Optimum Design

Chapter 4: Optimality Conditions

Minseok Seo
Artificial Intelligence Graduate School
Gwangju Institute of Science and Technology (GIST)

July 12, 2025

Overview

1. Definitions of Global and Local Minima

2. Review of Some Basic Caculus Concepts

4.1 Design of Global and Local Minima

- The design optimization problem is always converted to minimization of a cost function subject to equality and inequality constraints.
- The optimization problem is to find a point in the feasible design space that gives a minimum value to the cost function.

4.1.1 Minimum/Maximum

Global Minimum

A function $f(x)$ of n variables has a global minimum at x^* if the value of the function at x^* is less than or equal to the value of the function at any other point x in the feasible set S.

$$f(x^*) \leq f(x)$$

for all x in the feasible set S.
If strict inequality holds for all x other than x^*, in Eq. 4, then x^* is called a strict global minimum; otherwise, it is called a weak global minimum.

4.1.1 Minimum/Maximum

Local Minimum

A function $f(x)$ on n variables has a local minimum at x^* if inequality in Eq. 4 holds for all x in a neighborhood N (vicinity) of x^* in the feasible set S.
If strict inequality holds, then x^* is called a strict local minimum; otherwise, it is called a weak local minimum.

The neighborhood N of point x^* is defined as a set of points in its vicinity that is,

$$N = \{\mathbf{x} \mid \mathbf{x} \in \mathcal{S} \text{ with } \|\mathbf{x} - \mathbf{x}^*\| < \delta\}$$

for some small $\delta > 0$.

4.1.2 Existence of a Minimum

Theorem (Weierstrass Theorem - Existence of a 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. 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 of its boundary points and every sequence of points has subsequence that converges to a point in the set.
- A set is bounded if for any point, $\mathbf{x} \in S, \mathbf{x}^T\mathbf{x} < c$, where c is a finite number.

It is important, however, to realize that when they are not satisfied, a global solution may still exist.

4.2.1 Gradient Vector: Partial Derivatives of a Function

Gradient vector

$$c = \nabla f(\mathbf{x}^*) = \begin{bmatrix} \frac{\partial f(\mathbf{x}^*)}{\partial \mathbf{x}_1} \\ \frac{\partial f(\mathbf{x}^*)}{\partial \mathbf{x}_2} \\ \vdots \\ \frac{\partial f(\mathbf{x}^*)}{\partial \mathbf{x}_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial f(\mathbf{x}^*)}{\partial x_1} & \frac{\partial f(\mathbf{x}^*)}{\partial x_2} & \dots & \frac{\partial f(\mathbf{x}^*)}{\partial x_n} \end{bmatrix}^T$$

4.2.2 Hessian Matrix: Second-Order Partial Derivatives

Hessian Matrix

$$\frac{\partial^2 f}{\partial \mathbf{x} \partial \mathbf{x}} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} s & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Since $f(\mathbf{x})$ is assumed to be twice continuously differentiable, the cross-partial derivatives are equal;

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

4.2.3 Taylor's Expansion

Using Taylor's expansion, a function can be approximated by polynomials in a neighborhood of any point in terms of its value and derivatives.

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

where R is the remainder term that is smaller in magnitude than the previous terms if x is sufficiently close to x^*.
If we let $x - x^* = d$ (a small change in the point x^*), then the Taylor's expansion becomes a quadratic polynomial in d:

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

4.2.3 Taylor's Expansion

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

$$f(x_1, x_2) = f(x_1^*, x_2^*) + \frac{\partial f}{\partial x_1} d_1 + \frac{\partial f}{\partial x_2} d_2 + \frac{1}{2}\left[\frac{\partial^2 f}{\partial x_1^2} d_1^2 + 2\frac{\partial^2 f}{\partial x_1 \partial x_2} d_1 d_2 + \frac{\partial^2 f}{\partial x_2^2} d_2^2\right]$$

where $d_1 = x_1 - x_1^*, d_2 = x_2 - x_2^*$, and all partial derivatives are calculated at the given point (x_1^*, x_2^*). The remainder term R and the arguments of these partial derivatives $f(x_1^*, x_2^*)$ are omitted for notational compactness.
Also, we can be written using the summation notation as

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

4.2.3 Taylor's Expansion

Taylor's expansion can also be written in matrix notiation as

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

where $\mathbf{x} = (x_1, x_2), \mathbf{x}^* = (x_1^*, x_2^*), \mathbf{x} - \mathbf{x}^* = \mathbf{d}$, and $\mathbf{H}$ is the 2×2 Hessian matrix.

Taylor's expansion can be generalized to functions of n variables. In that case, $\mathbf{x}, \mathbf{x}^*$, and ∇f are n-dimensional vectors and $\mathbf{H}$ is the $n \times n$ Hessian matrix.
Defining the changes as $\Delta f = f(\mathbf{x}) - f(\mathbf{x}^*)$, Eq gives:

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

A first-order change in $f(\mathbf{x})$ at $\mathbf{x}^*$ (denoted as δf) is obtained by retaininig only the first term in the above equation:

$$\delta f = \nabla f^T \delta \mathbf{x} = \nabla f \cdot \delta \mathbf{x} \tag{1}$$

4.2.4 Quadratic Forms and Definite Matrices

Quadratic Form

- The quadratic form is a special nonlinear function having only second-order terms (either the square of a variable or the product of two varialbes).
- $F(\mathbf{x}) = x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_2 - 2x_2x_3 + 4x_3x_1$

Generalizing the quadratic form of n variables, we can write it in the double summation notation as:

$$F(\mathbf{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} x_i x_j$$

The quadratic form can be written in the matrix notation. Let $\mathbf{P} = [p_{ij}]$ be an $n \times n$ matrix and $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ be an n-dimensional vector.
Then the quadratic form can be written as:

$$F(\mathbf{x}) = \mathbf{x}^T \mathbf{P} \mathbf{x}$$

$\mathbf{P}$ is called the matrix of the quadratic form $F(\mathbf{x})$. $p_{ij} + p_{ji} =$ the coefficient of x_{ij}.

4.2.4 Quadratic Forms and Definite Matrices

We can obtain the symmetric matrix **A** for the quadratic form by using the asymmetric matrix **P** as follows:

$$\mathbf{A} = \frac{1}{2}(\mathbf{P} + \mathbf{P}^T) \quad \text{or} \quad a_{ij} = \frac{1}{2}(p_{ij} + p_{ji}, \quad i, j = 1, 2, \text{ to } , n)$$

The matrix **P** is replaced by the symmetric matrix **A**, and the quadratic form can be written as:

$$F(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$$

4.2.4 Quadratic Forms and Definite Matrices

Form of a Matrix

Quadratic form $F(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ may be either positive, negative, or zero for any $\mathbf{x} \neq 0$. The following are the possible forms for the function $F(\mathbf{x})$ and the associated symmetric matrix $\mathbf{A}$:

1. Positive definite. $F(\mathbf{x}) > 0$ for all $\mathbf{x} \neq 0$. The matrix $\mathbf{A}$ is called positive definite.
2. Positive semidefinite. $F(\mathbf{x}) \geq 0$ for all $\mathbf{x} \neq 0$. The matrix $\mathbf{A}$ is called positive semidefinite.
3. Negative definite. $F(\mathbf{x}) < 0$ for all $\mathbf{x} \neq 0$. The matrix $\mathbf{A}$ is called negative definite.
4. Negative semidefinite. $F(\mathbf{x}) \leq 0$ for all $\mathbf{x} \neq 0$. The matrix $\mathbf{A}$ is called negative semidefinite.
5. Indefinite. The quadratic form is called indefinite if it is positive for some values of $\mathbf{x}$ and negative for some others. In that case, matrix $\mathbf{A}$ is called indefinite.

4.2.4 Quadratic Forms and Definite Matrices

Theorem (Eigenvalue Check for the Form of a Matrix)

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

1. $F(\mathbf{x})$ is positive definite if and only if all eigenvalues of $\mathbf{A}$ are strictly positive that is, $\lambda_i > 0, i = 1$ to n.

2. $F(\mathbf{x})$ is positive semidefinite if and only if all eigenvalues of $\mathbf{A}$ are nonnegative that is, $\lambda_i \geq 0, i = 1$ to n. (note that at least one eigenvalue must be zero for it to be called positive semidefinite)

3. $F(\mathbf{x})$ is negative definite if and only if all eigenvalues of $\mathbf{A}$ are strictly negative that is, $\lambda_i < 0, i = 1$ to n.

4. $F(\mathbf{x})$ is negative semidefinite if and only if all eigenvalues of $\mathbf{A}$ are nonpositive that is, $\lambda_i \leq 0, i = 1$ to n. (note that at least one eigenvalue must be zero for it to be called negative semidefinite)

5. $F(\mathbf{x})$ is indefinite if some $\lambda_i < 0$ and some other $\lambda_j > 0$.

4.2.4 Quadratic Forms and Definite Matrices

Theorem (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 $\mathbf{A}$ defined as the determinant of a $k \times k$ submatrix obtained by deleting the last $(n-k)$ rows and columns of $\mathbf{A}$. Assume that no two consecutive principal minors are zero.

1. $\mathbf{A}$ *is positive definite if and only if all* $M_k > 0, k = 1$ *to* n.
2. $\mathbf{A}$ *is positive semidefinite if and only if* $M_k > 0, k = 1$ *to* r, *where* $r < n$ *is the rank of* $\mathbf{A}$.
3. $\mathbf{A}$ *is negative definite if and only if* $M_k < 0$ *for* k *odd and* $M_k > 0$ *for* k *even,* $k = 1$ *to* n.
4. $\mathbf{A}$ *is negative semidefinite if and only if* $M_k \leq 0$ *for* k *odd and* $M_k \geq 0$ *for* k *even,* $k = 1$ *to* n.
5. $\mathbf{A}$ *is indefinite if it does not satisfy any of the preceding criteria.*

4.2.4 Quadratic Forms and Definite Matrices

Differentiation of a Quadratic Form

- $\frac{\partial F(\mathbf{x})}{\partial x_i} = 2\sum_{j=1}^{n} a_{ij}x_i$ or $\nabla F(\mathbf{x}) = 2\mathbf{A}\mathbf{x}$
- $\frac{\partial^2 F(\mathbf{x})}{\partial x_j \partial x_i} = 2a_{ij}$ or $\mathbf{H} = 2\mathbf{A}$