.p.p. de üçten çok karar değişkeni varsa, grafiksel çözüm yaklaşımı kullanılarak, probleme çözüm bulunamaz. Problem ancak matematiksel olarak çözülebilir. D.o.p.p. nin tümünü çözebilen genel bir çözüm yöntemi yoktur. Bu amaçla farklı yöntemler geliştirilmiştir.

NOT: Bir $f(\mathbf{X})$ fonksiyonunun minimum noktası $\mathbf{X}^*$ ise, aynı nokta fonksiyonun negatifi olan $-f(\mathbf{X})$ in de maksimum noktasıdır.

8.1. Doğrusal Olmayan Programlama Problemleri

Örnek 1: Bir üretici üç tip ürün üretiyor. Bu ürünlerin aylık satış hacmi ve birim satış fiyatları sırasıyla X_j ve f_j, j=1,2,3 ile gösterilsin. Birinci ve ikinci ürünün satış hacmi kendi fiyatlarına, üçüncü ürünün satış hacmi ise diğer ürünlerin satış hacmine bağlıdır. Daha önceki bilgilere göre, X_j ve f_j arasında aşağıdaki bağıntıların olacağı tahmin ediliyor:

$$X_1 = 10 - f_1$$

$$X_2 = 16 - f_2$$

$$X_3 = 6 - \frac{1}{2} f_3 + \frac{1}{4} f_2$$

Bu üç ürünün birim maliyetleri sırasıyla 6, 7 ve 10 birimdir. Üretim, iş gücü ve makine zamanları ile sınırlıdır. Her ay 1000 makine saatine ve 2000 iş gücü saatine sahiptir. Birinci ürünün her birimi 0.4 makine saati ve 0.2 iş gücü saati, ikinci ürünün her birimi 0.2 makine saati ve 0.4 iş gücü saati, üçüncü ürünün her birimi ise 0.1 makine saati ve 0.1 iş gücü saati kullanıyor. Aylık kazancı maksimum yapacak olan satış miktarını bulmayı amaçlayan problemin matematiksel modelini oluşturunuz.

Çözüm:

Karar değişkenleri: $X_j : j.$ üründen satılacak miktar , j=1,2,3

Amaç fonksiyonu:

$$\begin{aligned} f(\mathbf{X}) &= \left(f_1 X_1 - 6X_1\right) + \left(f_2 X_2 - 7X_2\right) + \left(f_3 X_3 - 10X_3\right) \\ &= \left(\left(10 - X_1\right)X_1 - 6X_1\right) + \left(\left(16 - X_2\right)X_2 - 7X_2\right) + \left(\left(20 - \frac{X_2}{2} - 2X_3\right)X_3 - 10X_3\right) \end{aligned}$$

olmak üzere, d.o.p.p.

$$\begin{aligned} \max\ Z = 4X_1 - X_1^2 + 9X_2 - X_2^2 + 10X_3 - 2X_3^2 - \frac{1}{2} X_2 X_3 \\ 0.4X_1 + 0.2X_2 + 0.1X_3 \le 1000 \\ 0.2X_1 + 0.4X_2 + 0.1X_3 \le 2000 \\ X_1, X_2, X_3 \ge 0 \end{aligned}$$

biçiminde tanımlanır.

Örnek 2: Kenar uzunlukları *a*, *b*, *c* olan üçgen biçimindeki bir arazinin çevresinin yarısı *s*' dir. Bu arazinin çevresine 600 m uzunlukta dikenli tel çekilmek isteniyor. Dikenli tel ile çevrilen üçgen biçimindeki alanı maksimum yapacak biçimde problemi d.o.p.p. biçiminde modelleyiniz.

Çözüm:

NOT:

Üç Kenar Uzunluğu Verilen Üçgenin Alanı

ABC üçgenin çevresi

$Ç(\widehat{ABC}) = a + b + c$ olmak üzere,

$$u = \frac{a+b+c}{2} = \frac{Ç(\widehat{ABC})}{2}$$

$$\boxed{A(\widehat{ABC}) = \sqrt{u \cdot (u-a) \cdot (u-b) \cdot (u-c)}}$$

şeklindedir.

Yukarıdaki bilgiden yararlanarak, verilen problemin matematiksel modeli

$$\begin{aligned}
\max\ Z &= \sqrt{s(s-a)(s-b)(s-c)} \\
a+b+c &= 600 \\
a &\leq b+c \\
b &\leq a+c \\
c &\leq a+b \\
a,b,c &\geq 0
\end{aligned}$$

biçiminde elde edilir.

Örnek 3: Bir maden işletmesinin B1, B2, B3 ve B4 bölgelerinde birer maden ocağı bulunmaktadır. Bu dört bölgeden çıkarılan madenleri depolamak için büyük bir depo yapılması isteniyor. B2 ocağı, B1 ocağının 150 km doğusunda ve 260 km kuzeyindedir. B3 ocağı, B2 ocağının 150 km doğusunda ve 140 km güneyindedir. B4 ocağı, B3 ocağının 50 km doğusunda ve B1 ocağının 160 km kuzeyindedir. Depo ile maden ocakları arasındaki bağlantıyı sağlayacak yolların uzunluğu minimum olacak biçimde, depo için en uygun yeri belirlemek amacıyla oluşturulabilecek amaç fonksiyonu ne olur?

Çözüm: (Kısıtsı d.o.p.p.)

$$\begin{aligned}
\min f(\mathbf{X}) = &\sqrt{X_1^2 + X_2^2} + \sqrt{(X_1-150)^2 + (X_2-260)^2} + \sqrt{(X_1-300)^2 + (X_2-120)^2} \\
&+ \sqrt{(X_1-350)^2 + (X_2-160)^2}
\end{aligned}$$

8.2. Matematiksel Gösterimler

Tanım 1: (Öklid Uzayı)

Üzerinde bir iç çarpım fonksiyonu tanımlanmış sonlu boyutlu vektör uzayıdır.

Tanım 2: (Norm)

E^n Öklid uzayı verilsin. Burada, $\mathbf{x} = \left[x_1\, x_2 \ldots x_n\right]'$ ve $\mathbf{y} = \left[y_1\, y_2 \ldots y_n\right]'$ vektörleri tanımlansın. $\mathbf{x}$ ve $\mathbf{y}$ vektörleri arasındaki uzaklık

$$\|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}$$

biçiminde tanımlanır.

$\mathbf{y} = \mathbf{0}$ olması durumunda, $\mathbf{x}$' in normu $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$ olur.

Tanım 3: (İç Çarpım)

$\mathbf{x}$, $\mathbf{y} \in E^n$ olmak üzere, $\mathbf{x}$ ve $\mathbf{y}$ vektörleri arasındaki iç çarpım

$$\mathbf{x}'\mathbf{y} = \sum_{i=1}^{n} x_i y_i$$

biçiminde tanımlanır. Eğer, iç çarpım sıfır ise ($\mathbf{x}'\mathbf{y} = 0$), $\mathbf{x}$ ve $\mathbf{y}$ vektörleri birbirine diktir.

NOT: (Cauchy-Schwartz Eşitsizliği) $|\mathbf{x}'\mathbf{y}| \le \|\mathbf{x}\|\,\|\mathbf{y}\|$

Tanım 4: (Doğru)

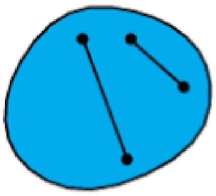
E^n Öklid uzayında, $S \neq \emptyset$ kümesi tanımlansın. x_1, $x_2 \in E^n$ olsun. Bu iki farklı noktadan geçen doğru,

$$S = \left\{x : x = \lambda x_2 + (1-\lambda) x_1 \text{ , } -\infty < x < \infty\right\}$$

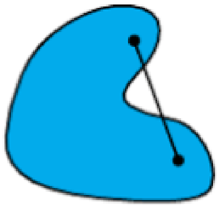
biçiminde verilen S noktalar kümesi olarak tanımlanır ya da S noktalar kümesine, x_1 ve x_2 den geçen doğru denir. Burada, λ skalerdir. $0 < \lambda < 1$ için, S, x_1 ve x_2' yi birleştiren doğru parçası adını alır.

Tanım 5: (Dışbükey Küme)

x_1 ve x_2 noktalarını birleştiren doğru parçası da S kümesinin bir öğesi ise, S kümesine dışbükey (konveks) küme denir.



Dışbükey küme



Dışbükey olmayan küme

Tanım 6: (Dışbükey Bileşim)

$x_1, x_2, \ldots, x_n$ aynı kümenin farklı noktaları iken $\lambda_i \geq 0$, $i = 1,2,\ldots,n$ ve $\sum_{i=1}^{n} \lambda_i = 1$ olmak üzere,

$$x_0 = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n = \sum_{i=1}^{n} \lambda_i x_i$$

biçiminde elde edilen x_0 ' a, verilen noktaların dışbükey bileşimi denir.

$(1-\lambda)x_1 + \lambda x_2$ biçimindeki noktalar, $\lambda_i \geq 0$, $i = 1,2,\ldots,n$ ve $\sum_{i=1}^{n} \lambda_i = 1$ olacak biçimde, x_1 ve x_2 noktalarının dışbükey bileşimleridir. Burada, $\lambda_1 = (1-\lambda)$ ve $\lambda_2 = \lambda$ dır.

Tanım 7: (Çok Boyutlu Düzlem)

E^n' de çok boyutlu düzlem,

$S = \{\mathbf{x} : \mathbf{c}'\mathbf{x} = z\}$

noktalar kümesi olarak tanımlanır. Burada, $\mathbf{c} \neq \mathbf{0}$ ve z sabittir.

Örneğin, $S = \{(x_1, x_2, x_3) : x_1 + 2x_2 - x_3 = 4\} \in E^3$ dır.

Tanım 8: (Kapalı Yarım Uzay ve Açık Yarım Uzay)

Çok boyutlu düzlem E^n' yi

$S = \{\mathbf{x} : \mathbf{c}'\mathbf{x} \geq z\}$ ve $S = \{\mathbf{x} : \mathbf{c}'\mathbf{x} \leq z\}$ biçiminde iki kapalı yarım uzaya ayırır. Kesin eşitsizlik durumunda, S' ye açık yarım uzay denir.

Tanım 9: (Uç Nokta)

S, E^n uzayında tanımlı bir dışbükey küme olsun. x_1, $x_2 \in S$, $\lambda \in [0,1]$ ve $x = (1-\lambda)x_1 + \lambda x_2$, $x = x_1 = x_2$ ise, $x \in S$, S' nin bir uç noktası adını alır.

$\{\mathbf{x} : \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ biçimindeki çok boyutlu küme, sonlu sayıda uç noktaya sahiptir.

Teorem: $S = \{\mathbf{x} : \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ boş olmayan bir küme olsun. Burada, $\mathbf{A}$, rankı m olan $m \times n$ boyutlu matris ve $\mathbf{b}$, m boyutlu vektördür. Bu özelliğe sahip S kümesinin en az bir uç noktası vardır.

Özellik 1: Bir kümenin farklı iki noktasının dışbükey bileşimi olarak yazılamayan noktası var ise, bu noktaya uç nokta (köşe nokta) denir.

Özellik 2: Eğer, $f(\mathbf{x}) = \mathbf{c}'\mathbf{x}$ fonksiyonunun en iyi çözümü var ise, bu nokta uygun çözüm alanının bir uç noktasıdır. Amaç fonksiyonu en iyi değerini birden fazla uç noktada alıyorsa, bu noktaların her dışbükey bileşimi de en iyi çözümdür.

Tanım 10: (Dışbükey Fonksiyon)

S, E^n' de tanımlı boş olmayan bir küme olsun. $f(\mathbf{x})$ fonksiyonu, S' de dışbükey bir fonksiyon ise,

$$f(\lambda x_2 + (1-\lambda)x_1) \leq \lambda f(x_2) + (1-\lambda) f(x_1)$$

biçiminde tanımlanır. Burada, x_1, $x_2 \in S$ ve $\lambda \in [0,1]$ dir.

$f(\mathbf{x})$ fonksiyonu, kesin dışbükey bir fonksiyon ise,

$$f(\lambda x_2 + (1-\lambda)x_1) < \lambda f(x_2) + (1-\lambda) f(x_1)$$

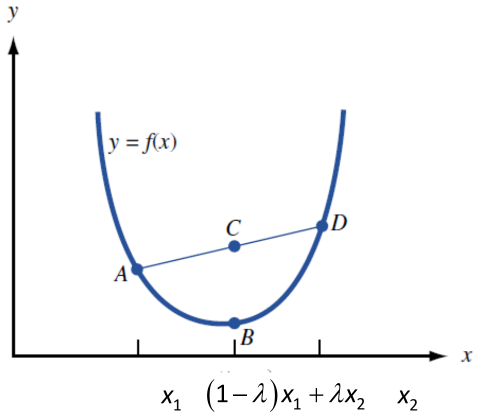
biçiminde tanımlanır.

Buna göre, $-f(\mathbf{x})$ fonksiyonu, içbükey (konkav) fonksiyon olacaktır.

Dışbükey fonksiyonların toplamı dışbükey bir fonksiyon, içbükey fonksiyonların toplamı da içbükey bir fonksiyon olur.

Doğrusal olmayan programlamada, ele alınan bir fonksiyonun dışbükey veya içbükey olduğunun belirlenebilmesi oldukça önemlidir.

Bir dışbükey fonksiyonun grafiksel gösterimi, Şekil 8.1’ de verilmiştir.



Şekil 8.1. Dışbükey bir fonksiyonun grafiksel gösterimi

Burada,

$$A=\left(x_1, f\left(x_1\right)\right)$$

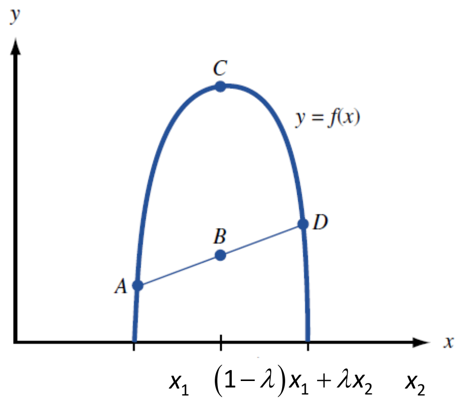
$$D=\left(x_2, f\left(x_2\right)\right)$$

$$B=\left(\lambda x_2+(1-\lambda)x_1, f\left(\lambda x_2+(1-\lambda)x_1\right)\right)$$

$$C=\left(\lambda x_2+(1-\lambda)x_1,\ \lambda f\left(x_2\right)+(1-\lambda)f\left(x_1\right)\right)$$

olarak tanımlıdır.

Bir içbükey fonksiyonun grafiksel gösterimi ise, Şekil 8.2’ deki gibi olacaktır.



Şekil 8.2. İçbükey bir fonksiyonun grafiksel gösterimi

Burada,

$$A=\left(x_1, f\left(x_1\right)\right)$$

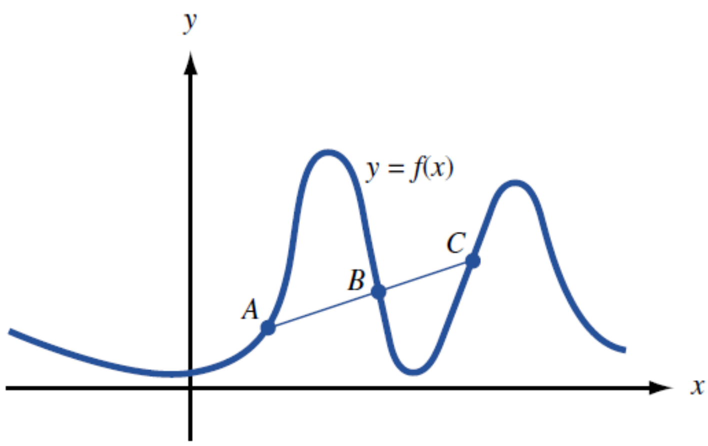
$$D=\left(x_2, f\left(x_2\right)\right)$$

$$C=\left(\lambda x_2+(1-\lambda)x_1, f\left(\lambda x_2+(1-\lambda)x_1\right)\right)$$

$$B=\left(\lambda x_2+(1-\lambda)x_1,\ \lambda f\left(x_2\right)+(1-\lambda)f\left(x_1\right)\right)$$

olarak tanımlıdır.

Şekil 8.3’ te, ne konveks ne de konkav bir fonksiyonun grafiği görülmektedir.



Şekil 8.3. Ne dışbükey ne de içbükey bir fonksiyonun grafiksel gösterimi

Örnek 4: *n* boyutlu uzayda, çok boyutlu düzlemin bir dışbükey küme olduğunu gösteriniz.

Çözüm:

$S=\{(x_1, x_2, \dots, x_n): d_1x_1+d_2x_2+\dots+d_nx_n=b\}\in E^n$ olmak üzere $\mathbf{x}_1$ ve $\mathbf{x}_2$ noktalarının dışbükey bileşimi $\mathbf{x}=\lambda\mathbf{x}_2+(1-\lambda)\mathbf{x}_1$ dir.

$\mathbf{D}'=[d_1\ d_2 \dots d_n]$ olmak üzere

$$\begin{aligned}\mathbf{D}'\mathbf{x}&=\mathbf{D}'(\lambda\mathbf{x}_2+(1-\lambda)\mathbf{x}_1)\\&=\mathbf{D}'\lambda\mathbf{x}_2+\mathbf{D}'\mathbf{x}_1-\mathbf{D}'\lambda\mathbf{x}_1\\&=\lambda\mathbf{D}'\mathbf{x}_2+(1-\lambda)\mathbf{D}'\mathbf{x}_1\\&=b(\lambda+1-\lambda)\\\mathbf{D}'\mathbf{x}&=b\end{aligned}$$

bulunur.

Örnek 5: Bir yarım uzayın üst tarafında bulunan noktaların oluşturduğu kümenin dışbükey küme olduğunu gösteriniz.

Çözüm:

$\mathbf{x}_1$ yarım uzay üzerinde ise, $\mathbf{D}'\mathbf{x}_1=b$

$\mathbf{x}_2$ yarım uzayın üst bölgesinde ise, $\mathbf{D}'\mathbf{x}_2>b$

yazılabilir. Bu iki noktanın, ($\mathbf{x}_1$ ve $\mathbf{x}_2$) dışbükey bileşimi

$$\begin{aligned}\mathbf{D}'\mathbf{x}&=\mathbf{D}'(\lambda\mathbf{x}_2+(1-\lambda)\mathbf{x}_1)\quad,\quad 0<\lambda<1\\&=\lambda\underbrace{\mathbf{D}'\mathbf{x}_2}_{>b}+(1-\lambda)\underbrace{\mathbf{D}'\mathbf{x}_1}_{=b}\end{aligned}$$

olacağından, $\mathbf{D}'\mathbf{x}>b$ dir.

Örnek 6: Bir doğrusal programlama probleminin (d.p.p.), tüm uygun çözümlerinin oluşturduğu kümenin dışbükey küme olduğunu gösteriniz.

Çözüm: $\mathbf{x}$ noktası, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ noktalarının doğrusal bileşimi olarak,

$$\lambda_1\mathbf{x}_1+\lambda_2\mathbf{x}_2+\dots+\lambda_n\mathbf{x}_n=\mathbf{x}$$

biçiminde yazılır. Burada, $\lambda_i\geq 0$, $i=1,2,\dots,n$ ve $\sum_{i=1}^{n}\lambda_i=1$ dir. Bu tanımlardan,

$$\begin{aligned}\mathbf{Ax} &= \mathbf{A}\left(\lambda_1\mathbf{x}_1+\lambda_2\mathbf{x}_2+...+\lambda_n\mathbf{x}_n\right)\\ &= \lambda_1\underbrace{\mathbf{Ax}_1}_{=\mathbf{b}}+\lambda_2\underbrace{\mathbf{Ax}_2}_{=\mathbf{b}}+...+\lambda_n\underbrace{\mathbf{Ax}_n}_{=\mathbf{b}}\end{aligned}$$

olup,

$$\mathbf{Ax}=\mathbf{b}\left(\lambda_1+\lambda_2+...+\lambda_n\right)$$

bulunur. $\sum_{i=1}^{n}\lambda_i=1$ olduğundan, $\mathbf{Ax}=\mathbf{b}$ dir.

Örnek 7: Dışbükey fonksiyonların toplamının da dışbükey fonksiyon olduğunu gösteriniz.

Çözüm: $S\subset E^m$ bir dışbükey küme ve $f_i(\mathbf{x})$, $i=1,2,...,m$ dışbükey fonksiyonlar olsun.

$G(\mathbf{x})=\sum_{i=1}^{m}f_i(\mathbf{x})$ olarak alınsın.

$$f_i\left(\lambda x_2+(1-\lambda)x_1\right)\le\lambda f_i(x_2)+(1-\lambda)f_i(x_1)\quad,\quad 0\le\lambda\le1$$

$$\sum_{i=1}^{m}f_i\left(\lambda x_2+(1-\lambda)x_1\right)\le\lambda\sum_{i=1}^{m}f_i(x_2)+(1-\lambda)\sum_{i=1}^{m}f_i(x_1)\quad,\quad x_1,x_2\in S$$

$$G\left(\lambda x_2+(1-\lambda)x_1\right)\le\lambda G(x_2)+(1-\lambda)G(x_1)$$

olup, G fonksiyonu dışbükeydir.

Örnek 8: $f(x)=x^2$, $x\in R$, fonksiyonunun dışbükey fonksiyon olduğunu gösteriniz.

Çözüm:

$f(x)$ fonksiyonu, $f\left(\lambda x_2+(1-\lambda)x_1\right)\le\lambda f(x_2)+(1-\lambda)f(x_1)$, $0\le\lambda\le1$ biçiminde gösterilmeli.

$$\begin{aligned}&\left(\lambda x_2+(1-\lambda)x_1\right)^2\le\lambda(x_2)^2+(1-\lambda)(x_1)^2\\ &\lambda^2x_2^2+2\lambda(1-\lambda)x_1x_2+(1-\lambda)^2x_1^2\le\lambda x_2^2+(1-\lambda)x_1^2\\ &\lambda^2x_2^2+2\lambda(1-\lambda)x_1x_2+(1-\lambda)^2x_1^2-\lambda x_2^2-(1-\lambda)x_1^2\le0\\ &-\lambda(1-\lambda)x_2^2+2\lambda(1-\lambda)x_1x_2-\lambda(1-\lambda)x_1^2\le0\\ &-\lambda(1-\lambda)\left(x_2^2-2x_1x_2+x_1^2\right)\le0\\ &-\underbrace{\lambda}_{>0}\underbrace{(1-\lambda)}_{>0}\underbrace{(x_1-x_2)^2}_{>0}\le0\end{aligned}$$

olduğundan, $f(x)=x^2$ fonksiyonu dışbükeydir.

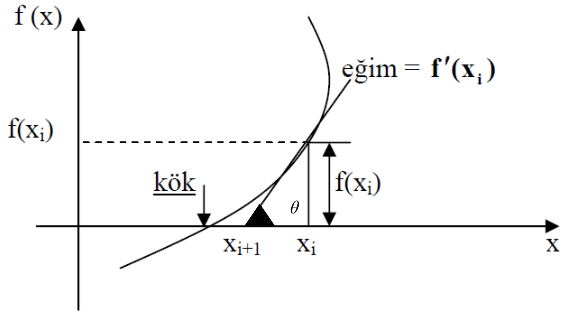
8.3. Doğrusal Olmayan Denklemlerin Çözümü

Newton-Raphson Yöntemi

Eş anlı doğrusal olmayan eşitliklerin çözüm kümesinin bulunmasında kullanılan yöntemlerden biri de Newton-Raphson (N-R) yöntemidir. N-R yöntemi, doğrusal olmayan denklemlerin ve denklem sistemlerinin çözümü için iteratif (yinelemeli) bir yaklaşım sunar.

- **Tek Değişkenli Doğrusal Olmayan Denklemlerin Çözümü**

Tek değişkenli bir $f(x)$ fonksiyonu Şekil 8.4' teki gibi tanımlansın. $f(x)=0$ denklemini sağlayan x değerinin elde edilmesi (kök değerinin bulunması) ile ilgilenilsin. $f(x)=0$ denkleminin köklerinden biri yaklaşık bir değer olarak x_0 olsun. x_0 değeri, Şekil 8.4' te x_i ile belirtilmiştir. x_i noktasından çizilen dikey çizginin eğriyi kestiği noktadaki teğetinin, x-eksenini kestiği nokta (x_{i+1} noktası) kök noktasına daha yakındır. Buna göre amaç, x_i noktası biliniyorken, köke daha yakın olan x_{i+1} noktasını elde etmek olmalıdır.



Şekil 8.4 Newton-Raphson formülünün geometrik açıklaması

x_{i+1} noktası da, $f(x)$ fonksiyonunun x_i noktasındaki eğiminden bulunacaktır. Burada,

$\tan\theta = f'(x_i)$

dir. Buna göre,

$$\tan\theta = f'(x_i) = \frac{f(x_i)}{x_i - x_{i+1}} \Rightarrow x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)} \tag{8.1}$$

elde edilir. Eşitlik (8.1) ile tanımlı yinelemeli nokta değeri x_{i+1}, Taylor derisi kullanılarak da elde edilebilir.

NOT 1: Tek değişkenli bir $f(x)$ fonksiyonunun x_0 civarındaki Taylor açılımı

$$f(x)=f(x_0)+\frac{(x-x_0)}{1!}f'(x_0)+\frac{(x-x_0)^2}{2!}f''(x_0)+\ldots+\frac{(x-x_0)^n}{n!}f^n(x_0)+\ldots$$

Buna göre, 1. dereceli Taylor açılımından (2 ve sonraki dereceli türevler ihmal edilebilir)

$$f(x)=f(x_0)+(x-x_0)f'(x_0)=0$$

$$x=x_0-\frac{f(x_0)}{f'(x_0)} \tag{8.2}$$

elde edilir. Genel ifade,

$$x_{i+1}=x_i-\frac{f(x_i)}{f'(x_i)} \tag{8.3}$$

biçiminde tanımlı olup, N-R yöntemi ile elde edilen denklemdir.

- **Newton-Raphson yönteminin hata analizi**

Verilen $f(x)$ fonksiyonunun gerçel köklerinden birinin ξ olduğunu kabul edilsin. Bu köke yaklaşmakta kullanılan N-R formülü ile gerçek sonuçtan ne kadar uzaklaşıldığı incelenmek istenilsin. Burada, hata miktarı h ile gösterilsin.

i. yinelemede yapılan hata miktarı,

$$h_i=x_i-\xi$$

olur. Buna göre, (*i*+1). adımda yapılan hata miktarı,

$$h_{i+1}=x_{i+1}-\xi$$

dir. Buradan,

$$h_{i+1}-h_i=(x_{i+1}-\xi)-(x_i-\xi)=x_{i+1}-x_i$$

elde edilir. Eşitlik (9.3) ile tanımlı genel ifade kullanılarak (*i*+1). adımda yapılan hata miktarı

$$h_{i+1}=h_i-\frac{f(x_i)}{f'(x_i)} \tag{8.4}$$

biçiminde yazılabilir.

- **Newton-Raphson yakınsaklık koşulu**

Eğer, $f(x)$ fonksiyonunun iki kökü var ise ve bu kökler birbirine çok yakın ise, ortalama değer teoremine göre bu iki kök arasında, $f'(\zeta)=0$ olacak biçimde bir ζ değeri vardır.

Eşitlik (10.2) deki x değeri

$$g(x)=x_0-\frac{f(x_0)}{f'(x_0)}$$

biçiminde tanımlanırsa, N-R yöntemi ile kök bulmak için seçilecek başlangıç noktasının yakınsaklık koşulu

$$\left|g'(x_0)\right|=\left|\frac{f(x_0)\times f''(x_0)}{\left(f'(x_0)\right)^2}\right|<1 \tag{8.5}$$

olarak tanımlanır.

- **Newton-Raphson algoritması**

Adım 1: $f(x)$ fonksiyonu için Eşitlik (8.5) ile tanımlı yakınsaklık koşulu dikkate alınarak bir başlangıç noktası belirlenir.

Adım 2: Gerçel köke yaklaşık bir değer, Eşitlik (8.3) ile tanımlı yaklaşım formülü kullanılarak elde edilir.

Adım 3: $\Delta_{i+1}=\left|x_{i+1}-x_i\right|<\varepsilon$ ise durulur. Son bulunan çözüm, kabul edilen küçük bir $\varepsilon>0$ sabitine göre, köke en yakın değer olarak kabul edilir.

- **Çok Değişkenli Doğrusal Olmayan Denklem Sisteminin Çözümü**

Adım 1:

$$\mathbf{f}(\mathbf{x})=\begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \cdot \\ \cdot \\ \cdot \\ f_m(\mathbf{x}) \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}$$

denklem sistemi belirlenir.

Adım 2: $\mathbf{x}_0$ başlangıç noktası, $\boldsymbol{\varepsilon} > 0$ durdurma koşulu belirlenir.

Adım 3: Denklem sisteminin 1. türev matrisi (J) tanımlanır ve ilgili çözüm noktası için matrisin tersi ($[J]^{-1}$) belirlenir.

Adım 4: Yeni çözüm değeri, $\mathbf{x}_{i+1} = \mathbf{x}_i - [J]^{-1}\mathbf{f}(\mathbf{x}_i)$ hesaplanır.

Adım 5: $\boldsymbol{\Delta}_{i+1} = |\mathbf{x}_{i+1} - \mathbf{x}_i| < \boldsymbol{\varepsilon}$ ise, durulur. Son bulunan çözüm verilen durdurma koşuluna göre köke en yakın çözüm olarak kabul edilir. Aksi halde, durdurma koşulu sağlanıncaya kadar yinelemeli işlemlere devam edilir.

Örnek 9: $f(x) = x^3 + 2x^2 + 10x - 20 = 0$ eşitliği ile tanımlı $f(x)$ fonksiyonunun gerçel kökünü N-R yöntemini kullanarak elde ediniz. Burada, $x_0 = 0$ ve $\varepsilon = 0.01$ alınız.

Çözüm:

$$f(x_0) = f(0) = -20$$
$$f'(x) = 3x^2 + 4x + 10 \Rightarrow f'(x_0) = f'(0) = 10$$

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 0 - \frac{(-20)}{10} = 2$$

$\Delta_1 = |x_1 - x_0| = |2 - 0| = 2 > \varepsilon$ olduğundan, 2. yinelemeye geçilir.

$$f(x_1) = f(2) = 16$$
$$f'(x_1) = f'(2) = 30$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 2 - \frac{16}{30} = 1.46$$

$\Delta_2 = |x_2 - x_1| = |1.46 - 2| = 0.54 > \varepsilon$ olduğundan, 3. yinelemeye geçilir.

$$f(x_2) = f(1.46) = 1.97$$
$$f'(x_2) = f'(1.46) = 22.23$$

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = 1.46 - \frac{1.97}{22.23} = 1.37$$

$\Delta_3 = |x_3 - x_2| = |1.37 - 1.46| = 0.09 > \varepsilon$ olduğundan, 4. yinelemeye geçilir.

$$f(x_3) = f(1.37) = 0.0462$$
$$f'(x_{2.3}) = f'(1.37) = 21.12$$

$$x_4 = x_3 - \frac{f(x_3)}{f'(x_3)} = 1.37 - \frac{0.0462}{21.12} = 1.3678$$

$\Delta_4 = |x_4 - x_3| = |1.3678 - 1.37| = 0.003 < \varepsilon$ olduğundan durulur.

Buna göre, son yinelemede elde edilen $x_4 = x^* = 1.3678$ en iyi çözüm noktasıdır.

$f(x^*) = -0.0213 \cong 0$ bulunur.

Örnek 10:

$$f_1(\mathbf{x}) = x_1^2 - 10x_1 + x_2^2 + 8 = 0$$
$$f_2(\mathbf{x}) = x_1 x_2^2 + x_1 - 10x_2 + 8 = 0$$

biçiminde tanımlı doğrusal olmayan denklem sisteminin çözüm kümesini N-R yöntemini kullanarak elde ediniz. Burada, $\mathbf{x}_0 = \mathbf{0}$ ve $\boldsymbol{\varepsilon} = 0.2$ alınız.

Çözüm:

$$J(\mathbf{x}) = \begin{bmatrix} 2x_1 - 10 & 2x_2 \\ x_2^2 + 1 & 2x_1 x_2 - 10 \end{bmatrix}$$

$$\mathbf{x}_1 = \mathbf{x}_0 - \left[J(\mathbf{x}_0)\right]^{-1} \mathbf{f}(\mathbf{x}_0)$$
$$= \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} -1/10 & 0 \\ -1/100 & -1/10 \end{bmatrix} \begin{bmatrix} 8 \\ 8 \end{bmatrix}$$
$$\mathbf{x}_1 = \begin{bmatrix} 0.80 \\ 0.88 \end{bmatrix}$$

$\boldsymbol{\Delta}_1 = |\mathbf{x}_1 - \mathbf{x}_0| = \left| \begin{bmatrix} 0.80 \\ 0.88 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.80 \\ 0.88 \end{bmatrix} \right| > \begin{bmatrix} 0.2 \\ 0.2 \end{bmatrix}$ olduğundan, 2. yinelemeye geçilir.

$$\mathbf{x}_2 = \mathbf{x}_1 - \left[J(\mathbf{x}_1)\right]^{-1} \mathbf{f}(\mathbf{x}_1)$$
$$= \begin{bmatrix} 0.80 \\ 0.88 \end{bmatrix} - \begin{bmatrix} -0.124 & -0.025 \\ -0.026 & -0.122 \end{bmatrix} \begin{bmatrix} 1.4144 \\ 0.61952 \end{bmatrix}$$
$$\mathbf{x}_2 = \begin{bmatrix} 0.99 \\ 0.99 \end{bmatrix}$$

$\mathbf{\Delta}_2 = \left|\mathbf{x}_2 - \mathbf{x}_1\right| = \left|\begin{bmatrix} 0.99 \\ 0.99 \end{bmatrix} - \begin{bmatrix} 0.80 \\ 0.88 \end{bmatrix} = \begin{bmatrix} 0.19 \\ 0.11 \end{bmatrix}\right| < \begin{bmatrix} 0.2 \\ 0.2 \end{bmatrix}$ olduğundan durulur.

Son bulunan değer, en iyi çözümdür.

$\mathbf{x}^* = \mathbf{x}_2 = \begin{bmatrix} 0.99 \\ 0.99 \end{bmatrix} \cong \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$

$f\left(\mathbf{x}^*\right) = \begin{bmatrix} 0.0602 \\ 0.0603 \end{bmatrix} \cong \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ bulunur.