Optimization problems – basic notations

Unconstrained optimization

First order condition for optimality

Second order conditions for optimality

Notation reminder

$$f(x) = f(x_1, x_2, \ldots, x_n) \qquad (f : \mathbb{R}^n \to \mathbb{R})$$

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix} \quad \nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \frac{\partial f}{\partial x_n^2}(x) \end{pmatrix}$$

Gradient vector

Hessian matrix

First order condition

Theorem. **(First Order Necessary Condition for (Local) Optimality)**
If x^* is an unconstrained local minimizer of a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, then we must have:

$$\nabla f(x^*) = 0$$


Fermat
(1607-1665)

Remarks

- This condition is necessary but not sufficient for local optimality.

Remarks

- This condition is necessary but not sufficient for local optimality.

- Nevertheless, it is useful because any local minimum must satisfy this condition. So, we can look for local (or global) minima only among points that make the gradient of the objective function vanish.

Remarks

- This condition is necessary but not sufficient for local optimality.
- Nevertheless, it is useful because any local minimum must satisfy this condition. So, we can look for local (or global) minima only among points that make the gradient of the objective function vanish.
- Terminology: A point x that satisfies $\nabla f(x) = 0$ is called a **stationary point** or a **critical point** of f. stationary point or a critical point of

Optimization problems – basic notations

Unconstrained optimization

First order condition for optimality

Second order conditions for optimality

Linear algebra review

Symmetric matrix: $A = A^T$

Theorem. Eigenvalues of a real symmetric matrix are real.

A square matrix A is said to be

- Positive semidefinite (psd) if:

- Positive definite (pd) if:

Linear algebra review

Recall that when we talk of psd or pd, we can assume that our matrix is symmetric: If A was not symmetric, we can replace A by $\frac{A^T + A}{2}$ as:

$$x^T A x = x^T \left(\frac{A^T + A}{2} \right) x$$

Linear algebra review

Theorem. A matrix is positive semidefinite iff all its eigenvalues are nonnegative. A matrix is positive definite iff all its eigenvalues are positive.

Examples:

1. $A = \begin{bmatrix} 2 & 4 \\ 4 & 5 \end{bmatrix}$
2. $A = \begin{bmatrix} 2 & 1 \\ 1 & 8 \end{bmatrix}$
3. $A = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$

Second order condition

Theorem. (Second Order Necessary Condition for (Local) Optimality)
If x^* is an unconstrained local minimizer of a twice continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, then, in addition to $\nabla f(x^*) = 0$, we must have:

$$\nabla^2 f(x^*) \succeq 0$$

(i.e., the Hessian at x^* is positive semidefinite.)

Second order condition

Theorem. (Second Order sufficient Condition for (Local) Optimality)
Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable, $\nabla f(x^*) = 0$ and

$$\nabla^2 f(x^*) \succ 0$$

(i.e. the Hessian at x^* is positive definite), then x^* is a strict local minimum of f.

Remarks

- $\nabla f(x^*) = 0$, $\nabla^2 f(x^*) \succeq 0$ is not sufficient for local optimality.

$$f(x) = x^3$$

- $\nabla^2 f(x) \succ 0$ is not necessary for (even strict global) optimality

$$f(x) = x^4$$

What are the questions in practice?

- How would we use all these optimality conditions to find local solutions and certify their optimality?

What are the questions in practice?

- How would we use all these optimality conditions to find local solutions and certify their optimality?

- Is it easy to find points satisfying these conditions? e.g., is it easy to solve $\nabla f(x) = 0$?

What are the questions in practice?

- How would we use all these optimality conditions to find local solutions and certify their optimality?
- Is it easy to find points satisfying these conditions? e.g., is it easy to solve $\nabla f(x) = 0$?
- Suppose you found that a given point is locally optimal, how would you go about checking if it is also globally optimal?

Exercise: State the analogues of the three theorems for local maxima.

Example 1: Least squares

Given: A $m \times n$ matrix (columns of A are independent)

b $m \times 1$ vector

Solve: $\min_{x \in \mathbb{R}^n} \|Ax - b\|^2$

Example 1: Least squares

Given: A $m \times n$ matrix (columns of A are independent)

b $m \times 1$ vector

Solve: $\min_{x \in \mathbb{R}^n} \|Ax - b\|^2$

Steps:

1. Write $f(x) = \|Ax - b\|^2$ as inner product of vectors
2. Solve $\nabla f(x) = 0$ for x
3. Show that $\nabla^2 f(x) = 2A^T A \succ 0$

Example 2

Find all the local minima and maxima of the following function:

$$f(x) = \frac{1}{2}x_1^2 + x_1x_2 + 2x_2^2 - 4x_1 - 4x_2 - x_2^3$$

$\nabla f(x) =$

$\nabla f(x) = 0 \Rightarrow$

$$f(x) = \frac{1}{2}x_1^2 + x_1x_2 + 2x_2^2 - 4x_1 - 4x_2 - x_2^3$$

$\nabla^2 f(x) =$

$x_2 = \qquad \Rightarrow \nabla^2 f(x) =$

$$f(x) = \frac{1}{2}x_1^2 + x_1x_2 + 2x_2^2 - 4x_1 - 4x_2 - x_2^3$$

$\nabla^2 f(x) =$

$x_2 = \qquad \Rightarrow \nabla^2 f(x) =$

Convex optimization I

Source: Ahmadi, ORF 363 slides.

Convex optimization

Convex sets

Intersection and Union of convex sets

General form of optimization

$$\min f(x)$$

$$\text{subject to } x \in \Omega$$

- We consider a very important special case of constrained optimization problems known as **convex optimization problems**

- For these problems,
  - f will be a **convex function**
  - Ω will be a **convex set**

Why convex optimization?

- Convex optimization problems are the broadest class of optimization problems that we know how to solve efficiently

Why convex optimization?

- Convex optimization problems are the broadest class of optimization problems that we know how to solve efficiently

- They have nice geometric properties;
  - e.g., a local minimum is automatically a global minimum

Why convex optimization?

- Convex optimization problems are the broadest class of optimization problems that we know how to solve efficiently

- They have nice geometric properties;
  - e.g., a local minimum is automatically a global minimum

- Numerous important optimization problems in engineering, operations research, machine learning, etc. are convex

Convex optimization

Convex sets

Intersection and Union of convex sets

Convex set

A set $\Omega \subseteq \mathbb{R}^n$ is **convex** if for all $x, y \in \Omega$, the line segment connecting x and y is also in Ω. In other words,

$$x, y \in \Omega, \lambda \in [0, 1] \Rightarrow \lambda x + (1 - \lambda) y \in \Omega$$

A point of the form $\lambda x + (1 - \lambda) y, \lambda \in [0, 1]$ is called a **convex combination** of x and y

Examples

Convex:

Not convex:

Common convex sets in optimization

- **Hyperplanes**: $\{x \mid a^T x = b\}$ $(a \in \mathbb{R}^n, b \in \mathbb{R})$

- **Halfspaces**: $\{x \mid a^T x \leq b\}$ $(a \in \mathbb{R}^n, b \in \mathbb{R})$

Common convex sets in optimization

- **Euclidean balls**: $\{x \mid \|x - x_c\| \leq r, x_c \text{ fixed}\}$ $(x, x_c \in \mathbb{R}^n, r \in \mathbb{R})$

- **Ellipsoids**: $\{x \mid (x - x_c)^T P (x - x_c) \leq r\}$ $(x, x_c \in \mathbb{R}^n, r \in \mathbb{R}, P \succ 0)$

Fancier convex sets

- The set of positive semidefinite matrices:

$$S_+^{n\times n} = \{P \in S^{n\times n} \mid P \succ 0\}$$

Proof:

Fancier convex sets

- The set of nonnegative polynomials in n variables.

$$\mathbb{P}_n = \{p(x) \text{ polynomial} \mid p(x) \geq 0, \forall x \in \mathbb{R}^n\}$$

Proof:

Convex optimization

Convex sets

Intersection and Union of convex sets

Intersections of convex sets

Intersection of two convex sets is convex:

$$\Omega_1 \text{ convex}, \Omega_2 \text{ convext} \Rightarrow \Omega_1 \cap \Omega_2 \text{ convex}$$

Proof:

Polyhedra

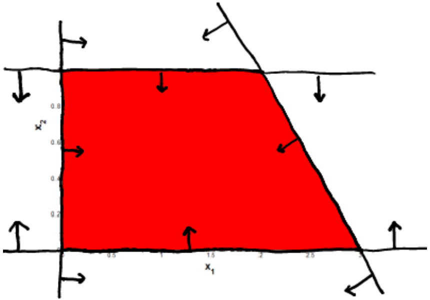
A polyhedron is the solution set of finitely many linear inequalities—ubiquitous in optimization theory

Polyhedra

A polyhedron is the solution set of finitely many linear inequalities—ubiquitous in optimization theory

Example: Buying two products: prices are 50 and 200 bahts

$x_1 =$ quantity of the first item bought
$x_2 =$ quantity of the second item bought



Polyhedra

- In matrix form:

$$\{x \mid Ax \leq b\}$$

  where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$

- These sets are convex: intersection of halfspaces $a_i^T x \leq b_i$ where a_i is the i-th row of A

Union of convex sets

Obviously, the union may not be convex: