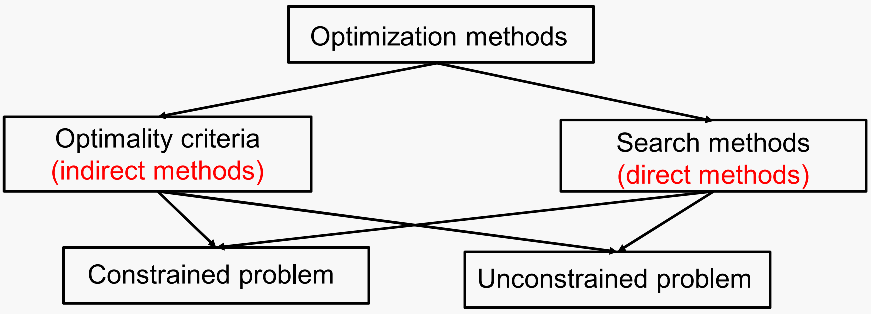
2nd Edition: 4 Optimum Design Concepts

فصل سوم- مفاهیم طراحی بهین

5 More on Optimum Design Concepts



Global (Absolute) Minimum: مینیمم فراگیر(مطلق)

A function *f(x)* of *n* variables has global (absolute) minimum at $\mathbf{x}^*$ if

$$f(\mathbf{x}^*) \le f(\mathbf{x}) \qquad (3.1)$$

for all $\mathbf{x}$ in the feasible design space *S.*

If $f(\mathbf{x}^*) < f(\mathbf{x})$, then $\mathbf{x}^*$ is: a strong (strict) global minimum

If $f(\mathbf{x}^*) = f(\mathbf{x})$, then $\mathbf{x}^*$ is: a weak global minimum

Local (Relative) Minimum: مینیمم محلی(نسبی)

A function *f(**x**)* of *n* variables has a local (relative) minimum at $\mathbf{x}^*$ if $f(\mathbf{x}^*) \le f(\mathbf{x})$ holds for all $\mathbf{x}$ in a small neighborhood *N* of $\mathbf{x}^*$ in the feasible design space *S*.

If $f(\mathbf{x}^*) < f(\mathbf{x})$, then $\mathbf{x}^*$ is: a strong (strict) local minimum

If $f(\mathbf{x}^*) = f(\mathbf{x})$, then $\mathbf{x}^*$ is: a weak local minimum

Neighborhood *N* of the point $\mathbf{x}^*$ is defined as the set of points

$$N = \left\{ X \,\middle|\, X \in S \text{ with} \left\| X - X^* \right\| < \delta \right\}$$

for some small δ>0. Geometrically, it is a small feasible region around the point $\mathbf{x}^*$.

Theorem 4.1- Weierstrass Theorem - Existence of Global Minimum

If *f (x)* is continuous on a nonempty feasible set *S* that is closed and bounded, then *f (x)* has a global minimum in *S*.

A set *S* is closed if it includes all its boundary points and every sequence of points has a subsequence that converges to a point in the set. A set is bounded if for any point, $x \in s, x^T x < c$, where *c* is a finite number.

Local ▲

Global ●

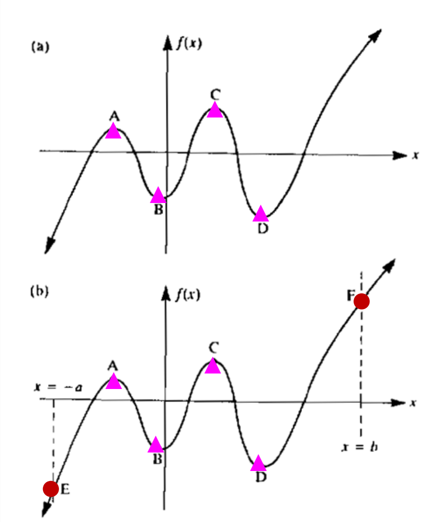


FIGURE 3-2 Graphical representation of optimum points. (a) Unbounded domain and function (no global optimum). (b) Bounded domain and function (global minimum and maximum exist).

EXAMPLE 4.4 Existence of Global Minimum Using Weierstrass Theorem

Consider a function *f(x)=-1/x* defined on the set *S*={*x*|0<*x*≤1}. Check existence of a global minimum for the function.

Solution. The feasible set *S* is not closed since it does not include the boundary point *x=0.* The conditions of the Weierstrass theorem are not satisfied, although *f* is continuous on *S.* Existence of a global minimum is not guaranteed and indeed there is no point *x** satisfying *f(x*)≤f (x)* for all *x ϵ S*.

If we define *S=* {*x*|0≤*x*≤1}, then the feasible set is closed and bounded. However, *f* is not defined at *x=0* (hence not continuous), so the conditions of the theorem are still not satisfied and there is no guarantee of a global minimum for *f* in the set *S*.

Gradient Vector

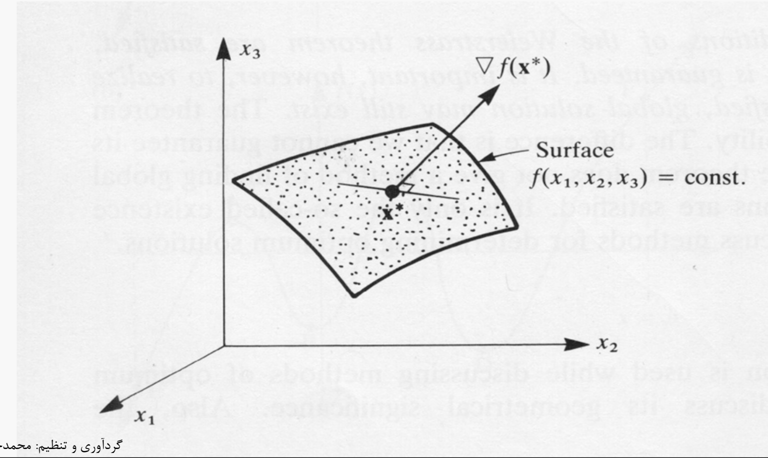
$$c_i = \frac{\partial f\,(X^*)}{\partial x_i}; \quad i = 1 \text{ to } n$$

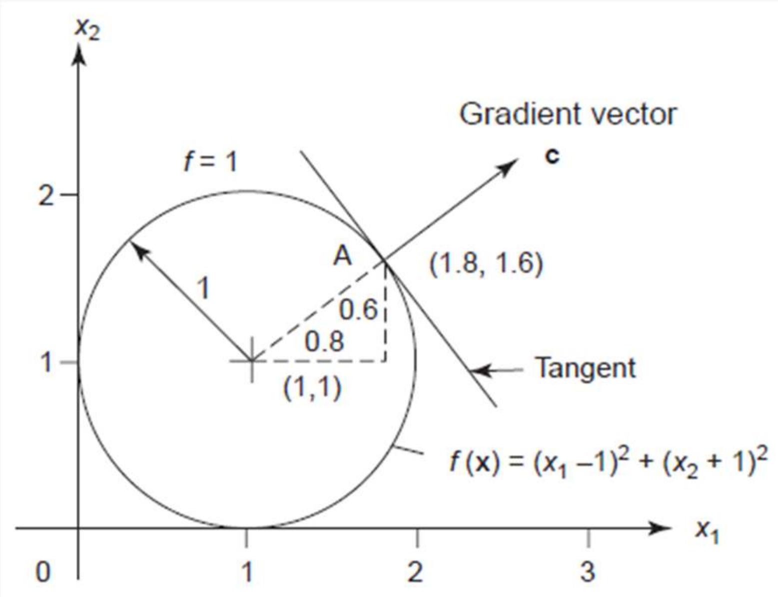
$$\nabla f\,(X^*) = \begin{bmatrix} \dfrac{\partial f\,(X^*)}{\partial x_1} \\ \dfrac{\partial f\,(X^*)}{\partial x_2} \\ \vdots \\ \dfrac{\partial f\,(X^*)}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f\,(X^*)}{\partial x_1} & \dfrac{\partial f\,(X^*)}{\partial x_2} & \cdots & \dfrac{\partial f\,(X^*)}{\partial x_n} \end{bmatrix}^T$$

Note that all partial derivatives are calculated at the given point x^*

Geometrically, the gradient vector is normal to the tangent plane at the point x^*.

Also, it points in the direction of maximum increase in the function.





ماتریس هسیان

$$\frac{\partial^2 f}{\partial X\,\partial X}=\begin{bmatrix}\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_n}\\ \frac{\partial^2 f}{\partial x_2\partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2\partial x_n}\\ \vdots & \vdots & & \vdots\\ \frac{\partial^2 f}{\partial x_n\partial x_1} & \frac{\partial^2 f}{\partial x_n\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2}\end{bmatrix}$$

$$\frac{\partial^2 f}{\partial x_i\partial x_j}=\frac{\partial^2 f}{\partial x_j\partial x_i};\quad i=1\,\text{to}\,n,\ j=1\,\text{to}\,n$$

$$H=\left[\frac{\partial^2 f}{\partial x_i\partial x_j}\right]_{(n\times n)};\quad i=1\,\text{to}\,n,\ j=1\,\text{to}\,n$$

Example: Hessian calculation

$$f(x) = x_1^3 + x_2^3 + 2x_1^2 + 3x_2^2 - x_1x_2 + 2x_1 + 4x_2$$

$$\begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 3x_1^2 + 4x_1 - x_2 + 2 \\ 3x_2^2 + 6x_2 - x_1 + 4 \end{bmatrix}$$

$$\frac{\partial^2 f}{\partial x_1^2} = 6x_1 + 4; \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = -1$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = -1; \quad \frac{\partial^2 f}{\partial x_2^2} = 6x_2 + 6$$

$$H(1,2) = \begin{bmatrix} 10 & -1 \\ -1 & 18 \end{bmatrix}$$

9/94

4.2.3 Taylor's Expansion

A function can be approximated by polynomials in a neighborhood of any point in terms of its value and derivatives using Taylor's expansion.
Consider first a function *f(x)* of a single variable.
Taylor's expansion for *f(x)* about the point *x** is

$$f(x) = f(x^*) + \frac{df(x^*)}{dx}(x - x^*) + \frac{1}{2}\frac{d^2 f(x^*)}{dx^2}(x - x^*)^2 + R$$

R is the remainder term that is smaller in magnitude than the previous terms if *x* is sufficiently close to *x**.

let *x-x*=d* (a small change in the point *x**)

$$f(x^* + d) = f(x^*) + \frac{df(x^*)}{dx}d + \frac{1}{2}\frac{d^2 f(x^*)}{dx^2}d^2 + R$$

10/94

For a function of two variables $f(x_1,x_2)$, Taylor's expansion at the point (x_1^*,x_2^*) is

$$f(x_1,x_2)=f(x_1^*,x_2^*)+\frac{\partial f}{\partial x_1}(x_1-x_1^*)+\frac{\partial f}{\partial x_2}(x_2-x_2^*)$$

$$+\frac{1}{2}\Big[\frac{\partial^2 f}{\partial x_1^2}(x_1-x_1^*)^2+2\frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1-x_1^*)(x_2-x_2^*)$$

$$+\frac{\partial^2 f}{\partial x_2^2}(x_2-x_2^*)^2\Big]+R$$

$$f(x_1,x_2)=f(x_1^*,x_2^*)+\sum_{i=1}^{2}\frac{\partial f}{\partial x_i}(x_i-x_i^*)$$

$$+\frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\partial^2 f}{\partial x_i \partial x_j}(x_i-x_i^*)(x_j-x_j^*)+R$$

$$f(x)=f(x^*)+\nabla f^T(x-x^*)+\frac{1}{2}(x-x^*)^T H(x-x^*)+R$$

$$f(x^*+d)=f(x^*)+\nabla f^T d+\frac{1}{2}d^T H d+R$$

$$\Delta f=f(x^*+d)-f(x^*)$$

$$\Delta f=\nabla f^T d+\frac{1}{2}d^T H d+R$$

A first-order change in $f(x)$ at x^* (denoted as δf) is obtained by retaining only the first term

$\delta f=\nabla f^T \delta x$ δx is a small change in x^* ($\delta x = x - x^*$).

EXAMPLE 4.8 Taylor's Expansion of a Function of Two Variables

Obtain second-order Taylor's expansion for the function $f(x)=3x_1^3x_2$ at the point $x^*=(1,1)$:

Solution. The gradient and Hessian of the function $f(x)$ at the point $x^*=(1,1)$ are

$$\nabla f(x)=\begin{bmatrix}\dfrac{\partial f}{\partial x_1}\\ \dfrac{\partial f}{\partial x_2}\end{bmatrix}=\begin{bmatrix}9x_1^2x_2\\ 3x_1^3\end{bmatrix}=\begin{bmatrix}9\\3\end{bmatrix}$$

$$H=\begin{bmatrix}18x_1x_2 & 9x_1^2\\ 9x_1^2 & 0\end{bmatrix}=\begin{bmatrix}18 & 9\\ 9 & 0\end{bmatrix}$$

$$\bar{f}(x)=3+\begin{bmatrix}9\\3\end{bmatrix}^T\begin{bmatrix}(x_1-1)\\(x_2-1)\end{bmatrix}+\frac{1}{2}\begin{bmatrix}(x_1-1)\\(x_2-1)\end{bmatrix}^T\begin{bmatrix}18 & 9\\ 9 & 0\end{bmatrix}\begin{bmatrix}(x_1-1)\\(x_2-1)\end{bmatrix}$$

$$\bar{f}(x)=9x_1^2+9x_1x_2-18x_1-6x_2+9$$

Accuracy of the second-order Taylor series expansion about (1, 1) for the function $f(\mathbf{x})=3x_1^3x_2$ of Example 3.5

x_1	x_2	f(x)	$\bar{f}$(x)	% Error
1.00	1.00	3.0000	3.0000	0.000
1.05	1.05	3.6465	3.6450	0.041
1.10	1.10	4.3923	4.3800	0.280
1.15	1.15	5.2470	5.2050	0.800
1.20	1.20	6.2208	6.1200	1.620
1.25	1.25	7.3242	7.1250	2.720
1.30	1.30	8.5683	8.2200	4.065
1.35	1.35	9.9645	9.4050	5.615
1.40	1.40	11.5248	10.6800	7.330
1.45	1.45	13.2615	12.0450	9.173
1.50	1.50	15.1875	13.5000	11.111

EXAMPLE 3.6 (4.9) Linear Taylor's Expansion of a Function

Obtain linear Taylor's expansion for the function

$$f(X)=x_1^2+x_2^2-4x_1-2x_2-4$$

at the point x^*=(1,2). Compare the approximate function with the original function in a neighborhood of the point (1,2).

$$\nabla f=\begin{bmatrix}\dfrac{\partial f}{\partial x_1}\\ \dfrac{\partial f}{\partial x_2}\end{bmatrix}=\begin{bmatrix}2x_1-4\\ 2x_2-4\end{bmatrix}=\begin{bmatrix}-2\\ 2\end{bmatrix}$$

linear Taylor series approximation for $f(x)$ is

$$\bar{f}(X)=1+\begin{bmatrix}-2 & 2\end{bmatrix}\begin{bmatrix}x_1-1\\ x_2-2\end{bmatrix}=-2x_1+2x_2-1$$

Accuracy of linear approximation for the function $f(\mathbf{x}) = x_1^2 + x_2^2 - 4x_1 - 2x_2 + 4$ of Example 3.6 at point (1, 2)

x_1	x_2	$f(\mathbf{x})$	$\tilde{f}(\mathbf{x})$	% Error
1.00	2.00	1.0000	1.0000	0.000
1.02	2.00	0.9604	0.9600	0.042
1.04	2.00	0.9216	0.9200	0.174
1.06	2.00	0.8836	0.8800	0.407
1.08	2.00	0.8464	0.8400	0.756
1.10	2.00	0.8100	0.8000	1.235
1.12	2.00	0.7744	0.7600	1.860
1.14	2.00	0.7396	0.7200	2.650
1.16	2.00	0.7056	0.6800	3.628
1.18	2.00	0.6724	0.6400	4.818
1.20	2.00	0.6400	0.6000	6.250
1.00	2.04	1.0816	1.0800	0.148
1.00	2.08	1.1632	1.1600	0.275
1.00	2.12	1.2544	1.2400	1.148
1.00	2.16	1.3456	1.3200	1.903
1.00	2.20	1.4400	1.4000	2.778
1.00	2.24	1.5376	1.4800	3.746
1.00	2.28	1.6384	1.5600	4.785
1.00	2.32	1.7424	1.6400	5.877
1.00	2.36	1.8496	1.7200	7.007
1.00	2.40	1.9600	1.8000	8.163
1.02	2.04	1.0420	1.0400	0.192
1.04	2.08	1.0880	1.0800	0.735
1.06	2.12	1.1380	1.1200	1.582
1.08	2.16	1.1920	1.1600	2.684
1.10	2.20	1.2500	1.2000	4.000
1.12	2.24	1.3120	1.2400	5.488
1.14	2.28	1.3780	1.2800	7.112
1.16	2.32	1.4480	1.3200	8.840
1.18	2.36	1.5220	1.3600	10.640
1.20	2.40	1.6000	1.4000	12.500

4.2.4 Quadratic Forms and Definite Matrices شکل‌های درجه دو و ماتریس‌های معین

$$F(X) = x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_2 + 2x_2x_3 + 2x_1x_3$$

$$F(X) = F(x_1, x_2, \cdots, x_n)$$

$$F(X) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}x_ix_j$$

$$\begin{aligned} F(X) = \frac{1}{2}[&(p_{11}x_1^2 + p_{12}x_1x_2 + \cdots + p_{1n}x_1x_n) \\ &+ (p_{21}x_2x_1 + p_{22}x_2^2 + \cdots + p_{2n}x_2x_n) \\ &+ \cdots + (p_{n1}x_nx_1 + \cdots + p_{nn}x_n^2)] \end{aligned}$$

$$y = PX$$

$$y_i = \sum_{j=1}^{n} p_{ij}x_j, \quad i = 1 \text{ to } n$$

$$F(X) = \frac{1}{2}\sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} p_{ij}x_j\right)$$

ماتریس شکل درجه دو

$$F(X) = \frac{1}{2}\sum_{i=1}^{n} x_i y_i$$

$$F(X) = \frac{1}{2}X^T y = \frac{1}{2}X^T PX$$

P is called the matrix of the quadratic form F(x).

$$\begin{aligned} F(X) = \frac{1}{2}\{&[p_{11}x_1^2 + p_{22}x_2^2 + \cdots + p_{nn}x_n^2] \\ &+ [(p_{12}+p_{21})x_1x_2 + (p_{13}+p_{31})x_1x_3 + \cdots + (p_{1n}+p_{n1})x_1x_n] \\ &+ [(p_{23}+p_{32})x_2x_3 + (p_{24}+p_{42})x_2x_4 + \cdots + (p_{2n}+p_{n2})x_2x_n] \\ &+ \cdots + [(p_{n-1,n}+p_{n,n-1})x_{n-1}x_n)]\} \end{aligned}$$

در حالی که بی نهایت ماتریس P وجود دارد،
ولی یکی از آن‌ها متقارن است که A است.

$$F(X) = \frac{1}{2}X^T PX = \frac{1}{2}X^T AX$$

EXAMPLE 4.10 Matrix of the Quadratic Form

Identify a matrix associated with the quadratic form

$$F(x_1,x_2,x_3) = \frac{1}{2}(2x_1^2 + 2x_1x_2 + 4x_1x_3 - 6x_2^2 - 4x_2x_3 + 5x_3^2)$$

Solution.

$$F(X) = \frac{1}{2}[x_1 \quad x_2 \quad x_3]\begin{pmatrix} 2 & 2 & 4 \\ 0 & -6 & -4 \\ 0 & 0 & 5 \end{pmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Since the coefficient of x_ix_j can be divided between p_{ij} and p_{ji} in any proportion, there are many matrices associated with a quadratic form. For example, the following matrices are also associated with the same quadratic form:

$$\begin{pmatrix} 2 & 0.5 & 1 \\ 1.5 & -6 & -6 \\ 3 & 2 & 5 \end{pmatrix}; \quad \begin{pmatrix} 2 & 4 & 5 \\ -2 & -6 & 4 \\ -1 & -8 & 5 \end{pmatrix}$$

Dividing the coefficients equally between p_{ij} and p_{ji}, we obtain

$$F(X) = \frac{1}{2}[x_1 \quad x_2 \quad x_3]\begin{pmatrix} 2 & 1 & 2 \\ 1 & -6 & -2 \\ 2 & -2 & 5 \end{pmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Form of a Matrix شکل یک ماتریس

Quadratic form $F(x)=x^TAx$ or its symmetric matrix A may be

Positive definite:
If it is always positive for any x [except for $F(0)$]

Negative definite:
If $x^TAx<0$ for all x except $x=0$

Positive semidefinite:
If $x^TAx\geq 0$ for all x and there exists at least one $x\neq 0$ (nonzero x) with $x^TAx = 0$

Negative semidefinite:
If $x^TAx\leq 0$ for all x and there exists at least one $x\neq 0$ (nonzero x) with $x^TAx = 0$

Indefinite
If it is positive for some vectors x and negative for others

EXAMPLE 4.11 Determination of the Form of a Matrix

Determine the form of the following matrices:

$$(i)\quad A=\begin{pmatrix}2&0&0\\0&4&0\\0&0&3\end{pmatrix};\qquad (ii)\quad A=\begin{pmatrix}-1&1&0\\1&-1&0\\0&0&-1\end{pmatrix}$$

Solution. The quadratic form associated with the matrix (i) is always positive, i.e.,

$$X^TAX=(2x_1^2+4x_2^2+3x_3^2)>0$$

unless $x_1=x_2=x_3=0$ ($\mathbf{x}=0$). Thus, the matrix is positive definite.

$$(ii)\quad A=\begin{pmatrix}-1 & 1 & 0\\ 1 & -1 & 0\\ 0 & 0 & -1\end{pmatrix}$$

The quadratic form associated with the matrix (ii) is negative semidefinite, since

$$X^{T}AX=(-x_1^2-x_2^2+2x_1x_2-x_3^2)=\{-x_3^2-(x_1-x_2)^2\}\le 0$$

for all x, and $x^TAx=0$ when $x_3=0$, and $x_1=x_2$ [e.g., x=(1,1,0)]. The quadratic form is not negative definite but is negative semidefinite since it can have zero value for nonzero x.

Therefore, the matrix associated with it is also negative semidefinite.

Theorem 4.2

Eigenvalue Check for the Form of a Matrix

Let λ_i, $i=1$ to n be n eigenvalues of a symmetric $n\times n$ matrix A associated with the quadratic form $F(x)=x^TAx$ (since A is symmetric, all eigenvalues are real). The following results can be stated regarding the quadratic form $F(x)$ or the matrix A:

1. Positive definite if and only if all eigenvalues of A are strictly positive, i.e., $\lambda_i>0$, $i=1$ to n.
2. Negative definite if and only if all eigenvalues of A are strictly negative, i.e., $\lambda_i<0$, $i=1$ to n.
3. Positive semidefinite if and only if all eigenvalues of A are nonnegative, i.e., $\lambda_i\ge 0$, $i=1$ to n (note that at least one eigenvalue must be zero for it to be called positive semidefinite).
4. Negative semidefinite if and only if all eigenvalues of A are nonpositive, i.e., $\lambda_i\le 0$, $i=1$ to n (note that at least one eigenvalue must be zero for it to be called negative semidefinite).
5. Indefinite if some $\lambda_i<0$ and some other $\lambda_j>0$.

استفاده از این قضیه مستلزم حل یک معادله n مجهولی و محاسبه مقادیر ویژه ماتریس است.

Rank of a Matrix

The general concept needed to develop the solution procedure for a general $m \times n$ system of equations is known as the rank[U4] of the matrix, defined as the order of the largest nonsingular square submatrix (A matrix having a zero determinant is called a singular matrix) of the given matrix.

Using the Gauss-Jordan elimination procedure, we can transform any $m \times n$ matrix A into the following equivalent form (for $m<n$):

$$A \sim \begin{bmatrix} I_{(r)} & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix} \qquad \text{(B.32)}$$

where $I_{(r)}$ is the $r \times r$ identity matrix. Then r is the rank of the matrix. Note that the identity matrix $I_{(r)}$ is unique for any given matrix.

Example B.6 Rank Determination by Elementary Operations

Determine rank of the following matrix:

$$A = \begin{pmatrix} 2 & 6 & 2 & 4 \\ -2 & -4 & 2 & 2 \\ 1 & 2 & -1 & -1 \end{pmatrix} \qquad \text{(a)}$$

Solution. The elementary operations lead to the following matrices:

$$A \sim \begin{pmatrix} 1 & 3 & 1 & 2 \\ -2 & -4 & 2 & 2 \\ 1 & 2 & -1 & -1 \end{pmatrix} \quad \left(\text{miltiply row 1 by} \frac{1}{2}\right) \qquad \text{(b)}$$

$$\sim \begin{pmatrix} 1 & 3 & 1 & 2 \\ 0 & 2 & 4 & 6 \\ 0 & -1 & -2 & -3 \end{pmatrix} \quad \begin{pmatrix} \text{add 2 times row 1 to row 2} \\ \text{and -1 times row 1 to row 3} \end{pmatrix} \qquad \text{(c)}$$

Slide 23

U4 Rank=رتبه
User١٢/٠٣/٢٠١٩ ؛

$$\sim\begin{pmatrix}1&0&0&0\\0&2&4&6\\0&-1&-2&-3\end{pmatrix}\begin{pmatrix}\text{add -3 times column 1 to column 2;}\\\text{-1 times column 1 to column 3;}\\\text{-2 times column 1 to column 4}\end{pmatrix} \quad \text{(d)}$$

$$\sim\begin{pmatrix}1&0&0&0\\0&1&2&3\\0&0&0&0\end{pmatrix}\begin{pmatrix}\text{multiply row 2 by } \frac{1}{2} \text{ and}\\\text{add to row 3}\end{pmatrix} \quad \text{(e)}$$

$$\sim\left(\begin{array}{cc|cc}1&0&0&0\\0&1&0&0\\\hline 0&0&0&0\end{array}\right)\begin{pmatrix}\text{add -2 times column 2 to column3;}\\\text{and -3 times column 2 to column 4}\end{pmatrix} \quad \text{(f)}$$

This matrix is in the form of Eq. (B.32). The rank of A is 2, since a 2×2 identity matrix is obtained at the upper left corner.

Leading Principal Minor

Every $n\times n$ square matrix A has certain scalars associated with it, called the leading principal minors. They are obtained as determinants of certain submatrices of A.

Let M_k, k=1 to n be called the leading principal minors of A. Then each M_k is defined as the determinant of the following submatrix:

$$M_k = |A_{kk}| \quad \text{(B.26)}$$

where A_{kk} is a $k\times k$ submatrix of A obtained by deleting the last $(n- k)$ columns and the corresponding rows.

For example, $M_1=A_{11}$, M_2=determinant of a 2×2 matrix obtained by deleting all rows and columns of A except the first two, and so on, are the leading principal minors of the matrix A.

Notation

We write D_k or M_k for the leading principal minor of order k. There are $\binom{n}{k}$ principal minors of order k, and we write k for any of the principal minors of order k.

U7 در جزوه فارسی بهینه سازی یادداشت نوشته شده در قضیه ماینرها برای شکل درجه دو
User۱۲/۰۳/۲۰۱۹ ؛

U5 minor=کهاد
User۱۲/۰۳/۲۰۱۹ ؛

U6 leading principal minor=کهاد اصلی مقدم
User۱۲/۰۳/۲۰۱۹ ؛

U1 را به ترتیب در سطر k و n شیئ که ترتیب آنها مهم نباشد؛ که با پرانتز که n شیئ از k انتخاب C(n,k) ترکیب
k!(n-k)! تقسیم بر n! اول و دوم مینویسند می گویند و عبارتست
User۱۲/۰۳/۲۰۱۹ ؛

From:Lecture 5 Principal Minors and the Hessian, Eivind Eriksen, BI Norwegian School of Management, Department of Economics, October 01, 2010
http://home.bi.no/a0710194/Teaching/BI-Mathematics/GRA-6035/2010/lecture5-hand.pdf, 19th, Oct., 2014.

A minor of A of order k is principal (Δ_k) if it is obtained by deleting $n-k$ rows and the $n-k$ columns with the same numbers.

The leading principal Minor of A of order k (D_k) is the minor of order k obtained by deleting the last $n-k$ rows and columns.

Theorem

Let A be a symmetric $n\times n$ matrix. Then we have:

- A is positive definite if $D_k>0$ for all leading principal minors
- A is negative definite if $(-1)^k D_k>0$ for all leading principal minors
- A is positive semidefinite if $\Delta_k \geq 0$ for all principal minors
- A is negative semidefinite if $(-1)^k \Delta_k \geq 0$ for all principal minors
- A is indefinite if non of the above conditions are satisfied.

❑ In the first two cases, it is enough to check the inequality for all the leading principal minors (i.e. for $1\leq k\leq n$).

❑ In the last two cases, we must check for all principal minors (i.e for each k with $1\leq k\leq n$ and for each of the $\binom{n}{k}$ principal minors.

For instance, in a principal minor where you have deleted row 1 and 3, you should also delete column 1 and 3

27/94

Theorem 4.3

Check for the Form of a Matrix Using Principal Minors

Let M_k be the kth leading principal minor of the $n\times n$ symmetric matrix A defined as the determinant of a $k\times k$ submatrix obtained by deleting the last $(n-k)$ rows and columns of A (Appendix B, Section $B.3$). Assume that no two consecutive principal minors are zero. Then A is:

1. Positive definite if and only if all $M_k>0$, $k=1$ to n.
2. Negative definite if and only if $M_k<0$ for k odd and $M_k>0$ for k even, $k=1$ to n.
3. Positive semidefinite if and only if $M_k>0$, $k=1$ to r, where $r<n$ is the rank of A (refer to Appendix B, Section $B.4$ for definition of rank of a matrix).
4. Negative semidefinite if and only if $M_k<0$ for k odd and $M_k>0$ for k even, $k=1$ to $r<n$.
5. Indefinite if it does not satisfy any of the preceding criteria.

This theorem is applicable only if the assumption of no two consecutive principal minors being zero is satisfied.

28/94

Definiteness: Example

Determine the definiteness of the quadratic form

$$Q(x_1,x_2,x_3) = 3x_1^2 + 6x_1x_3 + x_2^2 - 4x_2x_3 + 8x_3^2$$

Solution:

The symmetric matrix associated with the quadratic form Q is

$$A = \begin{pmatrix} 3 & 0 & 3 \\ 0 & 1 & -2 \\ 3 & -2 & 8 \end{pmatrix}$$

The leading principal minors are:

$$D_1 = 3, \quad D_2 = \begin{vmatrix} 3 & 0 \\ 0 & 1 \end{vmatrix} = 3, \quad D_3 = \begin{vmatrix} 3 & 0 & 3 \\ 0 & 1 & -2 \\ 3 & -2 & 8 \end{vmatrix} = 3$$

Since all leading principal minors are positive, *Q* is positive definite.

EXAMPLE 4.12 Determination of the Form of a Matrix

Determine the form of the matrices given in Example 4.11.

$$(i)\quad A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}; \qquad (ii)\quad A = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Solution. For a given matrix *A*, the eigenvalue problem is defined as $Ax=\lambda x$, where λ is an eigenvalue and *x* is the corresponding eigenvector (refer to Section *B.6* in Appendix *B* for more details). To determine the eigenvalues, we set the so-called characteristic determinant to zero $|(A - \lambda I)|=0$.

$$\begin{vmatrix} 2-\lambda & 0 & 0 \\ 0 & 4-\lambda & 0 \\ 0 & 0 & 3-\lambda \end{vmatrix} = 0$$

$\lambda_1=2$, $\lambda_2=3$, and $\lambda_3=4$.

Since all eigenvalues are strictly positive, the matrix is positive definite. The principal minor check of Theorem 4.3 also gives the same conclusion.

$$(ii)\quad A = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

For the matrix (ii), the characteristic determinant of the eigenvalue problem is

$$\begin{vmatrix} -1-\lambda & 1 & 0 \\ 0 & -1-\lambda & 0 \\ 0 & 0 & -1-\lambda \end{vmatrix} = 0 \qquad (-1-\lambda)[(-1-\lambda)^2 - 1] = 0$$

Therefore, the three roots give the eigenvalues as λ_1=-2, λ_2=-1, and λ_3=0. Since all eigenvalues are nonpositive, the matrix is negative semidefinite. To use Theorem 4.3, we calculate the three leading principal minors as

$$M_1 = -1, \quad M_2 = \begin{vmatrix} -1 & 1 \\ 1 & -1 \end{vmatrix} = 0, \; M_3 = \begin{vmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{vmatrix} = 0$$

Since there are two consecutive zero leading principal minors, we cannot use Theorem 4.3.

مسائل زیر را حل کرده و تا دو هفته دیگر تحویل فرمایید:

3) 1,8,16

۳.۳ – مسائل نامقید، شرایط لازم و کافی

ضرورت

- حل مسائل طراحی بهین مقید تعمیم منطقی روش حل مسائل نامقید است
- بیشتر مسائل طراحی بهین مقید را با تبدیل آن‌ها به مسائل نامقید حل می‌کنیم

موارد استفاده شرایط بهینگی:

- به دست آوردن نقاط نامزد مینیمم
- بررسی مینیمم بودن یک نقطه

4.2.5 Concept of Necessary and Sufficient Conditions

مفهوم شرایط لازم و کافی

1. Optimum points must satisfy the necessary conditions. Points that do not satisfy them cannot be optimum.

2. A point satisfying the necessary conditions need not be optimum, i.e., nonoptimum points may also satisfy the necessary conditions.

3. A candidate point satisfying a sufficient condition is indeed optimum.

4. If sufficiency conditions cannot be used or they are not satisfied, we may not be able to draw any conclusions about optimality of the candidate point.

مراحل بدست آوردن شرایط بهینگی محلی

فرض کنید x^* یک نقطه مینیمم محلی برای $f(x)$ باشد. برای مطالعه و بررسی همسایگی، فرض کنید x نقطه‌ای نزدیک x^* باشد. نمو d و Δf را در x^* و $f(x^*)$ به ترتیب زیر تعریف می‌کنیم

$d=x-x^*$ و $\Delta f=f(x)-f(x^*)$

$\Delta f=f(x)-f(x^*)\geq 0$

شرایط بهینگی محلی برای توابع یک‌متغیره

شرایط لازم مرتبه اول

$$f(x)=f(x^*)+f'(x^*)^T d+\frac{1}{2}f''(x^*)d^2+R$$

$$\Delta f=f'(x^*)^T d+\frac{1}{2}f''(x^*)d^2+R$$

شرایط لازم بهینگی: $\Delta f\geq 0$

یعنی Δf باید نامنفی باشد.

شرایط بهینگی برای توابع یک‌متغیره (ادامه)

از آنجا که جمله اول $(f'(x^*)^T d)$ از نظر قدر مطلق از مجموع جملات بعد بزرگتر است، بنابراین با توجه به اختیاری بودن d داریم:

عبارت	علامت و یا مقدار		
$f'(x)$	0	>0	<0
d		<0	>0
$f'(x)d$	0	<0	<0

اگر $f'(x^*)\neq 0$ باشد، جمله $f'(x^*)d$ (و در نتیجه Δf) می‌تواند منفی شود. بنابراین:

شرایط لازم بهینگی محلی (مرتبه اول): $\Delta f\geq 0 \Rightarrow f'(x^*)=0$

تعریف نقاط ایستا

نقاطی که شرط لازم مرتبه اول را برآورده کنند، یعنی

$$f'(x^*) = 0$$

نقاط نامزد مینیمم محلی می‌باشند، که به این نقاط، نقاط ایستا می‌گویند.

شرایط کافی

چون نقاط ایستا شرط لازم $f'(x^*) = 0$ را برآورده می‌کنند:

$$\Delta f = \frac{1}{2} f''(x^*) d^2 + R$$

$$\text{If for every } d \neq 0 : f''(x^*) > 0 \xrightarrow{\text{Then}} \Delta f > 0$$

و x^* حداقل یک مینیمم محلی تابع است. بنابراین معادله $f''(x^*) > 0$ شرایط کافی را بیان می‌کند و گویای این است که تابع در نقاط مینیمم شعاع انحنای مثبت دارد.

توجه:

اگر نقطه‌ای مثل x^* داشته باشیم که شرایط زیر را برآورده کند:

شرط لازم مرتبه اول: $f'(x^*) = 0$

شرط کافی: $f''(x^*) > 0$

آنگاه این نقطه یک مینیمم محلی است.

شرط لازم مرتبه دو

اگر شرط کافی ($f''(x^*) > 0$) برآورده نشود نمی‌توان نتیجه گرفت که این نقطه مینیمم نیست (زیرا این شرط لازم نبوده!). بنابراین شرط لازم دیگر (که مرتبه دو هم هست) عبارت است از:

$$f''(x^*) \geq 0$$

یعنی اگر در نقطه‌ای $f''(x^*) < 0$ آن نقطه نامزد نقطه مینیمم نیست. زیرا جمله دوم بسط تیلور $\frac{1}{2} f''(x^*) d^2$ منفی می‌شود و به این دلیل از شرط $f''(x^*) \geq 0$ به عنوان شرط لازم مرتبه دو یاد شده.

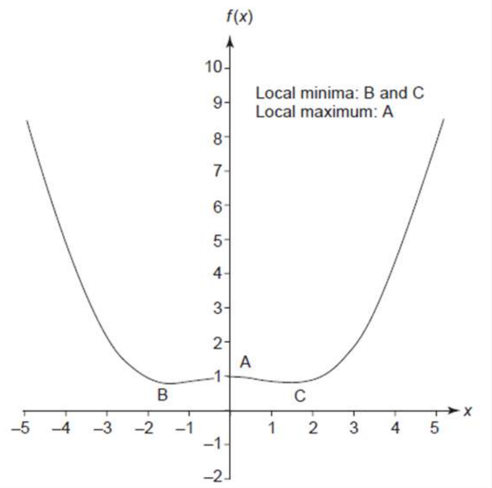
توجه:

اگر $f''(x^*) = 0$ ، شرط لازم و کافی به صورت زیر تغییر می‌کند:

شرط لازم: $$f'''(x^*) = 0$$

شرط کافی: $$f^{(4)}(x^*) > 0$$

به طور کلی پایین ترین مرتبه مشتق غیر صفر باید برای نقاط ایستا زوج (شرط لازم)، و برای نقاط مینیمم محلی مثبت باشد (شرط کافی)



The necessary condition is
f′=(2x/3)-sin x =0 (b)
It can be seen that *x=0* satisfies Eq. (b), so it is a stationary point.
Newton-Raphson method gives:
*x**=1.496 and -1.496
f″=(2/3)-cos x

نکته: هر چند نقاط B و C مینیمم فراگیرند ولی از آنجا که تابع کراندار نیست نمی‌توانیم فراگیر بودن آن‌ها را تضمین کنیم.

FIGURE 4-11 Graph of $f(x)=(x^2/3)+\cos x$ of Example 4.22.

1. x*=0; f″=-1/3<0, so this is a local maximum with f(0)=1.
2. x*=1.496; f″=0.592>0, so this is a local minimum with f(1.496)=0.821.
3. x*=-1.496; f″=0.592>0, so this is a local minimum with f(-1.496)=0.821.

شرایط بهینگی توابع چندمتغیره

برای حالت کلی یک تابع چندمتغیره $f(x)$ که x یک بردار n بعدی است.

$$f(x) = f(x^*) + \nabla f(x^*)^T d + \frac{1}{2} d^T H(x^*) d + R$$

یا، تغییر در تابع عبارت است از:

$$\Delta f(x) = \nabla f(x^*)^T d + \frac{1}{2} d^T H(x^*) d + R$$

اگر یک مینیمم محلی در x^* فرض کنیم آن گاه Δf باید نامنفی باشد

$$\Delta f \geq 0 \Rightarrow \nabla f(x^*) = 0 \qquad (3.33)$$

$$\nabla f(x^*) = 0 \Rightarrow \frac{\partial f(x^*)}{\partial x_i} = 0; \quad i = 1 \text{ to } n \qquad (3.35)$$

نقاطی که معادله (3.35) را برآورده می‌کنند نقاط ایستا نامیده می‌شوند.

شرط کافی

با توجه به جمله دوم معادله (3.33) که در نقطه ایستا ارزیابی می‌شود،

$$\Delta f \geq 0 \Rightarrow d^T H(x^*) d > 0 \qquad (3.36)$$

برای برآورده شدن معادله (3.36) باید هسیان $H(x^*)$ یک ماتریس معین مثبت باشد که این شرط کافی برای یک مینیمم محلی برای تابع $f(x)$ در نقطه x^* است.

قضیه (۴-۳): شرط لازم و کافی برای مینیمم محلی (4.4)

شرط لازم: اگر $f(x)$ یک مینیمم محلی در x^* داشته باشد.

$$\frac{\partial f(x^*)}{\partial x_i} = 0; \quad i = 1 \text{ to } n$$

شرط لازم مرتبه دو: اگر $f(x)$ در x^* یک مینیمم محلی داشته باشد، آن گاه ماتریس هسیان

$$H(x^*) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right]_{(n\times n)}$$

در نقطه x^* نیمه معین مثبت یا معین مثبت است.

شرط کافی مرتبه دو: اگر ماتریس $H(x^*)$ در نقطه ایستا x^* معین مثبت باشد، آن گاه برای تابع $f(x)$ یک مینیمم محلی است.

مثال ۴-۲۰: مینیمم تابع دومتغیره زیر را بیابید.

$$f(x) = x_1^2 + 2x_1x_2 + 2x_2^2 - 2x_1 + x_2 + 8$$

به دست آوردن نقطه ایستا:

$$\frac{\partial f}{\partial x} = \begin{bmatrix} 2x_1 + 2x_2 - 2 \\ 2x_1 + 4x_2 + 1 \end{bmatrix} = 0 \implies x^* = (x_1^*, x_2^*) = (5/2, -3/2)$$

از هر کدام از قضایای تشخیص نوع ماتریس متقارن به این نتیجه می‌رسیم که H معین مثبت است. بنابراین x^* یک مینیمم محلی برای تابع است.

$$H = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix} \qquad f(x^*) = 4.75$$

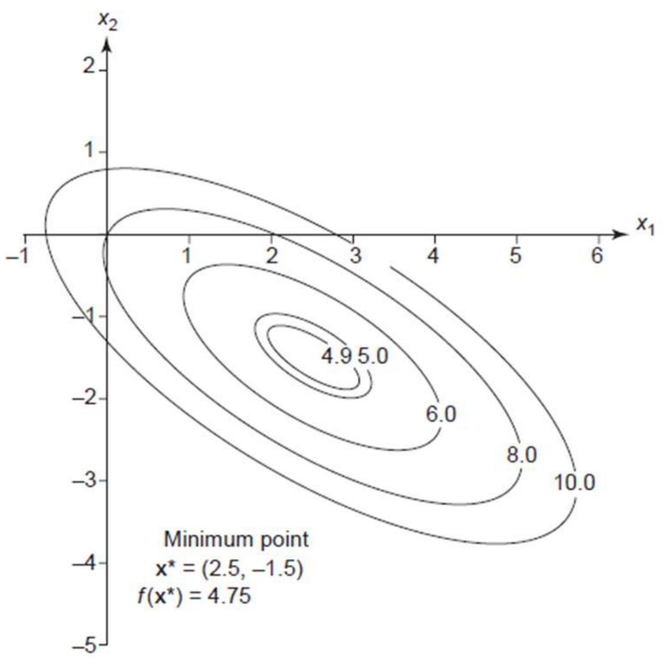
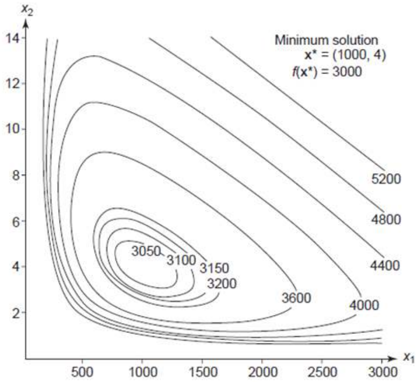


FIGURE 4-10 Isocost curves for the function of Example 4.20.



مثال ۲۳-۴:

$$f(x) = x_1 + \frac{(4*10^6)}{x_1 x_2} + 250x_2$$

The necessary conditions for optimality are

$$\frac{\partial f}{\partial x_1} = 0;\ 1 - \frac{(4*10^6)}{x_1^2 x_2} = 0$$

$$\frac{\partial f}{\partial x_2} = 0;\ 250 - \frac{(4*10^6)}{x_1 x_2^2} = 0$$

x_1^*=1000, and x_2^*=4 is a stationary point

$$H(1000,4) = \frac{(4*10^6)}{(4000)^2}\begin{bmatrix} 0.008 & 1 \\ 1 & 500 \end{bmatrix}$$

Since both eigenvalues are positive, the Hessian of *f(x)* at the point *x** is positive definite. Therefore, *x**=(1000,4) is a local minimum point with *f(x*)*=3000.

TABLE 4-1 Optimality Conditions for Unconstrained Problems

Function of one variable minimize f(x)	*Function of several variables minimize f(**x**)*
First-order necessary condition: $f' = 0$. Any point satisfying this condition is called a stationary point; it can be a local minimum, local maximum, or neither of the two (inflection point)	*First-order necessary condition*: $\nabla f = 0$. Any point satisfying this condition is called a stationary point; it can be a local minimum, local maximum, or neither of the two (inflection point)
Second-order necessary condition for a local minimum: $f'' \geq 0$	*Second-order necessary condition* for a local minimum: **H** must be at least positive semidefinite
Second-order necessary condition for a local maximum: $f'' \leq 0$	*Second-order necessary condition* for a local maximum: **H** must be at least negative semidefinite
Second-order sufficient condition for a local minimum: $f'' > 0$	*Second-order sufficient condition* for a local minimum: **H** must be positive definite
Second-order sufficient condition for a local maximum: $f'' < 0$	*Second-order sufficient condition* for a local maximum: **H** must be negative definite
Higher-order necessary conditions for a local minimum or local maximum: Calculate a higher ordered derivative that is not zero; all odd-ordered derivatives below this one must be zero	
Higher-order sufficient condition for a local minimum: Highest nonzero derivative must be even-ordered and positive	

مسئله زیر را حل کرده و تا دو هفته دیگر تحویل فرمایید:

3) 17

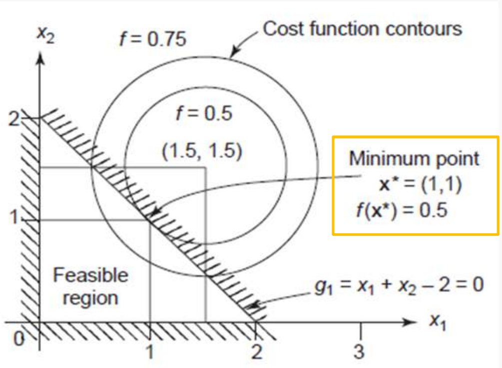
۳-۴ شرایط بهینگی مسایل مقید:

Role of Constraints

Constrained Optimum Point

minimize $f(x)=(x_1-1.5)^2+(x_2-1.5)^2$

subject to $g_1=x_1+x_2-2 \leq 0,\ g_2=-x_1 \leq 0,\ g_3=-x_2 \leq 0$



Unconstrained Optimum Point for a Constrained Problem (1.5, 1.5)

شرایط لازم

مقدمه:

نقاط منظم (Regular Point)

نقطه منظم : گرادیان توابع قیود فعال در آن‌ها مستقل خطی است. یعنی: در آن نقاط هیچ دو گرادیان قیدی با هم موازی و یا ترکیب خطی از یکدیگر نیستند.

Linear independence of the given set of vectors can be determined form the matrix A of dimension $n \times k$ whose columns are the given vectors. Then, if rank r is equal to k (r=k), the given set is linearly independent; otherwise it is dependent.

ابتدا مسایل مقید با قیدهای مساوی را در نظر می‌گیریم:

$$\min f(X) = f(x_1, x_2, \cdots, x_n)$$

$$h_i(x) = 0, \quad \text{i=1 to p}$$

مثال ۳.۲۴ یا ۴.۲۷:

$$\min f(x_1, x_2) = (x_1 - 1.5)^2 + (x_2 - 1.5)^2$$

$$h(x_1, x_2) = x_1 + x_2 - 2 = 0$$

فرض می‌کنیم بتوان یک متغیر را بر حسب متغیر دیگر به دست آورد.

$$x_2 = \varphi(x_1)$$

$$x_2 = \varphi(x_1) = -x_1 + 2$$

$$\min f(x_1, \varphi(x_1)) = (x_1 - 1.5)^2 + (-x_1 + 2 - 1.5)^2$$

$$\frac{df}{dx_1} = 0 \rightarrow x_1^* = 1 \quad x_2^* = 1$$

پس x^* یک نامزد مینیمم محلی است.

شرط کافی: $\frac{d^2 f}{dx_1^2} > 0$ برآورده می‌شود. پس x^* یک مینیمم محلی است.

$$\min f(x_1, \phi(x_1))$$

$$\frac{df\,(x_1, x_2)}{dx_1} = \frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2} \cdot \frac{dx_2}{dx_1} = 0$$

$$\rightarrow \frac{\partial f\,(x_1^*, x_2^*)}{\partial x_1} + \frac{\partial f\,(x_1^*, x_2^*)}{\partial x_2} \cdot \frac{d\varphi}{dx_1} = 0$$

$$h(x_1, x_2) = 0$$

φ

$$\frac{dh}{dx_1} = \frac{\partial h}{\partial x_1} + \frac{\partial h}{\partial x_2} \cdot \frac{dx_2}{dx_1} = 0$$

$$\rightarrow \frac{d\phi}{dx_1} = \frac{-\partial h(x_1^*, x_2^*) / \partial x_1}{\partial h(x_1^*, x_2^*) / \partial x_2}$$

$$\frac{\partial f(x_1^*, x_2^*)}{\partial x_1} + \frac{\partial f(x_1^*, x_2^*)}{\partial x_2} \left(\frac{-\dfrac{\partial h(x_1^*, x_2^*)}{\partial x_1}}{\dfrac{\partial h(x_1^*, x_2^*)}{\partial x_2}} \right) = 0$$

$$v = \frac{-\partial f\,(x_1^*, x_2^*) / \partial x_2}{\partial h\,(x_1^*, x_2^*) / \partial x_2}$$

ضریب لاگرانژ

$$\begin{cases} \dfrac{\partial f\,(x_1^*, x_2^*)}{\partial x_1} + v\,\dfrac{\partial h(x_1^*, x_2^*)}{\partial x_1} = 0 \\ \dfrac{\partial f\,(x_1^*, x_2^*)}{\partial x_2} + v\,\dfrac{\partial h(x_1^*, x_2^*)}{\partial x_2} = 0 \\ h(x_1, x_2) = 0 \end{cases}$$

شرایط لازم

در این مسئله:

$$\begin{cases} 2(x_1 - 1.5) + v = 0 \\ 2(x_2 - 1.5) + v = 0 \\ x_1 + x_2 - 2 = 0 \end{cases} \qquad x_1^* = x_2^* = 1, v^* = 1$$

ضریب لاگرانژ قیدهای مساوی از نظر علامت آزادند.

معنی هندسی ضریب لاگرانژ

$$L(x_1, x_2, V) = f(x_1, x_2) + vh(x_1, x_2)$$

$$\left.\begin{array}{l} \dfrac{\partial L(x_1^*, x_2^*)}{\partial x_1} = 0 \\ \dfrac{\partial L(x_1^*, x_2^*)}{\partial x_2} = 0 \end{array}\right\} \Rightarrow \nabla L(x_1^*, x_2^*) = 0$$

$$\text{or } \nabla f(X^*) + v\nabla h(X^*) = 0$$

$$\nabla h = \begin{bmatrix} \dfrac{\partial h(x_1^*, x_2^*)}{\partial x_1} \\ \dfrac{\partial h(x_1^*, x_2^*)}{\partial x_2} \end{bmatrix} \qquad \nabla f(X^*) = \begin{bmatrix} \dfrac{\partial f(x_1^*, x_2^*)}{\partial x_1} \\ \dfrac{\partial f(x_1^*, x_2^*)}{\partial x_2} \end{bmatrix}$$

$$\nabla f(X^*) = -v\nabla h(X^*)$$

At the candidate minimum point, gradient of the cost and constraint functions are along the same line and proportional to each other; the Lagrange multiplier *v* is the proportionality constant.

ضریب لاگرانژ: معنی هندسی

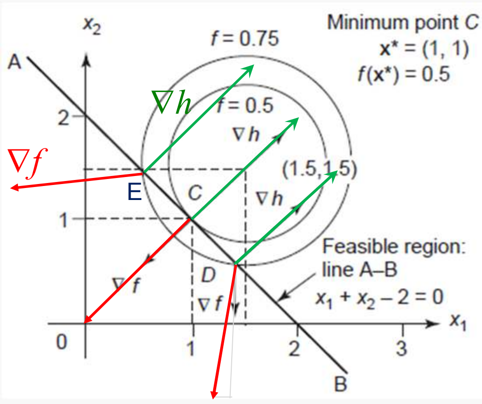


FIGURE 4-19 Graphical solution for Example 4.27. Geometrical interpretation of necessary conditions.

$$f(x_1,x_2)=(x_1-1.5)^2+(x_2-1.5)^2$$

$$h(x_1,x_2)=x_1+x_2-2=0$$

$$x^*=(1,1):\text{point C} \qquad f(x^*)=0.5$$

$$\nabla f\ (1,1)=\begin{bmatrix}-1\\-1\end{bmatrix} \qquad \nabla h=\begin{bmatrix}1\\1\end{bmatrix}$$

Nonoptimum point

$$E\,(0.64,1.35) \qquad f\,(E)=0.75$$

$$\nabla f=\begin{bmatrix}-1.7\\-0.29\end{bmatrix} \qquad \nabla h=\begin{bmatrix}1\\1\end{bmatrix}$$

معنی فیزیکی ضریب لاگرانژ

ضریب لاگرانژ به منزله نیرویی برای تحمیل قید می‌باشد.

$$L(x_1,x_2,V)=f(x_1,x_2)+vh(x_1,x_2)$$

We will see later that: The larger the value of the Lagrange multiplier, the higher is the dividend (bonus) to relax the constraint, or the higher is the penalty to tighten the constraint.

شرط کافی برای مینیمم محلی:

هرحرکت قابل قبول از ***C*** منجر به افزایش تابع هزینه می‌گردد و هر حرکت در جهت کاهش تابع هزینه باعث نقض قید می‌شود.

قضیه ضریب لاگرانژ: اگر مسئله مینیمم سازی ***f(X)*** را مشروط به قیود $h_i(x)=0,\ i=1\text{ to }p$ در نظر بگیریم، و ***X**** یک نقطه منظم باشد که برای مسئله یک مینیمم محلی است، آنگاه ضرایب لاگرانژ $v_j,\ j=1\text{ to }p$ وجود خواهند داشت به شکلی که:

$$\frac{\partial f\,(X^*)}{\partial x_i}+\sum_{j=1}^{p}v^*_j\,\frac{\partial h_j\,(X^*)}{\partial x_i}=0; \qquad i=1\text{ to }n \qquad h_j\,(X^*)=0;\ \ j=1\text{ to }p$$

It is convenient to write these conditions in terms of a Lagrange function defined as

$$L(X,V) = f(X) + \sum_{j=1}^{p} v_j h_j(X) = f(X) + V^T h(X)$$

$$\nabla L(X^*,V^*) = 0 \begin{cases} \dfrac{\partial L\left(X^*,V^*\right)}{\partial x_i} = 0; \quad i = 1 \text{ to } n \\ \dfrac{\partial L(X^*,V^*)}{\partial v_j} \equiv h_j(X^*) = 0; \quad j = 1 \text{ to } p \end{cases}$$

3.4.3 Necessary Conditions: Inequality Constraints-Karush-Kuhn-Tucker (KKT) Conditions

قضیه (3.6): شرایط لازم کان- تاکر (K-T)

فرض کنید x^* یک نقطه منظم از مجموعه قید باشد که یک مینیمم محلی برای $f(x)$ است به شرط قیود

$$h_i(x) = 0; \quad i = 1 \text{ to } p$$

$$g_i(x) \le 0; \quad i = 1 \text{ to } m$$

$g_i(x)+s_i=0$, where $s_i \geq 0$ is a slack variable or $g_i+s_i^2=0$

تابع لاگرانژ را برای مسئله به شکل زیر تعریف می‌کنیم.

$$L(x,v,u,s) = f(x) + \sum_{i=1}^{p} v_i h_i(x) + \sum_{i=1}^{m} u_i \left(g_i(x) + s_i^2\right)$$

$$= f(x) + \mathbf{v}^T \mathbf{h}(x) + \mathbf{u}^T \left(\mathbf{g}(x) + \mathbf{s}^2\right)$$

آن‌گاه ضرایب لاگرانژ $\mathbf{v}^*$ (یک بردار p تایی) و $\mathbf{u}^*$ (یک بردار m تایی) وجود خواهند داشت که لاگرانژین نسبت به s_i, u_i, v_i, x_j ایستاست، یعنی

$$\frac{\partial L}{\partial x_j} \equiv \frac{\partial f}{\partial x_j} + \sum_{i=1}^{p} v_i^* \frac{\partial h_i}{\partial x_j} + \sum_{i=1}^{m} u_i^* \frac{\partial g_i}{\partial x_j} = 0; \quad j=1 \text{ to } n \qquad (3.47)$$

$$h_i(x^*) = 0; \qquad i=1 \text{ to } p$$

$$g_i(x^*) + s_i^2 = 0; \qquad i=1 \text{ to } m$$

$$u_i^* s_i = 0; \qquad i=1 \text{ to } m$$

$$u_i^* \geq 0; \qquad i = 1 \text{ to } m$$

معنی هندسی شرایط کان- تاکر

معادله (3.47) را دو باره به شکل زیر می‌نویسیم:

$$-\frac{\partial f}{\partial x_j} = \sum_{i=1}^{p} v_i^* \frac{\partial h_i}{\partial x_j} + \sum_{i=1}^{m} u_i^* \frac{\partial g_i}{\partial x_j}; \quad j = 1 \text{ to } n$$

که نشان می‌دهد در نقطه ایستا، جهت منفی گرادیان (جهت تندترین کاهش) برای تابع هزینه ترکیب خطی گرادیان‌های قیود است.

شرایط سوئیچی

شرط $u_i^* s_i = 0;$
به عنوان شرایط سویچی (switching) یا شرایط کمبود مکمل خوانده می‌شوند.

تعداد 2^m حالت معمول برای m شرط سویچی وجود دارد که عبارتند از:

- الف: $s_i = 0$

کمبود صفر نشان دهنده فعال بودن نامساوی است؛ یعنی:

$$g_i = 0$$

- ب: $u_i = 0$

در این صورت باید $g_i \le 0$ باشد تا قابل قبول بودن را برآورده کند زیرا g+s²=0.

- ج: حالت غیر معمول:

در این حالت هم $u_i = 0$ و هم $s_i = 0$
از آنجا که با رها سازی قید انتظار داریم مقدار تابع هزینه کاهش یافته و یا حداقل ثابت بماند، این حالت را غیر معمول می‌گوییم؛ زیرا مفهوم آن این است که تابع هزینه با تغییرات طرف راست قید همیشه بی‌تغییر می‌ماند:

$$L = f + u(g + S^2) \xrightarrow{\text{Let } g=e \,\&\, s=0} = f + u(e)$$

$$\frac{\partial L}{\partial e} = \frac{\partial f}{\partial e} + u = 0$$

$$u = -\frac{\partial f}{\partial e} \xrightarrow{\text{But for } u=0} -\frac{\partial f}{\partial e} = 0$$

تعداد شرایط سویچی برابر تعداد قیود نامساوی مسئله است.

روند استفاده از شرایط (K-T)

مجهولات مسئله *V,S,U,X* است. این بردارها *p,m,m,n* بعدی‌اند. پس *(n+2m+p)* متغیر مجهول وجود دارد و ما به همین تعداد معادله برای به دست آوردن آن‌ها نیاز داریم. این معادلات مورد نیاز از شرایط (K-T) در دسترسند.

- ابتدا حالت‌های معمول شرایط سویچی را مشخص می‌کنیم.
- سپس با اعمال شرایط (K-T) در مورد تک‌تک شرایط سویچی نقاط نامزد مینیمم را به دست می‌آوریم. توجه شود که در این مرحله نقاطی نامزد نقطه مینیمم می‌شوند که تمام شرایط (K-T) را برآورده کنند.

چند نکته در مورد اعمال شرایط کان-تاکر

- محاسبه s_i در حقیقت محاسبه تابع $g_i(x)$ است زیرا $s_i^2=-g_i(x)$
- اگر قید نامساوی $g_i(x)$ در نقطه نامزد مینیمم غیر فعال باشد (یعنی $g_i(x^*)<0$ ، یا $s_i^2>0$)، آن‌گاه باید ضریب لاگرانژ مربوطه صفر باشد $u_i^*=0$ تا شرط سویچی برآورده شود.
- اگر قید فعال باشد (یعنی $g_i(x^*)=0$)، آن‌گاه ضرایب لاگرانژ باید نامنفی، $u_i^*\geq 0$ ، باشند

- در استفاده از شرایط کان-تاکر باید به این نکته توجه داشت که این شرایط فقط در مورد نقاط منظم برقرار است. یعنی اگر نقطه منظمی شرایط کان-تاکر را برقرار نکند نامزد مینیمم نیست مگر اینکه نامنظم باشد.
- نقاطی که شرایط را برآورده کنند نقاط کان-تاکر نامیده می‌شوند.
- نقاطی که شرایط کان-تاکر را برآورده می‌کنند ممکن است مقید یا نامقید باشند. وقتی هیچ قید مساوی وجود نداشته باشد و همه قیود نامساوی غیر فعال باشند به آن‌ها نامقید می‌گویند.

- اگر نقطه نامزد نامقید باشد، بسته به ماتریس هسیان تابع هزینه، ممکن است نقطه مینیمم محلی، ماکزیمم محلی یا نقطه عطف باشد.
- اگر قیدهای مساوی وجود داشته باشند و قیود نامساوی جملگی غیر فعال باشند (یعنی $u=0$)، آن‌گاه نقاطی که شرایط کان-تاکر را برآورده می‌کنند فقط ایستا هستند. آن‌ها می‌توانند نقاط ماکزیمم یا مینیمم یا عطف باشند.

- اگر بعضی از قیود نامساوی فعالند و ضرایب آن‌ها مثبت است،

آن‌گاه نقاطی که شرایط کان-تاکر را برآورده می‌کنند نمی‌توانند نقاط ماکزیمم محلی برای تابع هزینه باشند؛

زیرا این شرط تضمین می‌کند که کاهش بیشتر تابع هزینه در آن نقطه، تنها هنگامی میسر است که به ناحیه غیر قابل قبول قید وارد شویم.

(البته اگر قیود نامساوی فعال ضرایب صفر داشته باشند ممکن است نقاط ماکزیمم محلی باشند). آن‌ها ممکن است مینیمم محلی هم نباشند، که این امر به شرایط لازم و کافی مرتبه دو بستگی دارد.

- مهم است که به مقدار ضریب لاگرانژ هر قید که به شکل تابع قید بستگی دارد توجه شود. جواب بهین مسئله با تغییر شکل قید تغییر نمی‌کند ولی ضریب لاگرانژ متفاوت خواهد بود. سه شکل قید مقابل، سه ضریب متفاوت دارند.

$x/y-10\le 0, \quad (y>0)$

$x-10y\le 0$

$0.1x/y-1\le 0$

نقاط K-T یا

مقیدند یا

قید مساوی وجود دارد و قیود نامساوی جملگی غیر فعالند (یعنی $u=0$)

نقاط فقط ایستایند و می‌توانند ماکزیمم یا مینیمم یا عطف باشند.

بعضی قیود نامساوی فعالند (یعنی $u > 0$)

نقاط ماکزیمم نیستند. مینیمم بودن آن‌ها به شرایط لازم و کافی مرتبه دوم بستگی دارد.

نامقیدند

نقاط ایستایند، یعنی بسته به نوع H یا مینیمم محلی یا ماکزیمم محلی یا نقاط عطفند.

Example 4.30 **Write KKT necessary conditions and solve them for the problem:**

$$\min\ f(x) = \frac{1}{3}x^3 - \frac{1}{2}(b+c)x^2 + bcx + f_0$$

subject to $a \le x \le d$ where $0 < a < b < c < d$ and f_0 are specified constants

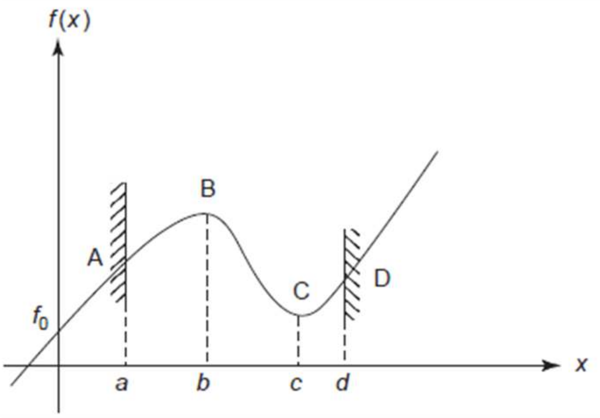


FIGURE 4-20 Graphical representation for Example 4.30. Point A, constrained local minimum; B, unconstrained local maximum; C, unconstrained local minimum; D, constrained local maximum.

$g_1 = a - x \le 0; \quad g_2 = x - d \le 0$

$$L = \frac{1}{3}x^3 - \frac{1}{2}(b+c)x^2 + bcx + f_0 + u_1(a - x + s_1^2) + u_2(x - d + s_2^2)$$

73/94

$$\frac{\partial L}{\partial x} = x^2 - (b+c)x + bc - u_1 + u_2 = 0 \qquad (c)$$

$$(a - x) + s_1^2 = 0,\ s_1^2 \ge 0;\ (x - d) + s_2^2 = 0,\ s_2^2 \ge 0 \qquad (d)$$

$$u_1 s_1 = 0,\ u_2 s_2 = 0 \qquad (e)$$

$$u_1 \ge 0,\ u_2 \ge 0 \qquad (f)$$

4 normal cases:

Case 1: $u_1=0$, $u_2=0$.

Eq. (c) gives two solutions as x=b and x=c.

for x=b: $s_1^2=b-a>0$; $s_2^2=d-b>0$ (g)

for x=c: $s_1^2=c-a>0$; $s_2^2=d-c>0$ (h)

Both points are candidate of minimum points

$x = b; \quad \frac{d^2f}{dx^2} = 2x - (b+c) = b - c < 0$ a local maximum point

$x = c; \quad \frac{d^2f}{dx^2} = c - b > 0$ a local minimum point

74/94

~~**Case 2: u_1=0, s_2=0.**~~

$\cancel{\text{from (d)} \rightarrow x = d}$

$\cancel{u_2 = -[d^2 - (b+c)d + bc] = -(d-c)(d-b)}$

$\cancel{\text{Since } d > c > b,\ u_2 \text{ is } < 0}$ The KKT necessary condition is violated

Case 3: s_1=0, u_2=0.

↓

x = a is a candidate minimum point

$$u_1 = a^2 - (b+c)a + bc = (a-b)(a-c) > 0$$

~~**Case 4: s_1=0, s_2=0.**~~

~~does not give any valid solution since x cannot be simultaneously equal to a and d.~~

Example 4.31

minimize $f(x)=x_1^2+x_2^2-3x_1x_2$

subject to $g=x_1^2+x_2^2-6\le 0$.

$L=x_1^2+x_2^2-3x_1x_2 +u(x_1^2+x_2^2-6+s^2)$

$$\frac{\partial L}{\partial x_1} = 2x_1 - 3x_2 + 2ux_1 = 0$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - 3x_1 + 2ux_2 = 0$$

$x_1^2+x_2^2-6+s^2 =0$, $s^2 \ge 0$, $u \ge 0$

$us=0$

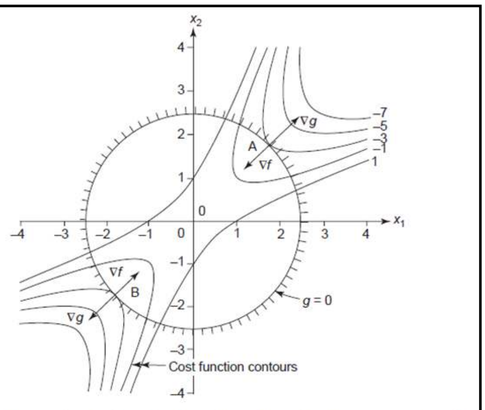


FIGURE 4-21 Graphical solution for Example 4.31. Local minimum points, A and B.

3 possible cases:

Case 1: u=0.

$$2x_1 - 3x_2 = 0;\ 2x_2 - 3x_1 = 0$$

$$x_1{}^* = 0, x_2{}^* = 0; f(0,0) = 0 \qquad s^2=6$$

candidate minimum point, point O in the Fig.

Case 2: s=0.

$$x_1 = x_2 = \sqrt{3}, u = \frac{1}{2} \qquad x_1 = x_2 = -\sqrt{3}, u = \frac{1}{2}$$

candidate minimum point, point A and B in the Fig.

$$\cancel{x_1 = -x_2 = \sqrt{3}, u = -\frac{5}{2}} \qquad \cancel{x_1 = -x_2 = -\sqrt{3}, u = -\frac{5}{2}}$$

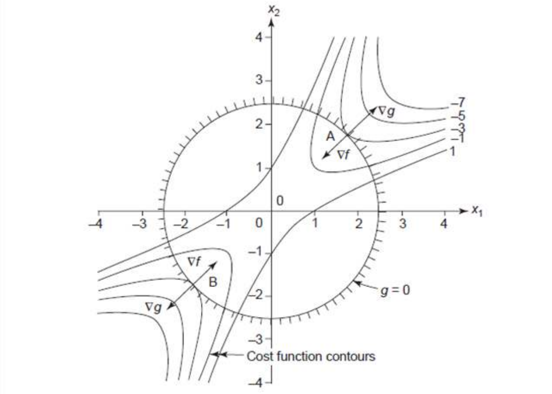


FIGURE 4-21 Graphical solution for Example 4.31. Local minimum points, A and B.

~~**Case 3: u=0, s=0.**~~

$s^2 = 6 \neq 0$ KKT conditions cannot be satisfied

1. $x_1^* = 0, x_2^* = 0, u^* = 0, f(0,0) = 0,$ Point O in Fig. 4-21

2. $x_1^* = x_2^* = \sqrt{3}, u^* = \frac{1}{2}, f(\sqrt{3}, \sqrt{3}) = -3,$ Point A in Fig. 4-21

3. $x_1^* = x_2^* = -\sqrt{3}, u^* = \frac{1}{2}, f(-\sqrt{3}, -\sqrt{3}) = -3,$ Point B in Fig. 4-21

77/94

Example 4.32

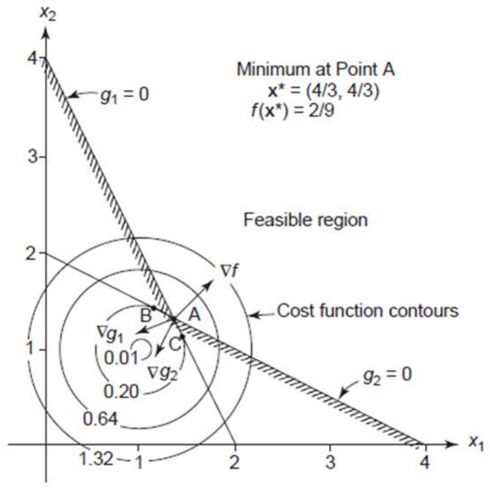


FIGURE 4-22 Graphical solution for Example 4.32.

minimize $f(x_1,x_2)=x_1^2+x_2^2-2x_1-2x_2+2$

subject to $g_1=-2x_1-x_2+4 \leq 0.$

$g_2=-x_1-2x_2+4 \leq 0.$

$L=x_1^2+x_2^2-2x_1-2x_2+2+u_1(-2x_1-x_2+4+s_1^2)+u_2(-x_1-2x_2+4+s_2^2)$

78/94

$$\frac{\partial L}{\partial x_1} = 2x_1 - 2 - 2u_1 - u_2 = 0 \qquad (b)$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - 2 - u_1 - 2u_2 = 0 \qquad (c)$$

$$g_1 = -2x_1 - x_2 + 4 + s_1^2 = 0;\ s_1^2 \ge 0, u_1 \ge 0, \quad (d)$$

$$g_2 = -x_1 - 2x_2 + 4 + s_2^2 = 0;\ s_2^2 \ge 0, u_2 \ge 0, \quad (e)$$

$$u_i s_i = 0,\ \text{i=1,2} \qquad (f)$$

۴ حالت معمول وجود دارد:
(حالت‌هایی که هم *u* و هم *s* صفر باشند در نظر گرفته نمی‌شوند):

1. $u_1 = 0,\quad u_2 = 0$
2. $u_1 = 0,\quad s_2 = 0 (g_2 = 0)$
3. $s_1 = 0 (g_1 = 0),\quad u_2 = 0$
4. $s_1 = 0 (g_1 = 0),\quad s_2 = 0 (g_2 = 0)$

Case 1: $u_1 = 0$, $u_2 = 0$.

Equations (b) and (c) give $x_l=x_2=1$. This is not a valid solution as it gives $s_1^2=-1(g_1=1)$, $s_2^2=-1(g_2=1)$ from Eqs. (d) and (e), which implies that both inequalities are violated, and so $x_1=1$ and $x_2=1$ is not a feasible design.

79/94

Case 2: $u_1=0$, $s_2=0$.

With these conditions, Eqs. (b), (c), and (e) become

$$2x_1-2-u_2=0,\ 2x_2-2-2u_2=0,\ -x_1-2x_2+4=0, \qquad (g)$$

Using the elimination procedure, we obtain $x_1=1.2$, $x_2=1.4$, and $u_2=0.4$. Therefore, the solution for this case is

$$x_1=1.2,\ x_2=1.4,\ u_1=0,\ u_2=0.4,\ f=0.2 \qquad (h)$$

We need to check for feasibility of the design point with respect to constraint g_1 before it can be claimed as a candidate local minimum point. Substituting $x_1=1.2$ and $x_2=1.4$ into Eq. (d), we find that $s_1^2=-0.2<0$ ($g_1=0.2$), which is a violation of constraint g_1.

point B, which is not in the feasible set.

گردآوری و تنظیم: محمدحسین ابوالبشری 80/94

Case 3: $s_1=0$, $u_2=0$.

With these conditions Eqs. (b), (c), and (d) give

$$2x_1-2-2u_1=0,\ 2x_2-2-u_1=0,\ -2x_1-x_2+4=0, \quad \text{(i)}$$

This is again a linear system of equations for the variables x_1, x_2, and u_1. Solving the system, we get the solution as

$$x_1=1.4,\ x_2=1.2,\ u_1=0.4,\ u_2=0,\ f=0.2 \quad \text{(j)}$$

Checking the design for feasibility with respect to constraint g_2, we find from Eq. (e) $s_2^2=-0.2<0$ ($g_2=0.2$). This is not a feasible design. Therefore, Case 3 also does not give any candidate local minimum point.

point C, which is not in the feasible set.

Case 4: $s_1=0$, $s_2=0$.

For this case, Eqs. (b) to (e) must be solved for the four unknowns x_1, x_2, u_1, and u_2. This system of equations is again linear and can be solved easily. Using the elimination procedure as before, we obtain $x_1=4/3$ and $x_2=4/3$ from Eqs. (d) and (e). Solving for u_1 and u_2 from Eqs. (b) and (c), we get $u_1=2/9>0$ and $u_2=2/9>0$. To check regularity condition for the point, we evaluate the gradients of the active constraints and define the constraint gradient matrix A as

$$\nabla g_1=\begin{bmatrix}-2\\-1\end{bmatrix},\nabla g_2=\begin{bmatrix}-1\\-2\end{bmatrix},\ A=\begin{bmatrix}-2 & -1\\-1 & -2\end{bmatrix} \quad (k)$$

Since rank (A)=no. of active constraints, the gradients ∇g_1 and ∇g_2 are linearly independent.

Thus, all the KKT conditions are satisfied and the preceding solution is a candidate local minimum point. The solution corresponds to point A in Fig. 4-22. The cost function at the point has a value of 2.

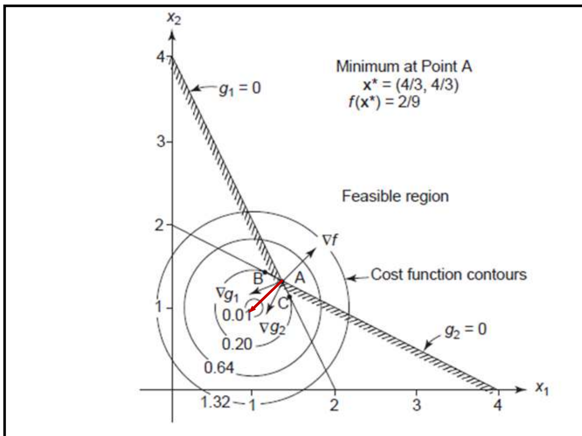


FIGURE 4-22 Graphical solution for Example 4.32.

It can be observed in Fig. 4-22 that the vector $-\nabla f$ can be expressed as a linear combination of the vectors ∇g_1 and ∇g_2 at point *A*. This satisfies the necessary condition of $-\frac{\partial f}{\partial x_j} = \sum_{i=1}^{p} v_i^* \frac{\partial h_i}{\partial x_j} + \sum_{i=1}^{m} u_i^* \frac{\partial g_i}{\partial x_j};\ \text{j=1 to n}$. It can also be seen from the figure that point *A* is indeed a local minimum because any further reduction in the cost function is possible only if we go into the infeasible region. Any feasible move from point *A* results in an increase in the cost function.

83/94

In summary, the following points should be noted regarding KKT first-order necessary conditions

1. The conditions can be used to check whether a given point is a candidate minimum; it must be feasible, the gradient of the Lagrangian with respect to the design variables must be zero, and the Lagrange multipliers for inequality constraints must be nonnegative.
2. For a given problem, the conditions can be used to find candidate minimum points. Several cases defined by the switching conditions must be considered and solved. Each case can give multiple solutions.

84/94

3. For each solution case, remember to
(i) check all inequality constraints for feasibility (i.e., $g_i \leq 0$ or $s_i^2 \geq 0$)
(ii) calculate all the Lagrange multipliers
(iii) ensure that the Lagrange multipliers for all the inequality constraints are nonnegative

۱-۳-۴-۳ شکل جایگزین شرایط کان- تاکر

شرایط لازم کان- تاکر را می‌توان بدون متغیر کمبود نوشت

تابع لاگرانژ را برای مسئله به شکل زیر تعریف می‌کنیم.

$$L(x, v, u, s) = f(x) + \sum_{i=1}^{p} v_i h_i(x) + \sum_{i=1}^{m} u_i (g_i(x) \boxed{+ s_i^2})$$

$$= f(x) + \mathbf{v}^T \mathbf{h}(x) + \mathbf{u}^T (\mathbf{g}(x) \boxed{+ \mathbf{s}^2})$$

$$\frac{\partial L}{\partial xj} \equiv \frac{\partial f}{\partial x_j} + \sum_{i=1}^{p} v_i^* \frac{\partial h_i}{\partial x_j} + \sum_{i=1}^{m} u_i^* \frac{\partial g_i}{\partial x_j} = 0; \text{j=1 to n} \quad (3.47)$$

$$h_i(x^*) = 0; \qquad \text{i= 1 to p}$$

$$g_i(x^*) + \boxed{s_i^2} = 0; \text{i=1 to m} \qquad \Rightarrow g_i(x^*) \le 0$$

$$u_i^* \boxed{s_i} = 0; \qquad \text{i=1 to m} \qquad \Rightarrow u_i^* g_i(x^*) = 0$$

$$u_i^* \ge 0; \qquad \text{i=1 to m}$$

TABLE 5-1 Alternate Form of KKT Necessary Conditions

Problem: minimize $f(\mathbf{x})$ subject to $h_i(\mathbf{x}) = 0$, $i = 1$ to p; $g_j(\mathbf{x}) \le 0$, $j = 1$ to m

1. Lagrangian function definition

$$L = f + \sum_{i=1}^{p} v_i h_i + \sum_{j=1}^{m} u_j g_j \qquad (5.1)$$

2. Gradient conditions

$$\frac{\partial L}{\partial x_k} = 0; \quad \frac{\partial f}{\partial x_k} + \sum_{i=1}^{p} v_i^* \frac{\partial h_i}{\partial x_k} + \sum_{j=1}^{m} u_j^* \frac{\partial g_j}{\partial x_k} = 0; \quad k = 1 \text{ to } n \qquad (5.2)$$

3. Feasibility check

$$g_j(\mathbf{x}^*) \le 0; \quad j = 1 \text{ to } m \qquad (5.3)$$

4. Switching conditions

$$u_j^* g_j(\mathbf{x}^*) = 0; \quad j = 1 \text{ to } m \qquad (5.4)$$

5. Nonnegativity of Lagrange multipliers for inequalities

$$u_j^* \ge 0; \quad j = 1 \text{ to } m \qquad (5.5)$$

6. Regularity check
   Gradients of active constraints must be linearly independent. In such a case, the Lagrange multipliers for the constraints are unique.

Use of Alternate form of KKT Conditions

minimize $f(x,y)=(x-10)^2+(y-8)^2$ subject to $g_1=x+y-12\le 0$, $g_2=x-8\le 0$.

1. Lagrangian function definition:

$L=(x-10)^2+(y-8)^2+u_1(x+y-12)+u_2(x-8)$.

2. Gradient condition:

$$\frac{\partial L}{\partial x}=2(x-10)+u_1+u_2=0$$

$$\frac{\partial L}{\partial y}=2(y-8)+u_1=0$$

3. Feasibility check: $g_1\le 0$, $g_2\le 0$
4. Switching conditions: $u_1g_1=0$, $u_2g_2=0$
5. Nonnegativity of Lagrange multipliers: $u_1,u_2\ge 0$
6. Regularity check.

4 normal cases:

1. $u_1=0$, $u_2=0$ (both g_1 and g_2 inactive)
2. $u_1=0$, $g_2=0$ (g_1 inactive, g_2 active)
3. $g_1=0$, $u_2=0$ (g_1 active, g_2 inactive)
4. $g_1=0$, $g_2=0$ (both g_1 and g_2 active)

Case 1: $u_1=0$, $u_2=0$ (both g_1 and g_2 inactive).
Equations (a) give the solution as, x=10, y=8. Checking feasibility of this point gives $g_1=6>0$, $g_2=2>0$; thus both constraints are violated and so this case does not give any candidate minimum point.

Case 2: $u_1=0$, $g_2=0$ (g_1 inactive, g_2 active).
$g_2=0$ gives x=8. Equations (a) give y=8 and $u_2=4$. At the point (8, 8), $g_1=4>0$ which is a violation. Thus the point (8,8) is infeasible and this case also does not give any candidate minimum points.

Case 3: $g_1=0$, $u_2=0$ (g_1 active, g_2 inactive).
Equations (a) and $g_1=0$ give x=7, y=5, $u_1=6>0$. Checking feasibility, $g_2=-1<0$ which is satisfied. Since there is only one active constraint, the question of linear dependence of gradients of active constraints does not arise; therefore regularity condition is satisfied. Thus point (7, 5) satisfies all the KKT necessary conditions.

Case 4: $g_1=0$, $g_2=0$ (both g_1 and g_2 active).

$g_1=0$, $g_2=0$ give $x=8$, $y=4$. Equations (a) give $u_1=8$, $u_2=-4<0$, which is a violation of the necessary conditions. Therefore, this case also does not give any candidate minimum points.

It may be checked that this is a convex programming problem since constraints are linear and the cost function is convex. Therefore, the point obtained in Case 3 is indeed a global minimum point according to the convexity results of next Section.

Check for KKT Conditions at Irregular Points

minimize $f(x_1,x_2)=x_1^2+x_2^2-4x_1+4$

subject to $g_1=-x_1\le 0$, $g_2=-x_2\le 0$, $g_3= x_2-(1-x_1)^3\le 0$

$L=x_1^2+x_2^2-4x_1+4+u_1(-x_1)+u_2(-x_2)+u_3(x_2-[1-x_1]^3)$

$$\frac{\partial L}{\partial x_1}=2x_1-4-u_1+u_3(3)(1-x_1)^2=0$$

$$\frac{\partial L}{\partial x_2}=2x_2-u_2+u_3=0 \qquad \text{(a)}$$

$g_i\le 0$, $i=1,2,3$ (b)

$u_i g_i=0$, $i=1,2,3$ (c)

$u_i\ge 0$, $i=1,2,3$

At $x^*=(1,0)$ the first constraint (g_1) is inactive and the second and third constraints are active. The switching conditions (c) identify the case as $u_1=0$, $g_2=0$, $g_3=0$.

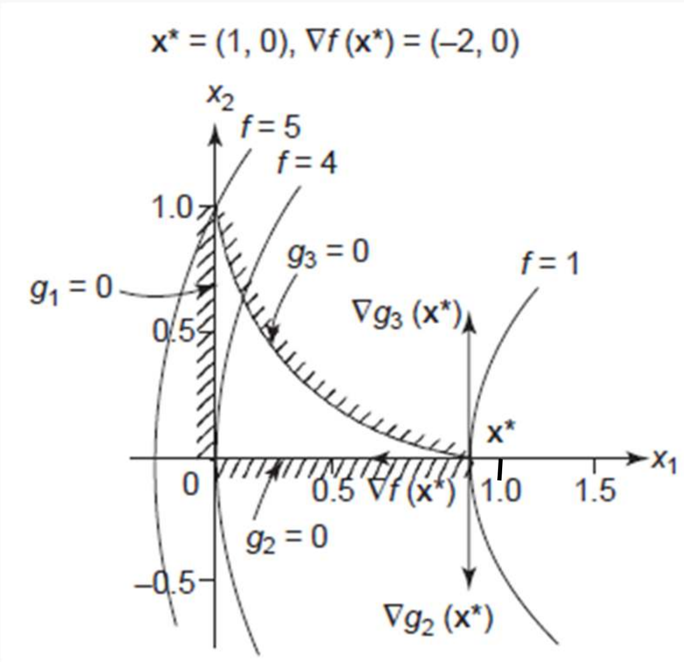
Substituting the solution into Eq. (a), we find that the first equation gives $-2=0$ and therefore it is not satisfied. Thus, KKT necessary conditions are not satisfied at the minimum point.

This apparent contradiction can be resolved by checking the regularity condition at the minimum point x*=(1,0). The gradients of the active constraints g_2 and g_3 are given as

$$\nabla g_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \nabla g_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

These vectors are not linearly independent. They are along the same line but in opposite directions, as shown in Fig. 5-1. Thus x* is not a regular point of the feasible set.

Since this is assumed in the KKT conditions, their use is invalid here. Note also that geometrical interpretation of the KKT conditions is violated; that is, ∇f at (1,0) cannot be written as a linear combination of the gradients of the active constraints g_2 and g_3. Actually ∇f is normal to both ∇g_2 and ∇g_3 as shown in the figure.



مسائل زیر را حل کرده و تا دو هفته دیگر تحویل فرمایید:

3) 29, 30, 46

3) 1,8,16,17 تمرین‌های قبلی