1 Preliminaries

- Let $Y_1, Y_2, \ldots, Y_n$ be independent normally distributed random variables with $\mathrm{E}[Y_i] = \mu_i$ and $\mathrm{Var}(Y_i) = \sigma_i^2$. Let $a_1, a_2, \ldots, a_n$ be known constants. Then,
$$\sum_{i=1}^{n} a_i Y_i \sim N(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2).$$
- If $Y \sim N(\mu, \sigma^2)$, then
$$Z = \frac{Y - \mu}{\sigma} \sim N(0, 1).$$
- Let $Z \sim N(0, 1)$. Then, $Z^2 \sim \chi_1^2$.
- Suppose that $Y_1, Y_2, \ldots, Y_n \overset{iid}{\sim} N(\mu, \sigma)$. Then,
$$\frac{\bar{Y} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1).$$
- Suppose that $Y_1, Y_2, \ldots, Y_n \overset{iid}{\sim} (\mu, \sigma)$. Then,
$$\frac{\bar{Y} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$
approximately for sufficiently large n.
- Let $Z \sim N(0, 1)$ and $V \sim \chi_\nu^2$. If Z and V are independent, then
$$\frac{Z}{\sqrt{V/\nu}} \sim t_\nu.$$
- Let $V \sim \chi_\nu^2$ and $W \sim \chi_\eta^2$. If V and W are independent, then
$$\frac{V/\nu}{W/\eta} \sim F_{\nu,\eta}.$$
- Let $X \sim N(\delta, 1)$, then
$$X^2 \sim \chi_{1,\delta^2}^{2'},$$
where $\chi_{1,\delta^2}^{2'}$ is the noncentral chi-square distribution with ν degrees of freedom and noncentrality parameter δ^2.
- Let $V \sim \chi_{\nu,\lambda}^{2'}$ and $W \sim \chi_\eta^2$. If V and W are independent, then
$$\frac{V/\nu}{W/\eta} \sim F'_{\nu,\eta,\lambda},$$
where $F'_{\nu,\eta,\lambda}$ is a noncentral F distribution with ν and η degrees of freedom and noncentrality parameter λ.

2 Matrix algebra

- $m \times n$ matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \{a_{ij}\}_{i=1,..,m,j=1,..,n}$$

  Note that

  a. For $\mathbf{A} = \{a_{ij}\}_{i=1,..,m,j=1,..,n}$ and $\mathbf{B} = \{b_{ij}\}_{i=1,..,m,j=1,..,n}$,

$$\mathbf{A} + \mathbf{B} = \{a_{ij} + b_{ij}\}_{i=1,..,m,j=1,..,n}.$$

  b. For $\mathbf{A} = \{a_{ij}\}_{i=1,..,m,j=1,..,n}$ and $\mathbf{B} = \{b_{ij}\}_{i=1,..,n,j=1,..,p}$,

$$\mathbf{AB} = \{\sum_{k=1}^{n} a_{ik}b_{kj}\}_{i=1,..,m,j=1,..,p}.$$

  For example, for $\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 5 & 2 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 4 & 1 \\ 1 & 3 \end{pmatrix}$,

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} 5 & 4 \\ 6 & 5 \end{pmatrix}, \quad \mathbf{AB} = \begin{pmatrix} 7 & 10 \\ 22 & 11 \end{pmatrix}.$$

- Column space of a matrix

  $\mathcal{C}(\mathbf{A}) = \{c_1\mathbf{a}_1 + \cdots + c_n\mathbf{a}_n : c_1, \ldots, c_n \text{ are real numbers}\} = \{\mathbf{Ac} : \mathbf{c} = (c_1, \ldots, c_n)^T \in \mathbb{R}^n\}.$

- Linearly independent vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_k\mathbf{x}_k = \mathbf{0} \Rightarrow c_1 = c_2 = \cdots = c_k = 0$$

- Identity matrix

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

- Transpose of a $m \times n$ matrix $\mathbf{A}$, denoted by $\mathbf{A}^T$ or $\mathbf{A}'$

$$\underset{m \times n}{\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}^T} = \underset{n \times m}{\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}}$$

  Note that $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.

- Inverse of a $m \times m$ matrix $\mathbf{A}$, denoted by $\mathbf{A}^{-1}$
  $\mathbf{A}^{-1}$ is defined as a $m \times m$ matrix such that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}.$$

  A matrix $\mathbf{A}$ that has the inverse matrix is called a non-singular matrix.
  Note that $(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ when $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$ exist.

- Orthogonal matrix
  A $m \times m$ matrix $\mathbf{A}$ is an orthogonal matrix if

$$\mathbf{A}\mathbf{A}^{T} = \mathbf{A}^{T}\mathbf{A} = \mathbf{I}.$$

- Rank of a $m \times n$ matrix $\mathbf{A}$, denoted by rank($\mathbf{A}$) or $r(\mathbf{A})$
  rank($\mathbf{A}$)= the number of linearly independent columns of $\mathbf{A}$

  <u>Facts related ranks</u>

  - $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}^T)$
  - $\text{rank}(\mathbf{A}\mathbf{B}) \leq \text{rank}(\mathbf{A})$ and $\text{rank}(\mathbf{A}\mathbf{B}) \leq \text{rank}(\mathbf{B})$
  - $\text{rank}(\mathbf{A}\mathbf{B}) = \text{rank}(\mathbf{A})$ if $\mathbf{B}$ is a non-singular matrix.
  - $\text{rank}(\mathbf{A}\mathbf{A}^T) = \text{rank}(\mathbf{A}^T\mathbf{A}) = \text{rank}(\mathbf{A})$

- Determinant of a $m \times m$ (square) matrix $\mathbf{A}$, denoted by $\det(\mathbf{A})$ or $|\mathbf{A}|$
  Special case: Determinant of a 2×2 matrix is defined as

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

  <u>Facts related to determinants</u>

  - $\det(\mathbf{I}) = 1$
  - $\det(\mathbf{A}) = \det(\mathbf{A}^T)$
  - $\det(\mathbf{A}\mathbf{B}) = \det(\mathbf{A})\det(\mathbf{B})$ for square matrices $\mathbf{A}, \mathbf{B}$
  - $\det(\mathbf{A}^{-1}) = \det(\mathbf{A})^{-1}$
  - $\det(\mathbf{A}) \neq 0 \Leftrightarrow \mathbf{A}^{-1}$ exists.
  - For an orthogonal matrix $\mathbf{A}$, $\det(\mathbf{A}) = -1$ or 1

- Symmetric matrix
  A $m \times m$ matrix $\mathbf{A}$ is said to be symmetric if $\mathbf{A} = \mathbf{A}^T$.

- Idempotent matrix
  A $m \times m$ matrix $\mathbf{A}$ is said to be idempotent if $\mathbf{A}^2 = \mathbf{A}$.

- Trace of a $m \times m$ matrix $\mathbf{A}$, denoted by $\text{tr}(\mathbf{A})$

  $\text{tr}(\mathbf{A}) = \sum_{i=1}^{m} a_{ii}$ = sum of all diagonal elements

  Facts related to trace

  - $\text{tr}(\mathbf{A}) = \text{tr}(\mathbf{A}^T)$
  - $\text{tr}(\mathbf{A}+\mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$
  - $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$
  - If $\mathbf{A}$ is idempotent, $\text{rank}(\mathbf{A}) = \text{tr}(\mathbf{A})$.

- Determinant and inverse of a partitioned matrix

  Let $\mathbf{A}$ be a $m \times m$ matrix partitioned such that

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}.$$

  i. $\det \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} = \det(\mathbf{A}_{11})\det(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})$ when $\mathbf{A}_{11}^{-1}$ exists.

  ii. $\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{A}_{22\cdot 1}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{A}_{22\cdot 1}^{-1} \\ -\mathbf{A}_{22\cdot 1}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{A}_{22\cdot 1}^{-1} \end{pmatrix}$

  where $\mathbf{A}_{22\cdot 1} = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$ when $\mathbf{A}_{11}^{-1}$ and $\mathbf{A}^{-1}$ exist.

  Note that

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{I} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22\cdot 1} \end{pmatrix}.$$

- Important facts for a partitioned matrix

  Let $\mathbf{X}$ be a $n \times p$ matrix partitioned such that $\mathbf{X} = [\mathbf{X}_1 | \mathbf{X}_2]$. Then,

$$\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{X}_1 = \mathbf{X}_1, \quad \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{X}_2 = \mathbf{X}_2.$$

  Consider a matrix of the form

$$\mathbf{X}^T\mathbf{X} = \begin{pmatrix} \mathbf{X}_1^T\mathbf{X}_1 & \mathbf{X}_1^T\mathbf{X}_2 \\ \mathbf{X}_2^T\mathbf{X}_1 & \mathbf{X}_2^T\mathbf{X}_2 \end{pmatrix}.$$

  It can be shown that the inverse of this matrix is

$$(\mathbf{X}^T\mathbf{X})^{-1} = \begin{pmatrix} (\mathbf{X}_1^T\mathbf{X}_1)^{-1} + (\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{X}_2\mathbf{G}\mathbf{X}_2^T\mathbf{X}_1(\mathbf{X}_1^T\mathbf{X}_1)^{-1} & -(\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T\mathbf{X}_2\mathbf{G} \\ -\mathbf{G}\mathbf{X}_2^T\mathbf{X}_1(\mathbf{X}_1^T\mathbf{X}_1)^{-1} & \mathbf{G} \end{pmatrix}$$

  where $\mathbf{G} = [\mathbf{X}_2^T(\mathbf{I} - \mathbf{X}_1(\mathbf{X}_1^T\mathbf{X}_1)^{-1}\mathbf{X}_1^T)\mathbf{X}_2]^{-1}$.

- Eigenvalues and eigenvectors of a square matrix $\mathbf{A}$

  λ is eigenvalue of $\mathbf{A}$ if the equation $\mathbf{Ax} = \lambda\mathbf{x}$ has a non-zero solution $\mathbf{x}$.
  $\mathbf{x}(\neq \mathbf{0})$ is eigenvector of $\mathbf{A}$ associated with λ.

  $\Leftrightarrow \mathbf{Ax} = \lambda\mathbf{x}$ and $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ for $\mathbf{x} \neq \mathbf{0}$.

- Diagonalization of a real symmetric matrix

  For a $m \times m$ real symmetric matrix $\mathbf{A}$, there exists an orthogonal matrix $\mathbf{P}$ such that

$$\mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m \end{pmatrix}.$$

  Note that, for $\mathbf{P} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m]$,

  i. $\lambda_1, \lambda_2, \ldots, \lambda_m$ are eigenvalues of $\mathbf{A}$
  ii. $\mathbf{x}_i$ is the eigenvector of $\mathbf{A}$ associated with λ_i
  iii. $\mathbf{x}_i^T\mathbf{x}_i = 1$ and $\mathbf{x}_i^T\mathbf{x}_j = 0$ for $i \neq j$
  iv. $\mathbf{A} = \sum_{i=1}^m \lambda_i \mathbf{x}_i \mathbf{x}_i^T$

- Positive definite and positive semi definite matrices

  Let $\mathbf{A}$ be $m \times m$ symmetric matrix.

  $\mathbf{A}$: positive definite (p.d.) $\Leftrightarrow \mathbf{x}^T\mathbf{A}\mathbf{x} > 0,\ \forall \mathbf{A} \in \mathbb{R}^m,\ \mathbf{x} \neq \mathbf{0}$
  $\Leftrightarrow$ all eigenvalues of $\mathbf{A}$ are > 0
  $\Leftrightarrow \mathbf{A} = \mathbf{B}\mathbf{B}^T$ for some non-singular matrix $\mathbf{B}$

  $\mathbf{A}$: positive semi definite (p.s.d.) $\Leftrightarrow \mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0,\ \forall \mathbf{A} \in \mathbb{R}^m$
  $\Leftrightarrow$ all eigenvalues of $\mathbf{A}$ are ≥ 0
  $\Leftrightarrow \mathbf{A} = \mathbf{B}\mathbf{B}^T$ for some matrix $\mathbf{B}$

- Matrix Derivatives

  Let $\mathbf{A}$ be a $k \times k$ matrix of constants, $\mathbf{a}$ be a $k \times 1$ vector of constants, and $\mathbf{y}$ be a $k \times 1$ vector of variables.

  i. If $z = \mathbf{a}'\mathbf{y}$, then
$$\frac{\partial z}{\partial \mathbf{y}} = \frac{\partial \mathbf{a}'\mathbf{y}}{\partial \mathbf{y}} = \mathbf{a}$$
  ii. If $z = \mathbf{y}'\mathbf{y}$, then
$$\frac{\partial z}{\partial \mathbf{y}} = \frac{\partial \mathbf{y}'\mathbf{y}}{\partial \mathbf{y}} = 2\mathbf{y}$$
  iii. If $z = \mathbf{a}'\mathbf{A}\mathbf{y}$, then
$$\frac{\partial z}{\partial \mathbf{y}} = \frac{\partial \mathbf{a}'\mathbf{A}\mathbf{y}}{\partial \mathbf{y}} = \mathbf{A}'\mathbf{a}$$
  iv. If $z = \mathbf{y}'\mathbf{A}\mathbf{y}$, then
$$\frac{\partial z}{\partial \mathbf{y}} = \frac{\partial \mathbf{y}'\mathbf{A}\mathbf{y}}{\partial \mathbf{y}} = \mathbf{A}\mathbf{y} + \mathbf{A}'\mathbf{y}$$

v. If $\mathbf{A}$ is symmetric, then

$$\frac{\partial \mathbf{y}'\mathbf{A}\mathbf{y}}{\partial \mathbf{y}} = 2\mathbf{A}\mathbf{y}$$

- Expectation of a random vector

  Let $\mathbf{y}$ be a $p \times 1$ random vector with mean $\mathrm{E}[\mathbf{y}] = \boldsymbol{\mu}$ and variance-covariance matrix $\mathrm{Var}(\mathbf{y}) = \boldsymbol{\Sigma}$.

  i. $\mathrm{E}[\mathbf{a}^T\mathbf{y}] = \mathbf{a}^T\mathrm{E}[\mathbf{y}] = \mathbf{a}^T\boldsymbol{\mu}$
  ii. $\mathrm{E}[\mathbf{A}\mathbf{y}] = \mathbf{A}\mathrm{E}[\mathbf{y}] = \mathbf{A}\boldsymbol{\mu}$
  iii. $\mathrm{Var}(\mathbf{a}^T\mathbf{y}) = \mathbf{a}^T\mathrm{Var}(\mathbf{y})\mathbf{a} = \mathbf{a}^T\boldsymbol{\Sigma}\mathbf{a}$
  iv. $\mathrm{Var}(\mathbf{A}\mathbf{y}) = \mathbf{A}\mathrm{Var}(\mathbf{y})\mathbf{A}^T = \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T$
  v. $\mathrm{E}[\mathbf{y}^T\mathbf{A}\mathbf{y}] = \mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu}$.

  Note that $\mathrm{Var}(\mathbf{y}) = \mathrm{E}[(\mathbf{y}-\boldsymbol{\mu})(\mathbf{y}-\boldsymbol{\mu})^T]$.

- Distribution of quadratic forms

  i. Suppose that $\mathbf{y} \sim N(\boldsymbol{\mu}, \mathbf{I})$. Then,

  $$\mathbf{y}^T\mathbf{A}\mathbf{y} \sim \chi^{2'}_{p,\lambda} \quad \Leftrightarrow \quad \mathbf{A} \text{ is idempotent}, \ p = \mathrm{tr}(\mathbf{A}), \ \lambda = \boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu}$$

  and

  $$\mathbf{y}^T\mathbf{A}\mathbf{y} \text{ and } \mathbf{y}^T\mathbf{B}\mathbf{y} \text{ are independent} \quad \Leftrightarrow \quad \mathbf{A}\mathbf{B} = \mathbf{0}.$$

  ii. Suppose that $\mathbf{y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\boldsymbol{\Sigma}$ is non-singular. Then,

  $$\mathbf{y}^T\mathbf{A}\mathbf{y} \sim \chi^{2'}_{p,\lambda} \quad \Leftrightarrow \quad \mathbf{A}\boldsymbol{\Sigma} \text{ is idempotent}, \ p = \mathrm{tr}(\mathbf{A}\boldsymbol{\Sigma}), \ \lambda = \boldsymbol{\mu}^T\mathbf{A}\boldsymbol{\mu}$$

  and

  $$\mathbf{y}^T\mathbf{A}\mathbf{y} \text{ and } \mathbf{y}^T\mathbf{B}\mathbf{y} \text{ are independent} \quad \Leftrightarrow \quad \mathbf{A}\boldsymbol{\Sigma}\mathbf{B} = \mathbf{0}.$$