Optimal Design : Interior Optima

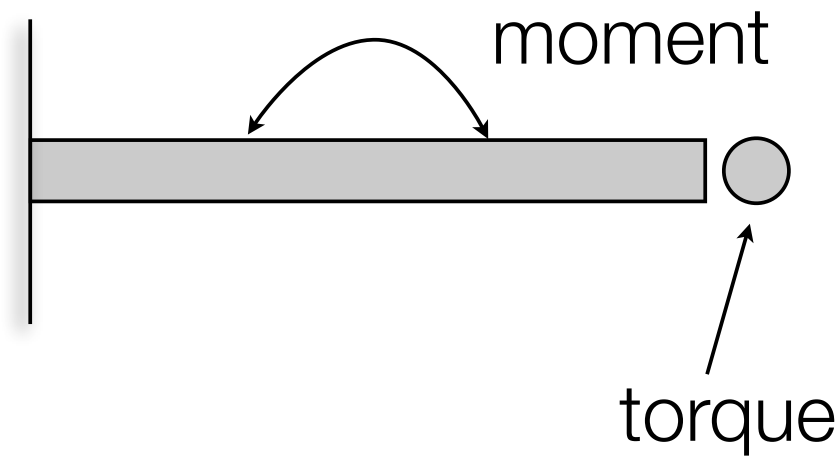
ME 7129

Kuei-Yuan Chan
National Taiwan University

Assuming all active inequality and equality constraints have been identified and are used to eliminate design DOFs, the remaining problem will be an **unconstrained optimization problem**.

Example 4.1

- Consider the following example



Unconstrained Problem

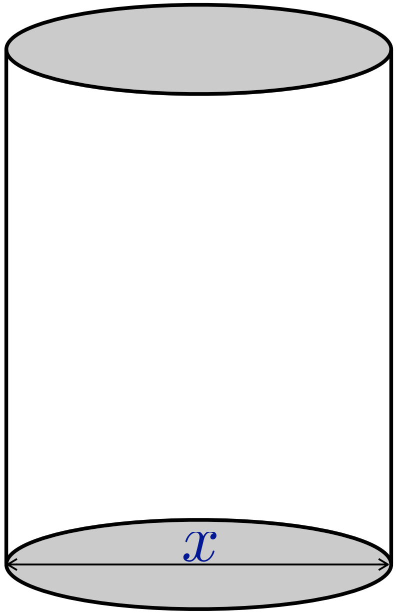
$$\text{minimize } f(\mathbf{x})$$
$$\text{subject to } \mathbf{x} \in \mathcal{X} \subseteq \Re^n$$

If $\mathcal{X}$ is an *open set* and a solution exists, that solution will be an **interior optimum**.

Weierstrass Theorem

A continuous function defined on a closed finite interval attains its maximum and minimum *in that interval*

Example - Cylindrical Refrigeration Tank



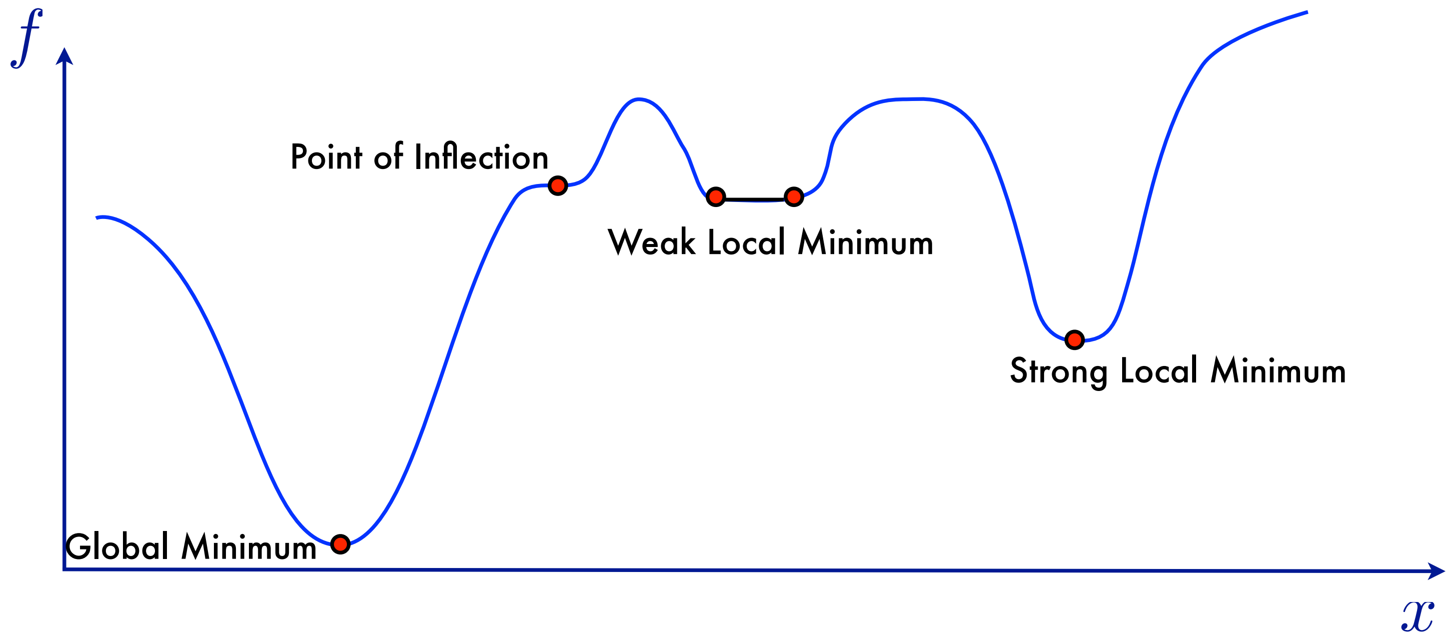
- Determine the objective function for building a minimum-cost cylindrical refrigeration tank of volume 50 m^3, if the circular ends cost \$10 per m^2, the cylinder wall costs \$6 per mm^2, and it cost \$80 per m^2 to refrigerate over the useful life.

- Total cost

$$f = (10)(2)\left(\frac{\pi x^2}{4}\right) + (6)(\pi x L) + 80\left(2 \cdot \frac{\pi x^2}{4} + \pi x L\right)$$

$$L = \frac{V}{\pi x^2/4}$$

Types of Minima



- Weak local minimum : if there exists a $\delta > 0$ such that $f(x^*) \leq f(x) \ \forall |x - x^*| < \delta$
- Strong local minimum : if there exists a $\delta > 0$ such that $f(x^*) < f(x) \ \forall |x - x^*| < \delta$
- Global minimum : if $f(x^*) < f(x) \ \forall x$

Multivariate Problems

- The Taylor series expansions can be extended to problems with several variables.
- Define the Euclidean norm as $\|\mathbf{x}\| \triangleq \left(\sum_{i=1}^{n} x_i^2 \right)$
- The Taylor series linear and quadratic approximations are now

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + \sum_{i=1}^{n} \frac{\partial f(\mathbf{x}_0)}{\partial x_i}(x_i - x_{i,0}) + o(\|\mathbf{x} - \mathbf{x}_0\|)$$

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + \sum_{i=1}^{n} \frac{\partial f(\mathbf{x}_0)}{\partial x_i}(x_i - x_{i,0}) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f(\mathbf{x}_0)}{\partial x_i \partial x_j}(x_i - x_{i,0})(x_j - x_{j,0}) + o(\|\mathbf{x} - \mathbf{x}_0\|^2)$$

Gradient and Hessian

- The gradient vector is the row vector of the first partial derivatives of f

$$\bigtriangledown f \triangleq (\partial f/\partial x_1 \partial f/\partial x_2, \cdots, \partial f/\partial x_n)$$

- The Hessian matrix is the square, symmetric matrix of the second derivatives of f

$$\mathbf{H} \triangleq \begin{pmatrix} \partial^2 f/\partial x_1^2 & \cdots & \partial^2 f/\partial x_1 \partial x_n \\ \vdots & & \vdots \\ \partial^2 f/\partial x_n \partial x_1 & \cdots & \partial^2 f/\partial x_n^2 \end{pmatrix}$$

Function Perturbation

- Define the perturbation vector $\partial \mathbf{x} \triangleq \mathbf{x} - \mathbf{x}_0$, the resulting function perturbation

$$\partial f \triangleq f(\mathbf{x}) - f(\mathbf{x}_0)$$

- Rewrite the perturbation using linear and quadratic expansions via compact form as

$$\partial f = f(\mathbf{x}_0)\partial \mathbf{x} + o(\|\partial \mathbf{x}\|)$$

$$\partial f = f(\mathbf{x}_0)\partial \mathbf{x} + \frac{1}{2}\partial \mathbf{x}^T \mathbf{H}(\mathbf{x}_0)\partial \mathbf{x} + o(\|\partial \mathbf{x}\|^2)$$

Example 4.2

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

Necessary vs. Sufficient Conditions

- Necessary : conditions that have to be satisfied for the point to be optimal. In other words, if a point does not satisfy these conditions, it cannot be a minimum.

- Sufficient : conditions that can well describe a point being optimal.

First Order Necessary Conditions

- Let the function $f(x)$ be $\mathcal{C}^1$ continuous, locally any perturbations must result in higher values of the function.

- The first order Taylor series expansions about x^*

$$f(x) \approx f(x^*) + \left.\frac{\partial f}{\partial x}\right|_{x=x^*} (x - x^*)$$

- For a small real number $h > 0$, we then have

$$f(x^* + h) = f(x^*) + hf'(x^*)$$

$$f(x^* - h) = f(x^*) - hf'(x^*)$$

- Local minimum requires that for small h

$$\left.\begin{aligned} f(x^* + h) > f(x^*) \\ f(x^* - h) > f(x^*) \end{aligned}\right. \longrightarrow f' = 0$$

ME7129
K.-Y. Chan

If $f(\mathbf{x}), \mathbf{x} \in \mathcal{X}$, has a local minimum at an interior point $\mathbf{x}_*$ of the set $\mathcal{X}$ and if $f(\mathbf{x})$ is continuously differentiable at $\mathbf{x}_*$ then

$$\nabla f(\mathbf{x}_*) = \mathbf{0}^T$$

Second Order Sufficiency Conditions

- Sufficient conditions can be derived using the second order Taylor series expansions.

- If $\mathbf{x}_\dagger$ is a stationary point of f , then

$$\partial f_\dagger = \frac{1}{2}\partial \mathbf{x}^T \mathbf{H}(\mathbf{x}_\dagger)\partial \mathbf{x}$$

- If the Hessian matrix of f is positive-definite at a stationary point $\mathbf{x}_\dagger$, then $\mathbf{x}_\dagger$ is a local minimum.

Example 4.6

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

Positive-Definite Matrix Tests

- A square, symmetric matrix is positive definite if and only if ***any*** of the following is true

  1. *All it eigenvalues are positive.*
  2. *All determinants of its leading principal minors are positive.*
  3. *All the pivots are positive when the matrix is reduced to row-echelon form, working symmetrically along the main diagonal.*

Example 4.7

$$f = -4x_1 + 2x_2 + 4x_1^2 - 4x_1x_2 + x_2^2$$

Taxonomy

Quadratic Form	**Hessian Matrix**	**Nature of x**
positive	positive-definite	local minimum
negative	negative-definite	local maximum
non-negative	positive-semidefinite	probable valley
non-positive	negative-semidefinite	probable ridge
any sign	indefinite	saddle point

Example 4.8

Convexity

- Positive second derivative means that the function must have increasing slopes, the **curvature** must be positive.

- Positive curvature can also be expressed geometrically such that a line tangent at any point of $f(x_2)$ will never cross the graph of the function.

- A line connecting any two points on the graph will be entirely above the graph of the function.

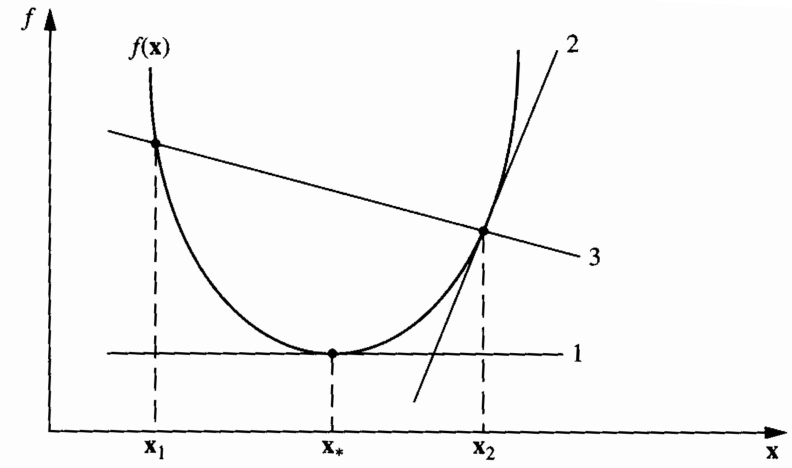


Figure 4.4. Geometric meaning of convexity.

Convex

A set $\mathcal{S} \subseteq \Re^n$ is convex if, for every point $\mathbf{x}_1, \mathbf{x}_2$ in $\mathcal{S}$, the point

$$\mathbf{x}(\lambda) = \lambda\mathbf{x}_2 + (1 - \lambda)\mathbf{x}_1, \ 0 \leq \lambda \leq 1$$

belongs also to the set.

A differentiable function is convex if and only if its Hessian is positive-semidefinite in its entire convex domain.

Examples

$$f = x_1^2 + 2x_2^2 + 3x_3^2 + 3x_1x_2 + 4x_1x_3 - 3x_2x_3$$

Unimodal Function

- A function is said to be unimodal in an interval if the function monotonically decreases as we approach the local minimum and then monotonically increase as we leave the minimum.

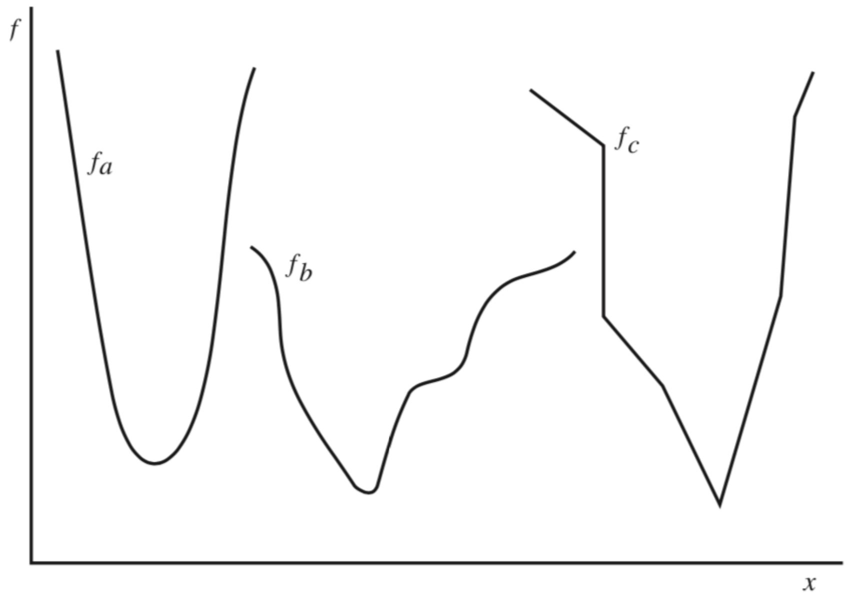


Figure 6.2. Unimodality.