Ch.3 Transformations

3.1 Linear Transformations

- A function $L : \Re^n \rightarrow \Re^m$ is called a linear transformation if
  1. $L(ax) = aL(x)$ for every $x \in \Re^n$ and $a \in \Re$
  2. $L(x + y) = L(x) + L(y)$ for every $x, y \in \Re^n$
- Let $\{e_1, e_2, \cdots, e_n\}$ and $\{e'_1, e'_2, \cdots, e'_n\}$ be two bases for $\Re^n$. We call T the transformation matrix from $\{e_1, e_2, \cdots, e_n\}$ to $\{e'_1, e'_2, \cdots, e'_n\}$.

$$[e_1, e_2, \cdots, e_n] = [e'_1, e'_2, \cdots, e'_n]T$$

- Let $y = Ax$ and $y' = Bx'$. Therefore, $y' = Ty = TAx = Bx' = BTx$, and hence $TA = BT$, or $A = T^{-1}BT$.

3.2 Eigenvalues and Eigenvectors

- A scalar λ (possibly complex) and a nonzero vector v satisfying the equation $Av = \lambda v$ are said to be, respectively, an eigenvalue and eigenvector of A.
- For λ to be an eigenvalue, $det\,[\lambda I - A] = 0$.
- Characteristic polynomial of the matrix A.

  $det\,[\lambda I - A] = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0.$

- **Theorem 3.1** *Suppose the characteristic equation* $\det[\lambda I - A] = 0$ *has* n *distinct roots* $\lambda_1, \lambda_2, \cdots, \lambda_n$. *Then, there exist* n *linearly independent vectors* $v_1, v_2, \cdots, v_n$ *such that*

$$Av_i = \lambda_i v_i, \;\; i = 1, 2, \cdots, n.$$

- **Theorem 3.2** *All eigenvalues of a symmetric matrix are real.*
- **Theorem 3.3** *Any real symmetric* $n \times n$ *matrix has a set* n *eigenvectors that are mutually orthogonal.*
- A matrix whose transpose equals to its inverse is said to be an orthogonal matrix.

$$T^T = T^{-1}$$

3.3 Orthogonal Projections

- If ν is a subspace of $\Re^n$, then the orthogonal complement of ν, denoted $\nu^\perp$,

$$\nu^\perp = \{x : v^T x = 0 \text{ for all } v \in \nu\}$$

- Let the range, or image, of A be denoted,

$$R(A) \equiv \{Ax : x \in \Re^n\}$$

- Nullspace, or kernel, of A be denoted,

$$N(A) \equiv \{x \in \Re^n : Ax = 0\}$$

- **Theorem 3.4** *Let A be a given matrix. Then,* $R(A)^{\perp} = N(A^T)$, *and* $N(A)^{\perp} = R(A^T)$.

3.4 Quadratic Forms

- A quadratic form $f : \Re^n \to \Re$ is a function,

$$f(x) = x^T Q x$$

where Q is an $n \times n$ real matrix. Q is assumed to be symmetric, that is, $Q = Q^T$.

- A quadratic form $x^T Q x$, $Q = Q^T$, is said to be positive definite if $x^T Q x > 0$ for all nonzero vectors x.
- It is positive semidefinite if $x^T Q x \geq 0$ for all x.
- The minors of a matrix Q are the determinants of the matrices obtained by successively removing rows and columns from Q.
- The principal minors are $det\, Q$ itself and the determinants of matrices obtained by successively removing an ith row and an ith column.
- The leading principal minors are $det\, Q$ and the minors obtained by successively removing the last row and last column.

$$\Delta_1 = q_{11},\ \Delta_2 = det \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix},\ \cdots, \Delta_n = det\, Q$$

- **Theorem 3.6 Sylvester's Criterion.** *A quadratic form x^TQx, $Q = Q^T$, is positive definite if and only if the leading principal minors of Q are positive.*
- If Q is not symmetric, Sylvester's Criterion cannot be used to check positive definiteness of the quadratic form x^TQx.
  For example,

$$Q = \begin{bmatrix} 1 & 0 \\ -4 & 1 \end{bmatrix}$$

- The leading principal minors of Q are $\Delta_1 = 1 > 0$ and $\Delta_2 = det\, Q = 1 > 0$. However, if $x = [1, 1]^T$, then $x^TQx = -2 < 0$, $\Rightarrow$ Not positive definite.

$$\begin{aligned} x^TQx &= x^T \begin{bmatrix} 1 & 0 \\ -4 & 1 \end{bmatrix} x \\ &\Rightarrow \frac{1}{2} x^T \left(\begin{bmatrix} 1 & 0 \\ -4 & 1 \end{bmatrix} + \begin{bmatrix} 1 & -4 \\ 0 & 1 \end{bmatrix} \right) x \\ &= x^T \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix} x = x^T Q_0 x \end{aligned}$$

- The leading principal minors of Q_0 are $\Delta_1 = 1 > 0$ and $\Delta_2 = det\, Q_0 = -3 < 0$, as expected.
- A necessary condition for a real quadratic form to be

positive semidefinite is that the leading principal minors be nonnegative.

- A symmetric matrix Q is said to be positive definite if the quadratic form $x^T Q x$ is positive defintie. If Q is positive definite, we write $Q > 0$.
- Symmetric matrix Q to be positive semidefinite ($Q \geq 0$), negative definite ($Q < 0$), and negative semidefinite ($Q \leq 0$).
- **Theorem 3.7** *A symmetric matrix Q is positive definite (or positive semidefinite) if and only if all eigenvalues of Q are positive (or nonnegative).*
- Nonnegativity of leading principal minors is necessary but not a sufficient condition for positive semidefiniteness.

3.5 Matrix Norms

- We define the norm of a matrix A, denoted $||A||$, to be any function $||\cdot||$ that satisfies the conditions:
  1. $||A|| > 0$ if $A \neq O$, and $||O|| = 0$, where O is the matrix with all entries equal to zero;
  2. $||cA|| = |c|||A||$, for any $c \in \Re$;
  3. $||A + B|| \leq ||A|| + ||B||$
  4. $||AB|| \leq ||A||||B||$.

- Frobenius norm(=Euclidean norm)

$$||A||_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} (a_{ij})^2 \right)^{\frac{1}{2}}$$

- We say that the matrix norm is included by, or is compatible with, the given vector norms if for any matrix $A \in \Re^{m \times n}$ and any vector $x \in \Re^n$, the following inequality is satisfied:

$$||Ax||_{(m)} \leq ||A|| \, ||x||_{(n)}$$

- We can define an induced matrix norm as:

$$||A|| = max_{||x||_{(n)}=1} ||Ax||_{(m)}$$

- **Theorem 3.8** *Let*

$$||x|| = \left(\sum_{k=1}^{n} |x_k|^2 \right)^{\frac{1}{2}} = \sqrt{< x, x >}$$

The matrix norm induced by this vector norm is,

$$||A|| = \sqrt{\lambda_1}$$

where λ_1 is the largest eigenvalue of the matrix $A^T A$.

- Example 3.1 Consider the matrix,

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

and let the norm in $\Re^2$ be given by

$$||x|| = \sqrt{x_1^2 + x_2^2}$$

Then,

$$A^T A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$$

and $det\,[\lambda I_2 - A^T A] = \lambda^2 - 10\lambda + 9 = (\lambda - 1)(\lambda - 9)$. Thus, $||A|| = \sqrt{9} = 3$. The eigenvector of $A^T A$ corresponding to $\lambda_1 = 9$ is

$$x_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Note that $||Ax_1|| = ||A||$.

$$||Ax_1|| = \|\frac{1}{2} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \| = 3.$$

Because $A = A^T$ in this example, we also have $||A|| = max_{1 \le i \le n} |\lambda_i(A)|$.

- In general, $max_{1 \le i \le n} |\lambda_i(A)| \neq ||A||$. Instead, we have $||A|| \geq max_{1 \le i \le n} |\lambda_i(A)|$.

- Example 3.2 Let,

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

then,

$$A^T A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$det\,[\lambda I_2 - A^T A] = det \begin{bmatrix} \lambda & 0 \\ 0 & \lambda - 1 \end{bmatrix} = \lambda(\lambda - 1)$$

Note that 0 is the only eigenvalue of A.

$$||A|| = 1 > |\lambda_i(A)| = 0.$$