Generalized Inverse

· (Full Rank Factorization) $A_{p\times q}$ of rank r can always be factorized as $A = K_{p\times r}L_{r\times q}$, where K and L have full column and full row rank respectively.

· $A^{-1}_{\text{left}} = (A^\top A)^{-1}A^\top$; $A^{-1}_{\text{right}} = A^\top(AA^\top)^{-1}$.

[Theorem] All idempotent matrices (except I) are singular.

[Theorem] For any idempotent matrix $A, r(A) = \text{tr}(A)$.

[Theorem] Eigenvalues of idempotent matrices are either 0 or 1.

[Theorem] For a symmetric matrix A, if all its eigenvalues are 1 or 0, then A is idempotent.

· (Moore-Penrose inverse) Let A be an $m\times n$ matrix. If a matrix A^+ exists that satisfies (1) $AA^+$is symmetric; (2) A^+A is symmetric; (3) $AA^+A = A$; (4) $A^+AA^+ = A^+$. $A^+$is defined as a Moore-Penrose inverse of A.

[Theorem] Each matrix A has an A^+. $A = BC$ by full-rank factorization. Then $A^+ = C^\top(CC^\top)^{-1}(B^\top B)^{-1}B^\top$.

· (Properties) (1) The Moore-Penrose inverse is unique. (2) $(A^\top)^+ = (A^+)^\top$. (3) $r(A^+) = r(A)$. (4) If $A = A^\top$, then $A^+ = (A^+)^\top$. (5) If A is nonsingular, $A^{-1} = A^+$. (6) If A is symmetric idempotent, $A^+ = A$. (7) If $r(A_{m\times n}) = m$, then $A^+ = A^\top(AA^\top)^{-1}, AA^+ = I$, If $r(A_{m\times n}) = n$, then $A^+ = (A^\top A)^{-1}A^\top, A^+A = I$. (8) The matrices $AA^+, A^+A, I - AA^+$and $I - A^+A$ are all symmetric idempotent.

· (Generalized inverse) Let A be an $m\times n$ matrix and the generalized inverse A^-satisfies $AA^-A = A$.

· (Properties) (1) The M-P inverse is also a generalized inverse. (2) G-inverse may not be unique.
(3) Let X be $m\times n, r(x) = k > 0$. (a) $r(X^-) \geq k$; (b) X^-X and $XX^-$are idempotent; (c) $r(X^-X) = r(XX^-) = k$; (d) $X^-X = I$ if and only if $r(X) = n$; (e) $XX^- = I$ if and only if $r(X) = m$; (f) $\text{tr}(X^-X) = \text{tr}(XX^-) = k = r(X)$; (g) If $X^-$is any G-inverse of X, then $(X^-)^\top$ is a G-inverse of $X^\top$.
(4) Let $K = X(X^\top X)^-X^\top, K$ is invariant for any G-inverse of $X^\top X$.
(5) $X(X^\top X)^-X^\top = XX^+$.
(6) For $K = X(X^\top X)^-X^\top$. (a) $K = K^\top, K = K^2$ (Symmetric Idempotent); (b) $\text{rank}(K) = \text{rank}(X) = r$ ($\text{rank}(K) = \text{tr}(K) = \text{rank}(X)$); (c) $KX = X; X^\top K = X^\top$; (d) $(X^\top X)^-X^\top$ is a G-inverse of X for any G-inverse of $X^\top X$; (e) $X(X^\top X)^-$is a G-inverse of $X^\top$ for any G-inverse of $X^\top X$.

· (Algorithm of Obtaining Generalized Inverse) Let $A \in \mathbb{R}^{n\times n}$ be a matrix of rank r, and $A_{11} \in \mathbb{R}^{r\times r}$. If A_{11} is invertible, then $\mathbf{G} = \begin{bmatrix} A_{11}^{-1} & O \\ O & O \end{bmatrix} \in \mathbb{R}^{n\times n}$ is a generalized inverse of A.

Distributions and Quadratic Forms

[Theorem] Let Y be a random vector with mean $\mu = E(Y)$ and $\Sigma = \text{Cov}(Y)$. Then, $E\left(Y^\top AY\right) = \text{tr}(A\Sigma) + \mu^\top A\mu$.

[Theorem] Let $g_1(Y_1), \cdots, g_m(Y_m)$ be m functions of the random vectors $Y_1, \cdots, Y_m$ respectively. If $Y_1, \cdots, Y_m$ are mutually independent, then $g_1, \cdots, g_m$ are mutually independent.

· (Multivariate Normal Distribution) (a) Probability density function (p.d.f) of $Y_{p\times 1} \sim N(\mu, \Sigma)$:

$$f_Y(y) = |\Sigma|^{-\frac{1}{2}}(2\pi)^{-\frac{p}{2}}e^{-\frac{1}{2}\left[(y-\mu)^\top\Sigma^{-1}(y-\mu)\right]}$$

(b) Moment generating function of $Y_{p\times 1} \sim N(\mu, \Sigma)$: one way:

$$M_Y(t) = e^{t^\top\mu + \frac{1}{2}t^\top\Sigma}.$$

$G = \begin{pmatrix} 0 & 0^\top \\ 0 & p(\frac{1}{n}) \end{pmatrix}$ $H = \begin{pmatrix} 0 & 0^\top \\ 1_a & I_a \end{pmatrix}$

Let B be a constant matrix and c be a constant vector.

$$BY + c \sim N\left(B\mu + c, B\Sigma B^\top\right)$$

(c) Marginal distribution, Conditional Distribution and independence.

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \sim N\left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right],$$

1. $Y_1 \sim N(\mu_1, \Sigma_{11})$.
2. $Y_1 \mid Y_2 = y_2 \sim N\left(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(y_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)$.
3. Y_1 and Y_2 are independent iff $\Sigma_{12} = 0$.

· (Non-Central χ^2 distribution) The density function of $u \sim \chi^2_{(n)}$, a central χ^2 distribution, is

$$f(u) = \frac{u^{\frac{1}{2}n-1}e^{-\frac{1}{2}u}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}$$

Let $x \sim N(0, I_n)$, then $x^\top x \sim \chi^2_{(n)}$. Let $x \sim N(\mu, I_n)$, then $u = x^\top x \sim \chi^2_{(n,\lambda)}$, where $\lambda = \frac{1}{2}\mu^\top\mu$ is a non-centered parameter and the density function of $\chi^2_{(n,\lambda)}$ is

$$f(u) = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\frac{u^{\frac{1}{2}n+k-1}e^{-\frac{1}{2}u}}{2^{\frac{1}{2}n+k}\Gamma\left(\frac{1}{2}n+k\right)}, \quad u > 0, \lambda \geq 0.$$

Define $\lambda^k = 1$ when $\lambda = 0$ and $k = 0$. The moment generating function of $u \sim \chi^2_{(n,\lambda)}$ is

$$(1-2t)^{-\frac{n}{2}}e^{-\lambda\left[1-(1-2t)^{-1}\right]}.$$

When $\lambda = 0$, the above M.G.F is $(1-2t)^{-\frac{n}{2}}$ which is precisely the M.G.F of $\chi^2_{(n)}$.

· (Non-Central F distribution) Let $u_1 \sim \chi^2_{(p_1,\lambda)}$ and $u_2 \sim \chi^2_{(p_2,0)}$. And u_1 is independent of u_2. Then

$$w = \frac{u_1/p_1}{u_2/p_2} \sim F_{(p_1,p_2,\lambda)}$$

Let $z \sim N(\mu, 1)$ and $u \sim \chi^2_{(n)}$. And z is independent of u. Then,

$$t = \frac{z}{\sqrt{u/n}} \sim \text{ Non-central } t \text{ distribution.}$$

[Theorem] A symmetric matrix A is positive definite if and only if there exists a nonsingular matrix P such that $A = P'P$.

[Lemma] The M.G.F of $x^\top Ax$ is

$$M_{x^\top Ax}(t) = |I - 2tA\Sigma|^{-\frac{1}{2}}e^{-\frac{1}{2}\mu^\top\left[I-(I-2tA\Sigma)^{-1}\right]\Sigma^{-1}\mu}.$$

[Lemma] If A and V are symmetric and V is positive definite, then AV has eigenvalues 0 and 1 implies that AV is idempotent.

[Lemma] If A is $n\times n$ symmetric idempotent matrix of rank r, then A has r eigenvalues equal to 1 and $n - r$ eigenvalues equal to 0.

· (Properties of eigenvalues) (1) For certain function $g(A), g(\lambda)$ is an eigenvalue of $g(A)$. (2) If $(I - A)$ is nonsingular, then $1/(1-\lambda)$ is an eigenvalue of $(I - A)^{-1}$. (3) If $-1 < \lambda < 1$, then $1/(1-\lambda)$ can be represented by the series $\frac{1}{1-\lambda} = 1 + \lambda + \lambda^2 + \lambda^3 + \cdots$ Correspondingly, if all eigenvalues of A satisfying $-1 < \lambda < 1$, then$(I - A)^{-1} = I + A + A^2 + A^3 + \cdots$ (4) If A is any $n\times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$, then $|A| = \prod_{i=1}^n\lambda_i$ and $\text{tr}(A) = \sum_{i=1}^n\lambda_i$

[Theorem] Let $x_{p\times 1} \sim N(\mu, \Sigma)$ and let A be symmetric. Then, $q = x^\top Ax$ follows $\chi^2_{(r,\lambda)}$ with r being the rank of A and $\lambda = \frac{\mu^\top A\mu}{2}$ if and only if $A\Sigma$ is idempotent.

· (Properties of normal random vector) (1) $E(x^\top Ax) = \text{tr}(A\Sigma) + \mu^\top A\mu$. (2) $\text{Cov}(x, x^\top Ax) = 2\Sigma A\mu$. (3) $\text{Var}(x^\top Ax) = 2\,\text{tr}[(A\Sigma)^2] + 4\mu^\top A\Sigma A\mu$. (5) $\text{Cov}(x^\top A_1x, x^\top A_2x) = 2\,\text{tr}(A_1\Sigma A_2\Sigma) + 4\mu^\top A_1\Sigma A_2\mu$. (6)$\text{Cov}(x, x^\top Ax) = 2\Sigma A\mu$.

[Theorem] Let $x \sim N_n(\mu, \Sigma)$. Then $x^\top Ax$ and Bx are distributed independently if and only if $B\Sigma A = 0$.

[Theorem] Let $x \sim N(\mu, \Sigma)$. Then $x^\top Ax$ and $x^\top Bx$ are distributed independently if and only if $B\Sigma A = 0$ (or equivalently, $A\Sigma B = 0$).

[Theorem] Let $x \sim N(\mu, V)$ and let A_i be $n\times n$ symmetric matrix of rank k_i, for $i = 1, 2, \cdots, p$. Denote $A = \sum_{i=1}^p A_i$, which is symmetric with rank k. Then $x^\top A_ix \sim \chi^2_{(k_i,\frac{1}{2}\mu^\top A_i\mu)}$. $x^\top A_ix$ are pairwise independent and $x^\top Ax$ is $\chi^2_{(k,\frac{1}{2}\mu^\top A\mu)}$, if and only if (I) : any 2 of (a) A_iV idempotent, for all i; (b) $A_iVA_j = 0$ for all $i < j$; (c) AV is idempotent; are true; or (II) : (c) is true and (d) $k = \sum_{i=1}^p k_i$; or(III) : (c) is true and (e) $A_1V, \cdots, A_{p-1}V$ are indempotent and A_pV is non-negative definite.

[Corollary] (Cochran's Theorem) $x \sim N(0, I_n)$ and A_i is symmetric of rank r_i for $i = 1, \cdots, p$ with $\sum_{i=1}^p A_i = I_n$, then $x^\top A_ix$ are distributed independently as $\chi^2_{r_i}$ if and only if $\sum_{i=1}^p r_i = n$.

Full Rank Model

· (Properties of the least squares estimate) Given the least squares estimate, we define the vector of residuals as $\hat{\varepsilon} = Y - X\hat{\beta} = [I - H]Y$ where $H = X(X^\top X)^{-1}X^\top$. (1) The hat matrix H is symmetric idempotent; (2) $X^\top\hat{\varepsilon} = 0$; (3) $\hat{Y}^\top\hat{\varepsilon} = 0$; (4) $I - H$ is symmetric idempotent; (5) $E(\hat{\beta}) = \beta_0$ (unbiased estimate); (6) $\text{Cov}(\hat{\beta}) = (X^\top X)^{-1}\sigma^2$; (7) $\text{tr}(I_n - H) = n - p$; (8) $\hat{\varepsilon}^\top\hat{\varepsilon} = \text{tr}(YY^\top(I - H))$; (9) $E(YY^\top) = \sigma^2I + X\beta\beta^\top X^\top$; (10) $\hat{\sigma}^2 = \hat{\varepsilon}^\top\hat{\varepsilon}/(n-p)$ is an unbiased estimate of σ^2, that is $E(\frac{\hat{\varepsilon}^\top\hat{\varepsilon}}{n-p}) = \sigma^2$.

· The weighted least squars estimator is defined as the minimizer of $(Y - X\beta)^\top\Sigma^{-1}(Y - X\beta)$; $\tilde{\beta} = \varepsilon(X^\top\Sigma^{-1}X)^{-1}X^\top\Sigma^{-1}Y$. Note $E(\tilde{\beta}) = \beta_0$ and $\text{Cov}(\tilde{\beta}) = \varepsilon(X^\top\Sigma^{-1}X)^{-1}$.

· To find the BLUE of $t^\top\beta$ is to determine λ such that $\lambda^\top Y$ is unbiased for $t^\top\beta$ and has minimum variance among all the linear unbiased estimates. (hence $\lambda^\top X = t^\top$) The BLUE of $t^\top\beta$ is $t^\top(X^\top\Sigma^{-1}X)^{-1}X^\top\Sigma^{-1}Y$ with variance $t^\top(X^\top\Sigma^{-1}X)^{-1}t$.

[Theorem] $W = \lambda^\top\Sigma\lambda$ is minimized if $\lambda^\top = t^\top(X^\top\Sigma^{-1}X)^{-1}X^\top\Sigma^{-1}$ subject to the constraint that $X^\top\lambda = t$.

· $\text{Cov}(X, Y) = E[\text{Cov}(X, Y \mid Z)] + \text{Cov}(E(X \mid Z), E(Y \mid Z))$.

[Lemma] If

$$M = \begin{bmatrix} X^\top \\ Z^\top \end{bmatrix}\begin{bmatrix} X & Z \end{bmatrix} = \begin{bmatrix} X^\top X & X^\top Z \\ Z^\top X & Z^\top Z \end{bmatrix} = \begin{bmatrix} A & B \\ B^\top & D \end{bmatrix},$$

and put

$$W = (D - B^\top A^{-1}B)^{-1} = \left[Z^\top Z - Z^\top X\left(X^\top X\right)^{-1}X^\top Z\right]^{-1},$$

then,

$$M^{-1} = \begin{bmatrix} A^{-1} + A^{-1}BWB^\top A^{-1} & -A^{-1}BW \\ WB^\top A^{-1} & W \end{bmatrix}$$
$$= \begin{bmatrix} -A^{-1}B \\ I \end{bmatrix}W\begin{bmatrix} -B^\top A^{-1} & I \end{bmatrix} + \begin{bmatrix} A^{-1} & 0 \\ 0 & 0 \end{bmatrix}.$$