Appendix

A. Matrix operators

a. Multiplication

(1) $\boldsymbol{AB} \neq \boldsymbol{BA}$; for example,

$$\begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 14 & 3 \\ -2 & -6 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}\begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ 29 & -2 \end{bmatrix}$$

(2) "$\boldsymbol{AB} = \boldsymbol{AC}$" $\not\Rightarrow$ $\boldsymbol{B} = \boldsymbol{C}$; for example,

$$\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \neq \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} -5 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

(3) $\boldsymbol{A}\,\boldsymbol{B} = \boldsymbol{0} \not\Rightarrow \boldsymbol{A} = \boldsymbol{0}$ or $\boldsymbol{B} = \boldsymbol{0}$; for example,

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

b. Transpose

(1) $(\boldsymbol{AB})^T = \boldsymbol{B}^T \boldsymbol{A}^T$

(2) $(\boldsymbol{A}+\boldsymbol{B})^T = \boldsymbol{A}^T + \boldsymbol{B}^T$

(3) $(r\boldsymbol{A})^T = r\,\boldsymbol{A}^T$

c. Inverse

(1) $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1} \boldsymbol{A}^{-1}$

(2) $(\boldsymbol{A}+\boldsymbol{B})^{-1} \neq \boldsymbol{A}^{-1} + \boldsymbol{B}^{-1}$; for example,

$$\boldsymbol{A} = \begin{bmatrix} 5 & 4 \\ 4 & 3 \end{bmatrix},\ \boldsymbol{B} = \begin{bmatrix} -5 & 4 \\ 4 & -3 \end{bmatrix} \Rightarrow (\boldsymbol{A}+\boldsymbol{B})^{-1} = \begin{bmatrix} 0 & 8 \\ 8 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1/8 \\ 1/8 & 0 \end{bmatrix}$$

$$\Rightarrow \boldsymbol{A}^{-1} = \begin{bmatrix} -3 & 4 \\ 4 & -5 \end{bmatrix},\ \boldsymbol{B}^{-1} = \begin{bmatrix} 3 & 4 \\ 4 & 5 \end{bmatrix} \Rightarrow \boldsymbol{A}^{-1} + \boldsymbol{B}^{-1} = \begin{bmatrix} 8 & 0 \\ 8 & 0 \end{bmatrix}.$$

(3) $(r\boldsymbol{A})^{-1} = (1/r)\,\boldsymbol{A}^{-1}$

(4) $(\boldsymbol{A}^T)^{-1} = (\boldsymbol{A}^{-1})^T$

d. Determinant

(1) det $\boldsymbol{AB}$ = det $\boldsymbol{A}$ det $\boldsymbol{B}$

(2) det $(\boldsymbol{A}+\boldsymbol{B}) \neq$ det $\boldsymbol{A}$ + det $\boldsymbol{B}$; for example,

$$\begin{vmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{vmatrix} \neq \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} + \begin{vmatrix} 5 & 6 \\ 7 & 8 \end{vmatrix}$$

(3) det $\boldsymbol{A}^T$ = det $\boldsymbol{A}$

(4) det $(r\boldsymbol{A}) \neq r$ det $\boldsymbol{A}$

(5) det $(\boldsymbol{A}^{-1})$ = 1 / det $\boldsymbol{A}$

B. Matrix factorization

a. Triangulation factorization

$$\boldsymbol{A}_{m\times n} = \boldsymbol{L}_{m\times m}\, \boldsymbol{D}_{m\times m}\, \boldsymbol{U}_{m\times n}\,,$$

where ***L*** is a lower triangle matrix with diagonal entries 1, ***D*** is a diagonal matrix, and ***U*** is a upper triangle matrix with diagonal entry 1.

Constraint: No

Purposes:
Solving a sequence of system equations, all with the same coefficient matrix. For example, solving $\boldsymbol{LDx} = \boldsymbol{b}_1$, $\boldsymbol{LUx} = \boldsymbol{b}_2$, …, $\boldsymbol{LUx} = \boldsymbol{b}_n$ instead of solving $\boldsymbol{Ax} = \boldsymbol{b}_1$, $\boldsymbol{Ax} = \boldsymbol{b}_2$, …, $\boldsymbol{Ax} = \boldsymbol{b}_n$, since the operation just involves substitution.

b. Diagonalization

$$\boldsymbol{A}_{n\times n} = \boldsymbol{P}_{n\times n}\, \boldsymbol{D}_{n\times n}\, \boldsymbol{P}^{-1}_{n\times n} \equiv \boldsymbol{A}_{n\times n}\, \boldsymbol{P}_{n\times n} = \boldsymbol{P}_{n\times n}\, \boldsymbol{D}_{n\times n}\,,$$

where ***D*** is an eigenvalue matrix of ***A*** and ***P*** is the eigenvector matrix of ***A***.

Constraint:
1. ***A*** is a square matrix.
2. ***A*** must have *n* linearly independent eigenvectors; that is, the columns of ***P*** are linearly independent.

Purposes:
Quickly compute $\boldsymbol{A}^k$ for large *k*.

c. QR Factorization

$$\boldsymbol{A}_{m\times n} = \boldsymbol{Q}_{m\times n}\, \boldsymbol{R}_{n\times n},$$

where the columns of $\boldsymbol{Q}$ form an orthonormal basis for Col $\boldsymbol{A}$; moreover, $\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}_{n\times n}$ and then $\boldsymbol{R} = \boldsymbol{Q}^T\boldsymbol{A}$, where $\boldsymbol{R}$ is an $n\times n$ upper triangular invertible matrix with positive entries on its diagonal.
Constraint: $\boldsymbol{A}$ must have n linearly independent columns.

Purposes:
For the linear system $\boldsymbol{A}_{m\times n}\, \boldsymbol{x}_{n\times 1} = \boldsymbol{b}_{m\times 1}$, the least-squares solution is $\hat{\boldsymbol{x}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$. In some cases, the normal equation for a least-squares problem can be ill-conditioned; that is, small errors in the calculations of the entries of $\boldsymbol{A}^T\boldsymbol{A}$ can sometimes cause relatively large errors in the solution $\hat{\boldsymbol{x}}$. If the columns of $\boldsymbol{A}$ are linearly independent, the least-squares solution can often be computed more reliably through a $\boldsymbol{QR}$ factorization of $\boldsymbol{A}$ and the unique least-squares solution is $\hat{\boldsymbol{x}} = \boldsymbol{R}^{-1}\boldsymbol{Q}^T\boldsymbol{b}$.
Another purpose: For finding eigenvalues.
$\boldsymbol{A} = \boldsymbol{Q}_1\, \boldsymbol{R}_1,\quad \boldsymbol{Q}_1^{\,T} = \boldsymbol{Q}_1^{-1}$.
Let $\boldsymbol{A}_1 = \boldsymbol{R}_1\, \boldsymbol{Q}_1$; that is, $\boldsymbol{R}_1 = \boldsymbol{A}_1\, \boldsymbol{Q}_1^{-1} \Rightarrow \boldsymbol{A} = \boldsymbol{Q}_1\, \boldsymbol{R}_1 = \boldsymbol{Q}_1\, \boldsymbol{A}_1\, \boldsymbol{Q}_1^{-1}$.
If $\boldsymbol{A}_1 = \boldsymbol{Q}_2\, \boldsymbol{R}_2$, let $\boldsymbol{A}_2 = \boldsymbol{R}_2\, \boldsymbol{Q}_2$; that is, $\boldsymbol{R}_2 = \boldsymbol{A}_2\, \boldsymbol{Q}_2^{-1}$, …

d. Diagonalization of symmetric matrices

(related to Principal Component Analysis)

$$\boldsymbol{A}_{n\times n} = \boldsymbol{P}_{n\times n}\, \boldsymbol{D}_{n\times n}\, \boldsymbol{P}^{-1}_{n\times n} = \boldsymbol{P}_{n\times n}\, \boldsymbol{D}_{n\times n}\, \boldsymbol{P}^{T}_{n\times n}$$

where $\boldsymbol{D}$ is an eigenvalue matrix of $\boldsymbol{A}$ and $\boldsymbol{P}$ is the orthonormal eigenvector matrix of $\boldsymbol{A}$; that is, $\boldsymbol{P}^{-1} = \boldsymbol{P}^T$. The eigenvalues may be negative.

$\boldsymbol{A}$ can be spectral decomposed as $\boldsymbol{A} = \sum_{i=1}^{n} \lambda_i \mathbf{e}_i \mathbf{e}_i^T$, where $\boldsymbol{e}_i$ is the ith orthonormal eigenvector. Every column of $\lambda_i \mathbf{e}_i \mathbf{e}_i^T$ is a multiple of $\boldsymbol{e}_i$ and each matrix $\mathbf{e}_i \mathbf{e}_i^T$ is a projection matrix; that is, the vector $(\mathbf{e}_i \mathbf{e}_i^T)\boldsymbol{x}$ is the orthogonal projection of $\boldsymbol{x}$ onto the subspace spanned by $\boldsymbol{e}_i$.

Constraint:

$\boldsymbol{A}$ is a symmetric matrix; $\boldsymbol{A}$ is a symmetric matrix if and only if $\boldsymbol{A}$ is orthogonally diagonalizable.

Purposes: for orthogonal decomposition.

e. Quadratic forms on R^n

$$Q(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{A}_{n\times n} \boldsymbol{x},$$

where ***A*** is a symmetric matrix and ***x*** is a vector.

A is symmetric ⇒

$$\boldsymbol{A}_{n\times n} = \boldsymbol{P}_{n\times n} \boldsymbol{D}_{n\times n} \boldsymbol{P}^T_{n\times n} \Rightarrow \boldsymbol{Q}(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} = \boldsymbol{x}^T \boldsymbol{P} \boldsymbol{D} \boldsymbol{P}^T \boldsymbol{x} = (\boldsymbol{P}^T \boldsymbol{x})^T \boldsymbol{D} (\boldsymbol{P}^T \boldsymbol{x})$$

We can take an orthogonal change of variable $\boldsymbol{P}^T\boldsymbol{x} = \boldsymbol{y}$ $(i.e., \boldsymbol{x} = \boldsymbol{P}\boldsymbol{y})$

such that the component of ***y*** are independent $\boldsymbol{Q}(\boldsymbol{x}) = \boldsymbol{y}^T \boldsymbol{D} \boldsymbol{y}$

The columns of ***P*** are called the principal axes.

The quadratic form ***Q*** is positive definite if ***Q***(***x***) > 0 for all $\boldsymbol{x} \neq \boldsymbol{0}$

negative definite if ***Q***(***x***) < 0 for all $\boldsymbol{x} \neq \boldsymbol{0}$,

indefinite if ***Q***(***x***) has both > 0 and < 0 for all $\boldsymbol{x} \neq \boldsymbol{0}$,

$\boldsymbol{Q}(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$ is

positive definite if and only if the eigenvalues of ***A*** are all positive.

negative definite if and only if the eigenvalues of ***A*** are all negative.

f. Cholesky factorization

$$\boldsymbol{A} = \boldsymbol{R}^T \boldsymbol{R},$$

where ***A*** is a symmetric matrix with positive eigenvalues,

and ***R*** is a upper triangular with positive diagonal.

g. Singular value decomposition

(The decomposition is not unique)

$$\boldsymbol{A}_{m\times n} = \boldsymbol{U}_{m\times m}\boldsymbol{\Sigma}_{m\times n}\boldsymbol{V}^T_{n\times n},$$

where $\boldsymbol{A}$ be an $m\times n$ matrix with rank r,

$\boldsymbol{V}$ is an orthonormal matrix with right singular vectors of $\boldsymbol{A}$ (the orthonormal basis of $\boldsymbol{R}^n$ consisting of eigenvectors of $\boldsymbol{A}^T\boldsymbol{A}$),

$\boldsymbol{U}$ is an orthonormal matrix with left singular vectors of $\boldsymbol{A}$ (the non-zero orthonormal basis for Col $\boldsymbol{A}$), and

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{D} & 0 \\ 0 & \underbrace{0}_{n-r \text{ columns}} \end{bmatrix} \} \; m\text{-}r \text{ rows}$$

where $\boldsymbol{D}$ is an $r\times r$ diagonal matrix and the diagonal entries are the first r singular values of $\boldsymbol{A}$.

Constraint: No

Purposes:

Not all matrices can be diagonalized $\boldsymbol{A}_{n\times n} = \boldsymbol{P}_{n\times n}\boldsymbol{D}_{n\times n}\boldsymbol{P}^{-1}_{n\times n}$ as $\boldsymbol{A}_{m\times n} = \boldsymbol{Q}_{m\times m}\boldsymbol{D}_{m\times n}\boldsymbol{P}^{-1}_{n\times n}$, but a factorization is possible for any $m\times n$ matrix $\boldsymbol{A}$. Such a special factorization is called the *singular value decomposition*. The singular value decomposition is the main tool for performing principal component analysis in practical applications.

Another purpose: to estimate the rank of a matrix.

The singular values of $\boldsymbol{A}$ are the square roots of the eigenvalues of $\boldsymbol{A}^T\boldsymbol{A}$. All eigenvalues of $\boldsymbol{A}^T\boldsymbol{A}$ are positive, thus all singular values are positive.

THE END !