OPTIMIZATION METHODS

- Instructor: Prof. Kwan-Ho You
- Lecture notes: http://www.icampus.skku.edu
- E-mail: khyou@skku.edu
- Phone: (031) 290 - 7148
- Textbook: *"An Introduction to Optimization"*
- Author: E.K. Chong and S.H. Zak
- Press: Wiley, 4th edition

Ch.1 Methods of Proof and Some Notation

1.1 Methods of Proof

- The statement "A $\Leftrightarrow$ B" reads "A if and only if B," or " A is equivalent to B," or "A is necessary and sufficient for B."
- To prove a statement (A $\Rightarrow$ B), three different techniques can be used.
  1. Direct method. (A $\Rightarrow$ B), ($\sim$ A $\cup$ B)
  2. Proof by contraposition. ($\sim$ B $\Rightarrow$ $\sim$A)
  3. Proof by contradiction. ($\sim$(A $\cap$ $\sim$B))
- Principle of induction

1.2 Notation

- $\{x : x \in \Re,\ x > 5\}$ reads "the set of all x such that x is real and x is greater than 5."
- If $X \subset Y$, X is a subset of Y.
- If $X \backslash Y$, X minus Y.
- $f :\ X \rightarrow Y$ means "f is a function from the set X into the set Y."

Ch.2 Vector Spaces and Matrices

2.1 Vector and Matrix

- Column n-vector is an array of n numbers as,

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

- $\Re$ is the set of real numbers, and $\Re^n$ is the set of column n-vectors with real components.
- $\Re^n$ an n-dimensional real vector space.
- Transpose of a given column vector $\mathbf{a}$ is $\mathbf{a^T}$.

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \quad \mathbf{a^T} = [a_1, a_2, \cdots, a_n].$$

- Two vectors $\mathbf{a} = [a_1, a_2, \cdots, a_n]^T$ and $\mathbf{b} = [b_1, b_2, \cdots, b_n]^T$ are equal if $a_i = b_i,\ i = 1, 2, \cdots, n$.
- A set of vectors $\{\mathbf{a_1}, \mathbf{a_2}, \cdots, \mathbf{a_k}\}$ is said to be linearly independent if the equality implies that all coefficients

$\alpha_i,\ i = 1, \cdots, k$ are equal to zero.

$$\alpha_1 \mathbf{a_1} + \alpha_2 \mathbf{a_2} + \cdots + \alpha_k \mathbf{a_k} = 0$$

- A vector $\mathbf{a}$ is said to be a linear combination of vectors $\mathbf{a_1}, \mathbf{a_2}, \cdots \mathbf{a_k}$ if there are scalars $\alpha_1, \cdots, \alpha_k$ such that

$$\mathbf{a} = \alpha_1 \mathbf{a_1} + \alpha_2 \mathbf{a_2} + \cdots + \alpha_k \mathbf{a_k}$$

Proposition 1 *A set of vectors* $\{\mathbf{a_1}, \mathbf{a_2}, \cdots, \mathbf{a_k}\}$ *is linearly dependent if and only if one of the vectors from the set is a linear combination of the remaining vectors.*

- A subset ν of $\Re^n$ is called a subspace of $\Re^n$ if ν is closed under the operations of vector addition and scalar multiplication.
- Span of $\mathbf{a_1}, \mathbf{a_2}, \cdots, \mathbf{a_k}$ is

$$span[\mathbf{a_1}, \mathbf{a_2}, \cdots, \mathbf{a_k}] = \left\{ \sum_{i=1}^{k} \alpha_i \mathbf{a_i} : \alpha_1, \cdots, \alpha_k \in \Re \right\}$$

- Given a subspace ν, any set of linearly independent vectors $\{\mathbf{a_1}, \mathbf{a_2}, \cdots, \mathbf{a_k}\} \subset \nu$ such that $\nu = span[\mathbf{a_1}, \mathbf{a_2}, \cdots, \mathbf{a_k}]$ is referred to as a basis of the subspace ν.
- All bases of a subspace ν contain the same number of vectors.
- This number is called the dimension of ν, denoted dim ν.

Proposition 2 *If* $\{\mathbf{a_1}, \mathbf{a_2}, \cdots, \mathbf{a_k}\}$ *is a basis of* ν*, then any vector* $\mathbf{a}$ *of* ν *can be represented uniquely as*

$$\mathbf{a} = \alpha_1 \mathbf{a_1} + \alpha_2 \mathbf{a_2} + \cdots + \alpha_k \mathbf{a_k},$$

where $\alpha_i \in \Re$, $i = 1, 2, \cdots, k$.

2.2 Rank of a Matrix

- A matrix with m rows and n columns is called an $m \times n$ matrix,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

- The maximal number of linearly independent columns of A is called the rank of the matrix A, denoted rank A.

Proposition 3 *The rank of a matrix A is invariant under the following operations:*

1. *Multiplication of the columns of A by nonzero scalars.*
2. *Interchange of the columns*
3. *Addition to a given column a linear combination of other columns*

- A matrix A is said to be square if the number of its rows is equal to the number of its columns ($n \times n$).

- Associated with each square ($n \times n$) matrix A is a scalar called the determinant of the matrix A, denoted det A or $|A|$.
- A pth-order minor of an $m \times n$ matrix A, with $p \leq min(m, n)$, is the determinant of a $p \times p$ matrix obtained from A by deleting $m - p$ rows and $n - p$ columns.

  Proposition 4 *If an $m \times n(m \geq n)$ matrix A has a nonzero nth-order minor, then the columns of A are linearly independent, that is, rank A=n.*
- The rank of a matrix is equal to the highest order of its nonzero minor.
- A nonsingular (or invertible) matrix is a square matrix whose determinant is nonzero.
- Matrix B is the inverse matrix of A, and write $B = A^{-1}$.
- Matrix A is symmetric if $A = A^T$.

2.3 Linear Equations

- Suppose we are given m equations in n unknowns of the form,

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$

$$
\begin{aligned}
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{aligned}
$$

- We can represent the above system of equations as

$$Ax = b$$

Theorem 1 *The system of equations $Ax = b$ has a solution if and only if*

$$rank\ A = rank[A,\ b].$$

Theorem 2 *Consider the equation $Ax = b$, where $A \in \Re^{m\times n}$, and rank A=m. A solution to $Ax = b$ can be obtained by assigning arbitrary values for $n - m$ variables and solving for the remaining ones.*

For example,(rank A=2)

$$\begin{pmatrix} x_1 + x_2 + x_3 = 1 \\ x_1 - x_2 - x_3 = 1 \end{pmatrix} \Rightarrow A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}.$$

2.4 Inner Products and Norms

- For $x, y \in \Re^n$, we define the Euclidean inner product by

$$< x, y >= \sum_{i=1}^{n} x_i y_i = x^T y$$

- The inner product is a real-valued function $< \cdot, \cdot >: \Re^n \times \Re^n \to \Re$ having the following properties:
  1. Positivity: $< x, x > \geq 0$, $< x, x >= 0$ iff $x = 0$.
  2. Symmetry: $< x, y >=< y, x >$
  3. Additivity: $< x + y, z >=< x, z > + < y, z >$
  4. Homogeneity: $< rx, y >= r < x, y >$ for every $r \in \Re$.
- The vector x and y are said to be orthogonal if $< x, y >= 0$.
- The Euclidean norm of a vector x is defined as

$$||x|| = \sqrt{< x, x >} = \sqrt{x^T x}$$

Theorem 3 Cauchy-Schwarz Inequality. *For any two vectors x and y in $\Re^n$, the Cauchy-Schwarz inequality*

$$| < x, y > | \leq ||x|| \, ||y||$$

holds. Futhermore, equality holds iff $x = \alpha y$ for some $\alpha \in \Re$.

- The Euclidean norm of a vector $||x||$ has the following properties:
  1. Positivity: $||x|| \geq 0$, $||x|| = 0$ iff $x = 0$.
  2. Homogeneity: $||rx|| = |r| \, ||x||$, $r \in \Re$
  3. Triangle Inequality: $||x + y|| \leq ||x|| + ||y||$
- The Euclidean norm is an example of a general vector

norm.($\equiv$ 2-norm, $||x||_2$).

- Other examples of vector norms on $\Re^n$ include the 1-norm, defined by $||x||_1 = |x_1| + \cdots + |x_n|$, and the ∞-norm, defined by $||x||_\infty = max_i |x_i|$.