Chapter 2 Vector Spaces and Matrices

An Introduction to Optimization

Spring, 2014

Wei-Ta Chu

Vectors and Matrices

- *n*-dimensional column vector and row vector

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \qquad \boldsymbol{a}^T = [a_1, a_2, \ldots, a_n]$$

- Properties

$$\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}$$

$$(\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c} = \boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c})$$

$$\boldsymbol{0} = (0, 0, \ldots, 0)$$

$$\boldsymbol{a} + \boldsymbol{0} = \boldsymbol{0} + \boldsymbol{a} = \boldsymbol{a}$$

$$\alpha(\boldsymbol{a} + \boldsymbol{b}) = \alpha\boldsymbol{a} + \alpha\boldsymbol{b}$$

$$\alpha(\beta\boldsymbol{a}) = (\alpha\beta)\boldsymbol{a}$$

$$1\boldsymbol{a} = \boldsymbol{a} \qquad (-1)\boldsymbol{a} = -\boldsymbol{a}$$

$$\alpha\boldsymbol{0} = \boldsymbol{0} = 0\boldsymbol{a}$$

Linearly Independent

- A set of vectors $\{\boldsymbol{a}_1, ..., \boldsymbol{a}_k\}$ is said to be ***linearly independent*** if the equality $\alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + ... + \alpha_k\boldsymbol{a}_k = \mathbf{0}$ implies that all coefficients $\alpha_i, i = 1, ..., k,$ are equal to zero.
- Any set of vectors containing the vector **0** is ***linearly dependent***.
- A set of vectors $\{\boldsymbol{a}_1, ..., \boldsymbol{a}_k\}$ is linearly dependent if and only if one of the vectors from the set is a ***linear combination*** of the remaining vectors.

Subspace

- A subset $\mathcal{V}$ of R^n is called a ***subspace*** of R^n if $\mathcal{V}$ is closed under addition and closed under scalar multiplication.
  - If $\boldsymbol{a}$ and $\boldsymbol{b}$ are vectors in $\mathcal{V}$, then the vectors $\boldsymbol{a}+\boldsymbol{b}$ and $\alpha\boldsymbol{a}$ are also in $\mathcal{V}$ for every scalar α.
- Let $\boldsymbol{a}_1, ..., \boldsymbol{a}_k$ be arbitrary vectors in R^n. The set of all their linear combinations is called the ***span*** of $\boldsymbol{a}_1, ..., \boldsymbol{a}_k$
- Given a subspace $\mathcal{V}$, any set of linearly independent vectors $\{\boldsymbol{a}_1, ..., \boldsymbol{a}_k\} \subset \mathcal{V}$ such that $\mathcal{V} = span[\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_k]$ is referred to as a ***basis*** of the subspace $\mathcal{V}$.
- All bases of a subspace $\mathcal{V}$ contain the same number of vectors. This number is called the ***dimension*** of $\mathcal{V}$.

Subspace

- If $\{\boldsymbol{a}_1, ..., \boldsymbol{a}_k\}$ is a basis of $\mathcal{V}$, then any vector $\boldsymbol{a}$ of $\mathcal{V}$ can be represented ***uniquely*** as $\boldsymbol{a} = \alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + ... + \alpha_k\boldsymbol{a}_k$, where $\alpha_i \in R, i = 1, 2, ..., k$
- The coefficients $\alpha_i, i = 1, 2, ..., k,$ are called the ***coordinates*** of $\boldsymbol{a}$ with respect to the basis $\{\boldsymbol{a}_1, ..., \boldsymbol{a}_k\}$

Rank of A Matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \qquad \boldsymbol{a}_k = \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix}$$

- Consider the $m \times n$ matrix
- The maximal number of linearly independent columns of *A* is called the ***rank*** of *A*, denoted by *rank*(*A*). The *rank*(*A*) is the dimension of $span[\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_k]$
- The rank of a matrix *A* is invariant under the following operations:
  - Multiplication of the columns of *A* by nonzero scalars.
  - Interchange of the columns.
  - Addition to a given column a linear combination of other columns.

Determinant

- The determinant of the matrix $A = [\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n]$ is a function of its columns and has the following properties:
  - $\det[\boldsymbol{a}_1, ..., \boldsymbol{a}_{k-1}, \alpha\boldsymbol{a}_k, ..., \boldsymbol{a}_n] = \alpha \det[\boldsymbol{a}_1, ..., \boldsymbol{a}_{k-1}, \boldsymbol{a}_k, ..., \boldsymbol{a}_n]$
  - If we have $\boldsymbol{a}_k = \boldsymbol{a}_{k+1}$, $\det(A) = 0$
  - Determinant of an identity matrix is 1.
  - If one of the columns is **0**, then the determinant is equal to zero.
  - The determinant does not change if we add to a column another column multiplied by a scalar.
  - The determinant changes its sign if we interchange columns.

Determinant

- 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
- If an $m \times n$ matrix *A* $(m \geq n)$ has a nonzero *n*th-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(s).
- A ***nonsingular*** (or ***invertible***) matrix is a square matrix whose determinant is nonzero.

Linear Equations

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1,$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2,$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m.$$

- Given *m* equations in *n* unknowns of the form
- We can represent the system of equations as $A\boldsymbol{x} = \boldsymbol{b}$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n] \qquad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \qquad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- The system of equations $A\boldsymbol{x} = \boldsymbol{b}$ has a solution if and only if *rank*(*A*) = *rank*([*A*, **b**])
- If *rank*(*A*)=*m*, a solution to $A\boldsymbol{x} = \boldsymbol{b}$ can be obtained by assigning arbitrary values for *n-m* variables and solving for the remaining ones.

General and Particular Solutions

- Theorem 4.7.2
  - If $\mathbf{x}_0$ denotes any single solution of a consistent linear system $A\mathbf{x} = \mathbf{b}$, and if $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ form a basis for the null space of A, (that is, the solution space of the homogeneous system $A\mathbf{x} = \mathbf{0}$), then every solution of $A\mathbf{x} = \mathbf{b}$ can be expressed in the form

$$\mathbf{x} = \mathbf{x}_0 + c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

  Conversely, for all choices of scalars $c_1, c_2, \ldots, c_k$, the vector $\mathbf{x}$ in this formula is a solution of $A\mathbf{x} = \mathbf{b}$.

Example (General Solution of $A\mathbf{x} = \mathbf{b}$)

- The solution to the nonhomogeneous system

$$\begin{aligned} x_1 + 3x_2 - 2x_3 \quad\quad + 2x_5 \quad\quad &= 0 \\ 2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 &= -1 \\ 5x_3 + 10x_4 \quad\quad + 15x_6 &= 5 \\ 2x_1 + 5x_2 \quad\quad + 8x_4 + 4x_5 + 18x_6 &= 6 \end{aligned}$$

is

$x_1 = -3r - 4s - 2t$, $x_2 = r$,
$x_3 = -2s$, $x_4 = s$,
$x_5 = t$, $x_6 = 1/3$

- The result can be written in vector form as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} -3r-4s-2t \\ r \\ -2s \\ s \\ t \\ 1/3 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1/3 \end{bmatrix}}_{\mathbf{x}_0} + \underbrace{r\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -4 \\ 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}}_{\mathbf{x}}$$

which is the general solution.

- The vector $\mathbf{x}_0$ is a **particular solution** of nonhomogeneous system, and the linear combination $\mathbf{x}$ is the general solution of the homogeneous system.

Inner Products and Norms

- For $\boldsymbol{x}, \boldsymbol{y} \in R^n$, their ***Euclidean inner product*** is

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} x_i y_i = \boldsymbol{x}^T \boldsymbol{y}$$

- Properties:
  - Positivity: $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$, $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ if and only if $\boldsymbol{x} = 0$
  - Symmetry: $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$
  - Additivity: $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$
  - Homogeneity: $\langle r\boldsymbol{x}, \boldsymbol{y} \rangle = r \langle \boldsymbol{x}, \boldsymbol{y} \rangle$ for every $r \in R$
- The vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are said to be ***orthogonal*** if $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$
- The ***Euclidean norm*** of a vector $\boldsymbol{x}$ is defined as

$$\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} = \sqrt{\boldsymbol{x}^T \boldsymbol{x}}$$

Inner Products and Norms

- ***Cauchy-Schwarz Inequality***: For any two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in R^n, the Cauchy-Schwarz inequality

  $$|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\|$$

  holds. Furthermore, equality holds if and only if $\boldsymbol{x} = \alpha \boldsymbol{y}$ for some $\alpha \in R$
- The Euclidean norm of a vector has the following properties
  - Positivity: $\|\boldsymbol{x}\| \geq 0$, $\|\boldsymbol{x}\| = 0$ if and only if $\boldsymbol{x} = \mathbf{0}$
  - Homogeneity: $\|r\boldsymbol{x}\| = |r|\|\boldsymbol{x}\|, r \in R$
  - Triangle inequality: $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$

Inner Products and Norms

- The Euclidean norm is an example of a general ***vector norm***, which is any function satisfying the three properties of positivity, homogeneity, and triangle inequality.
- Other vector norms:
  - 1-norm: $\|\boldsymbol{x}\|_1 = |x_1| + \cdots + |x_n|$
  - $\infty-$ norm: $\|\boldsymbol{x}\|_\infty = \max_i |x_i|$
- The Euclidean norm is often referred as the 2-norm, denoted by $\|\boldsymbol{x}\|_2$
- *p*-norm:

$$\|\boldsymbol{x}\|_p = \begin{cases} (|x_1|^p + \cdots + |x_n|^p)^{1/p} & \text{if } 1 \le p < \infty \\ \max\{|x_1|, \ldots, |x_n|\} & \text{if } p = \infty \end{cases}$$

Inner Products and Norms

- A function $f : R^n \rightarrow R^m$ is ***continuous*** at $\boldsymbol{x}$ if for all $\epsilon > 0$, there exists $\delta > 0$ such that $\|\boldsymbol{y} - \boldsymbol{x}\| < \delta \Longrightarrow \|f(\boldsymbol{y}) - f(\boldsymbol{x})\| < \epsilon$.
- If the function f is continuous at every point in R^n, we say that it is continuous on R^n.

Inner Products and Norms

- For the complex vector space C^n, we define an inner product $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} x_i \bar{y}_i$, where the bar denotes complex conjugation.
- Properties:
  - $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$, $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ if and only if $\boldsymbol{x} = 0$
  - $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle}$
  - $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$
  - $\langle r\boldsymbol{x}, \boldsymbol{y} \rangle = r \langle \boldsymbol{x}, \boldsymbol{y} \rangle$, where $r \in C$