BAB II
MATERI PENUNJANG

2.1. ALJABAR MATRIK

Matrik \( A_{mn} \) dapat ditinjau dari vektor-vektor baris atau dari vektor-vektor kolom. Sub-ruang dari \( \mathbb{R}^n \) yang direntang oleh vektor-vektor baris dinamakan ruang baris dari \( A \) dan sub-ruang dari \( \mathbb{R}^m \) yang direntang oleh vektor-vektor kolom dinamakan ruang kolom dari \( A \), oleh karena itu perlu diingat kembali definisi-definisi berikut ini.

**Definisi 2.1.1**
Sebuah vektor \( w \) dinamakan kombinasi linier dari vektor-vektor \( v_1, v_2, ..., v_r \) jika vektor tersebut dapat dinyatakan dalam bentuk

\[
 w = k_1v_1 + k_2v_2 + ... + k_rv_r
\]

dengan \( k_1, k_2, ..., k_r \) adalah skalar.

**Definisi 2.1.2**
Jika \( v_1, v_2, ..., v_r \) adalah vektor-vektor di dalam sebuah ruang vektor \( V \) dan jika tiap vektor di dalam \( V \) dapat dinyatakan sebagai kombinasi linier dari \( v_1, v_2, ..., v_r \) maka vektor-vektor tersebut merentang \( V \).

**Definisi 2.1.3**
Sebuah himpunan vektor \( S = \{ v_1, v_2, ..., v_r \} \) dan persamaan vektor

\[
 k_1v_1 + k_2v_2 + ... + k_rv_r = 0
\]

mempunyai paling sedikit satu pemecahan yakni : \( k_1=0, k_2=0, ..., k_r=0 \).
Jika pemecahan tersebut adalah satu-satunya pemecahan maka pemecahan tersebut dinamakan pemecahan trivial dan \( S \) dinamakan sebuah himpunan yang
bebas linier. Jika ada pemecahan lainnya maka pemecahan lainnya dinamakan pemecahan tak trivial dan S dinamakan sebuah himpunan tak bebas linier.

**Definisi 2.1.4**
Jika $V$ adalah sembarang ruang vektor dan $S = \{ v_1, v_2, \ldots, v_r \}$ adalah sebuah himpunan berhingga dari vektor-vektor di dalam $V$, maka $S$ dinamakan sebuah basis untuk $V$ jika:

(i). $S$ bebas linier

(ii). $S$ merentang $V$

**Definisi 2.1.5**
Dimensi dari sebuah ruang vektor $V$ yang berdimensi berhingga didefinisikan sebagai banyaknya vektor di dalam sebuah basis untuk $V$.

2.1.1. Rank Matrik

**Definisi 2.1.6**
Dimensi baris dan ruang kolom dari sebuah matrik $A$ dinamakan rank baris dan rank kolom dari $A$.

**Teorema 2.1.1**
(i) Jika $A_{m \times k}$ dan $B_{k \times n}$ maka $r(AB) \leq \min\{r(A), r(B)\}$

(ii) Jika matrik $P$ dan $Q$ yang non singular maka $r(PAQ) = r(A)$

Bukti :

(i) Baris-baris $AB$ merupakan kombinasi linier dari baris-baris $B$ sehingga jumlah baris-baris bebas linier $AB$ kurang dari atau sama dengan baris-baris $B$ karena jika tidak, tidak mungkin untuk menyatakan baris-baris $AB$ sebagai kombinasi linier dari baris-baris $B$ maka $r(AB) \leq r(B)$, kemudian kolom-kolom $AB$ merupakan kombinasi linier kolom-kolom $A$ maka $r(AB) \leq r(A)$.
Jadi $r(AB) \leq \min (r(A), r(B))$

(ii) Dari (i) diatas didapatkan $r(A) \geq r(AQ) \geq r(AQQ^{-1}) = r(A)$ jadi $r(A) = r(AQ)$ dan $r(A) \geq r(PA) \geq r(P^{-1}PA) = r(A)$ jadi $r(A) = r(PA)$ sehingga $r(A) \geq r(PAQ) \geq r(P^{-1}PAQQ^{-1}) = r(A)$.

Terbukti $r(A) = r(PAQ)$.

2.2.1. Diagonalisasi Matrik Bujursangkar

Definisi 2.1.7
Suatu vektor tak nol $x$ didalam $\mathbb{R}^n$ dinamakan sebuah vektor eigen dari matrik $A_{an}$ jika $Ax = \lambda x$, dengan skalar $\lambda$ dinamakan nilai eigen dari $A$ dan $x$ dikatakan suatu vektor eigen yang berpasangan dengan $\lambda$.

Definisi 2.1.8
1. Vektor $v_i$, $v_j$ adalah vektor ortogonal jika $v_i^Tv_j = 0$ (semua $i,j : i \neq j$) dan vektor $v_i$, $v_j$ adalah vektor ortonormal jika $v_i^Tv_i = 1$ dan $v_i^Tv_j = 0$ (semua $i,j : i \neq j$)

2. Matrik bujursangkar $A_{an}$ adalah matrik ortogonal jika kolom-kolom dari $A$ adalah himpunan vektor-vektor ortonormal, dan dimiliki sifat-sifat berikut,

(i) $A^TA = I_n$

(ii) $A^{-1} = A^T$

(iii). $A^T$ adalah matrik ortogonal

Definisi 2.1.9
Jika ada matrik nonsingular $P$ sedemikian sehingga $B = PAP^{-1}$, maka matrik bujursangkar $A$ dan $B$ dikatakan similar ($A$ dan $B$ mempunyai nilai eigen yang sama).

Materi Penunjang
Definisi 2.1.10
Matrik $A_{nm}$ dapat didiagonalisasi secara ortogonal jika terdapat matrik ortogonal $U$ sehingga $U^{-1}AU = U^T AU = D$ diagonal. Matrik $U$ dikatakan mendiaagonalisasi $A$ secara ortogonal.

Teorema 2.1.2
Matrik $A$ adalah matrik simetris jika dan hanya jika $A$ dapat didiagonalisasi secara ortogonal.

Bukti :
$(\Rightarrow)$ Diketahui matrik $A_{nm}$ adalah matrik simetris.
Misalkan ke-$n$ nilai eigen dari matrik $A$ semua berlainan ($\lambda_i \neq \lambda_j$; $i,j = 1, ..., n$). Kemudian misal $x_i$ dan $x_j$ adalah vektor eigen yang berpasangan dengan nilai eigen $\lambda_i$ dan $\lambda_j$ sedemikian sehingga

$$Ax_i = \lambda_i x_i \quad \text{dan} \quad Ax_j = \lambda_j x_j \quad \text{dengan} \quad x_i, x_j \neq 0$$

makanya

$$x_j^T Ax_i = \lambda_j x_j^T x_i$$

dan

$$x_i^T Ax_j = \lambda_i x_i^T x_j$$

Karena nilai eigen $\lambda_i \neq \lambda_j$ dan $A$ simetris maka $x_j^T Ax_i = x_i^T Ax_j$, sehingga

$$(\lambda_i - \lambda_j) x_j^T x_i = 0$$

dengan demikian $x_j^T x_i = 0$. Jadi vektor eigen dengan nilai eigen yang berbeda adalah ortogonal, kemudian dibuktikan $x_1, x_2, ..., x_n$ bebas linier,

$$k_1 x_1 + k_2 x_2 + ... + k_n x_n = 0$$

dengan membentuk hasil kali skalar dengan $x_1$ diperoleh

$$k_1 x_1^2 + k_2 x_1^T x_2 + ... + k_n x_1^T x_n = 0$$

karena $x_1 \neq 0$ sehingga $x_1^2 \neq 0$ dan $x_j^T x_i = 0$ untuk $i \neq j$ maka $k_1 = 0$, kemudian dibentuk perkalian lagi dengan $x_2$ didapat $k_2 = 0$ sedemikian seterusnya, hingga

Material Penunjang
didapat setiap \( k_i = 0 \), maka \( x_1, x_2, ..., x_n \) bebas linier.

Misalkan \( u_i = x_i / ||x_i|| \) maka \( u_i^T u_i = 1 \) dan \( u_i^T u_j = 0 \) (\( i \neq j \)) sehingga \( u_1, u_2, ..., u_n \) adalah vektor orthonormal, jadi \( u_1, u_2, ..., u_n \) adalah basis orthonormal untuk \( \mathbb{R}^n \). Jadi bila semua nilai eigen berlainan, maka tidak lebih dari satu vektor eigen bebas linier dapat berpasangan dengan suatu nilai eigen yang sama dan sehingga ada matrik ortogonal \( U = [u_1, u_2, ..., u_n] \) sedemikian sehingga \( U^T A U = D \).

Misal tidak semua nilai eigen berlainan (misal ada nilai eigen \( \lambda_j \) sebanyak \( k \) buah), akan dibuktikan bahwa ada \( k \) vektor eigen orthonormal berpasangan dengan nilai eigen \( \lambda_j \).

Andaikan sedikitnya ada satu vektor eigen orthonormal dengan nilai eigen \( \lambda_j \) sebut \( u_j \) sehingga \( u_j, v_1, ..., v_{n-1} \) adalah suatu basis orthonormal untuk \( \mathbb{R}^n \).

Tinjau matrik \( U = [u_j, v_1, ..., v_{n-1}] \) maka

\[
AU = [Au_j, Av_1, ..., Av_{n-1}] = [\lambda_j u_j, Av_1, ..., Av_{n-1}]
\]

\[
U^T A U = \begin{bmatrix}
\lambda_j & u_j^T Av_1 & \cdots & u_j^T Av_{n-1} \\
\lambda_j v_1^T u_j & v_1^T Av_1 & \cdots & v_1^T Av_{n-1} \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_j v_{n-1}^T u_j & v_{n-1}^T Av_1 & \cdots & v_{n-1}^T Av_{n-1}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\lambda_j & 0 \\
0 & Z
\end{bmatrix}
\]

dengan \( v_1^T u_j = 0 \); \( u_j^T Av_i = (u_j^T Av_i)^T = v_i^T A u_j = \lambda_j v_i^T u_i = 0 \) (\( i = 1, ..., n-1 \)) dan \( Z \) adalah matrik simetris dengan \( Z = [z_{s,s}] = [v_i^T Av_i] \), (\( i, s = 1, ..., n-1 \)).
Misalkan $U^T A U = B$ dan karena $U$ matrik ortogonal maka $B$ dan $A$ similar dengan demikian $B$ dan $A$ mempunyai nilai eigen $\lambda$ yang sama, sehingga

$$|B - \lambda I_n| = \begin{bmatrix} \lambda_j & 0 \\ 0 & Z \end{bmatrix} - \lambda I_n = (\lambda_j - \lambda) |Z - \lambda I_{n-1}|$$

Jadi jika $\lambda_j$ adalah nilai eigen dari $A$ sebanyak $k \geq 2$ maka

$$B - \lambda_j I_n = \begin{bmatrix} 0 & 0 \\ 0 & Z - \lambda_j I_{n-1} \end{bmatrix}$$

dan $|Z - \lambda_j I_{n-1}| = 0$ sehingga kenalan dari $B - \lambda_j I_n \geq 2$ karena $r(A - \lambda_j I_n) = r(B - \lambda_j I_n)$, kenalan dari $A - \lambda_j I_n \geq 2$, akibatnya ada suatu vektor eigen $u_j$ dari $A$ dengan nilai eigen $\lambda_j$ bebas linier dan tegak lurus pada $u_i$. Dalam cara ini dapat ditunjukkan jika nilai eigen $\lambda_j$ sebanyak $k$ buah maka ada $k$ vektor eigen ortonormal dengan nilai eigen $\lambda_j$. Sehingga $k$ vektor eigen ortonormal tersebut bersama vektor eigen yang berpasangan dengan nilai eigen lainnya menyusun basis ortonormal untuk $\mathbb{R}^n$. Sehingga ada matrik ortonormal $U$ sedemikian sehingga $U^T A U = D$.

($\Leftarrow$) Diketahui matrik $A_{nm}$ dapat didiagonalisasi secara ortonormal maka terdapat matrik ortonormal $U$ sedemikian sehingga $D = U^T A U$.

Matrik $U$ ortonormal sehingga $U^T = U^{-1}$ sedemikian sehingga,

$$D = U^T A U$$

$$U D U^T = U U^T A U U^T$$

$$= A$$

dan sehingga $A^T = (U D U^T)^T$

$$= U D U^T$$

$$= A$$

Terbukti bahwa $A$ simetris.
Teorema 2.1.3
Jika A simetris maka \( r(A) \) sama dengan banyaknya nilai eigen A yang bukan nol.

Bukti:
Misalkan \( r(A) = r \) dan karena A simetris maka menurut teorema (2.1.2) ada matrik ortogonal U sedemikian sehingga \( U^T A U = D = \text{diagonal} (\lambda_1, \lambda_2, ..., \lambda_n) \) dengan \( \lambda_i \) adalah nilai eigen matrik A. Menurut teorema 2.1.1 bagian (ii) diatas \( r(A) = r(U^T A U) = r(D) = r \) dengan \( D = \text{diagonal} (\lambda_1, \lambda_2, ..., \lambda_n)^{\text{diagonal}} (\lambda_1, \lambda_2, ..., \lambda_r, 0, ..., 0) \). Terbukti \( r(A) \) adalah banyaknya nilai eigen yang bukan nol.

2.1.3. Matrik Idempoten
Definisi 2.1.11
Matrik A adalah idempoten jika \( A^2 = A^T A = A \) dan A disebut matrik proyeksi jika A idempoten dan simetris.

Definisi 2.1.12
Trace suatu matrik \( [a_{ij}]_{m,n} \) adalah jumlah elemen-elemen diagonal matrik A

\[
\text{tr}(A) = \sum_{i} a_{ii}
\]

Jika A, B dan c adalah matrik dan skalar maka didapatkan sifat sebagai berikut:

(i) \( \text{tr}(A+B) = \text{tr}(A) + \text{tr}(B) \)

(ii) \( \text{tr}(A^T) = \text{tr}(A) \)

(iii) \( \text{tr}(cA) = c \text{tr}(A) \)

(iv) \( \text{tr}(AB) = \text{tr}(BA) \)

Teorema 2.1.4
Jika A matrik simetris maka

(i) Jika A idempoten maka rank dan trace dari A adalah sama \( r(A) = \text{tr}(A) \).

(ii) \( I_n - A \) adalah idempoten dengan \( r(I_n - A) = n - r(A) \).
Bukti:

(i) Diketahui $A$ matrik simetris dan idempoten ($A^2 = A$), misalkan persamaan karakteristik $Ax = \lambda x$ ($x \neq 0$) dengan $x$ adalah vektor eigen matrik $A$ dan skalar $\lambda$ adalah nilai eigen dari $A$ maka $x^TAx = \lambda x^Tx$ .....(2.1.1)
dan karena $A$ idempoten maka $x^TAx = x^TA^2x$

$$= x^TATAx$$

$$= (Ax)^T(Ax) = \lambda^2 x^Tx$$ ....(2.1.2)
dari (2.1.1) dan (2.1.2) diatas didapatkan

$$\lambda x^Tx - \lambda^2 x^Tx = \lambda x^Tx(\lambda - 1)$$

$$= 0$$

untuk $(\lambda - 1) = 0$ maka $\lambda = 1$ atau $\lambda x^Tx = 0$, karena $x \neq 0$ maka $\lambda x^Tx = 0$
hanya jika $\lambda = 0$. Dengan demikian nilai eigen $\lambda$ adalah 0 atau 1. Menurut teorema (2.1.3) terdapat sebanyak $r(A)$ nilai eigen yang bernilai 1 dan sebanyak $n - r(A)$ yang bernilai 0, sehingga $A$ dapat didiagonalkan dalam bentuk:

$$U^TAT = D = \begin{bmatrix} I_{r(A)} & 0 \\ 0 & 0 \end{bmatrix}$$
dari bentuk diagonal diatas didapat $tr(U^TUA) = r(A)$ dan karena $tr(U^TUA) =
tr(AUU^T) = tr(A)$ maka terbukti $tr(A) = r(A)$.

(ii) $(I - A)(I - A) = I - A - A + A^2$

$$= I - 2A + A = I - A$$

jadi $(I - A)$ adalah idempoten.

Sehingga menurut (i) diatas $r(I - A) = tr(I - A) = tr(I) - tr(A) = n - tr(A) = n - r(A)$
2.1.4. Bentuk Kuadrat

**Definisi 2.1.13**

Suatu bentuk kuadrat dalam variabel \( x_1, x_2, \ldots, x_n \)

\[
 f(x_1,x_2,\ldots,x_n) = a_{11}x_1^2 + \ldots + a_{nn}x_n^2 + 2a_{12}x_1x_2 + \ldots + 2a_{1n}x_1x_n + 2a_{22}x_2^2 + \ldots + 2a_{nn}x_n^2
\]

\[
 = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_i x_j = x^T A x
\]

dengan \( x = (x_1, x_2, \ldots, x_n)^T \) dan \( A_{nn} \) adalah matriks simetris bentuk kuadrat dalam \( x \) dan \( A \) disebut matrik bentuk kuadrat.

**Definisi 2.1.14**

Matrik simetris \( A \) disebut matrik definit positif jika \( x^T A x > 0, \forall x \neq 0 \) dan disebut matrik semidefinisit positif jika \( x^T A x \geq 0, \forall x \). Juga disebut matrik definit negatif jika \( x^T A x < 0, \forall x \neq 0 \).

Contoh 2.1 :

1. \( F = x^T A x = [x_1, x_2] \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \)

\[
= 3x_1^2 + 5x_2^2
\]

adalah definit positif karena tak pernah negatif dan nol jika hanya jika \( x = 0 \).

2. \( F = x^T A x = [x_1, x_2, x_3] \begin{bmatrix} 4 & -2 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \)

\[
= 4x_1^2 + x_2^2 - 4x_1x_2 + 3x_3^2
\]

\[
= (2x_1 - x_2)^2 + 3x_3^2
\]

adalah semidefinisit positif, karena selalu positif dan nol jika \( x_2 = 2x_1, x_3 = 0 \).
Bentuk Kuadrat

3. \( F = x^T A x = [x_1, x_2] \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -x_1^2 - x_2^2 \)

adalah definit negatif, karena tak pernah positif dan nol jika hanya jika \( x = 0 \).

**Teorema 2.1.5**

Jika \( A \) matrik simetris maka

(i) \( A \) definit positif jika dan hanya jika ada matrik nonsingular \( P \) sedemikian sehingga \( A = P P^T \).

(ii) Jika \( A \) definit positif maka \( A^{-1} \) juga definit positif.

(iii) Jika \( A \) definit positif maka ada matrik nonsingular \( Q \) sedemikian sehingga

\( Q^T A Q = I. \)

Bukti:

(i) \( (\Rightarrow) \) Diketahui bahwa \( A \) definit positif dan karena \( A \) simetris maka \( A \) dapat didiagonalisasi secara ortogonal sehingga \( U^T A U = D \) dan \( A = U^T D U \)

Untuk \( \forall x \neq 0 \), \( x^T A x = x^T U^T D U x \)

\( = (Ux)^T D (Ux) \)

\( = y^T D y = \sum_i \lambda_i y_i^2 \)

dengan \( y = Ux \), dan karena \( x \neq 0 \) maka \( y \neq 0 \). Sehingga \( \sum_i \lambda_i y_i^2 > 0 \), untuk \( \forall \lambda_i > 0 \) (i = 1, 2, ..., n). Jadi ada \( D^{1/2} = \text{diagonal} (\lambda_1^{1/2}, \lambda_2^{1/2}, ..., \lambda_n^{1/2}) \).

maka \( A = U D U^T \)

\( = U D^{1/2} (D^{1/2})^T U^T \)

\( = U D^{1/2} (U D^{1/2})^T \)

\( = P P^T \) dengan \( P = U D^{1/2} \)

\( U \) dan \( D^{1/2} \) nonsingular sehingga matrik \( P \) nonsingular.

---

**Material Penunjang**
(⇐) Diketahui matrik nonsingular \( P \) sedemikian sehingga \( A = PP^T \)

sehingga \( x^T A x = x^T PP^T x \)

\[ = (P^T x)^T P^T x \]

\[ = y^T y \text{ dengan } y = P^T x \]

Sehingga \( x^T A x > 0 \) untuk semua \( x \neq 0 \) dan \( x^T A x = 0 \) jika dan hanya jika \( P^T x = 0 \) dan karena \( P \) nonsingular \( (P^{-1})^T \) maka \( P^T x = 0 \) hanya jika \( x = 0 \).

Terbukti \( A \) definit positif.

(iii) Diketahui bahwa \( A \) definit positif maka menurut bagian (i) diatas ada \( P \)

sedemikian sehingga \( A = PP^T \).

Maka \( A^{-1} = (PP^T)^{-1} \)

\[ = (P^T)^{-1} P^{-1} \]

\[ = (P^{-1})^T P^{-1} \]

\[ = Q Q^T \text{ dengan } Q = (P^{-1})^T \text{ nonsingular}. \]

Terbukti \( A^{-1} \) definit positif.

(iii) Dari bagian (i) dan (ii) diketahui \( A = PP^T \) dan \( Q = (P^{-1})^T \) sehingga

\[ Q^T A Q = Q^T PP^T Q \]

\[ = P^{-1} P P^T (P^{-1})^T \]

\[ = I \]

\textit{Teorema 2.1.6}

Jika matrik \( X_{nxp} \) dengan rank \( p \) disebut matrik rank kolom penuh maka \( X^T X \) definit positif.

Bukti:

\[ z^T X^T X z = (X z)^T X z \]

\[ = y^T y \text{ dengan } y = X z \]

Jelas \( y^T y \geq 0 \), ambil \( y^T y = 0 \), \( y^T y = 0 \) jika dan hanya jika \( y = 0 \)

\textit{Material Penunjang}
Jadi \( y = xz = 0 \) dan karena \( X \) matrik rank kolom penuh sehingga \( Xz = 0 \) hanya mempunyai penyelesaian trivial yakni \( z = 0 \). Maka \( z^T X^T Xz > 0, \forall z \neq 0 \).

Terbukti \( X^T X \) definit positif.

Teorema 2.1.7
Matrik simetris \( A \) adalah semidefinit positif jika dan hanya jika \( \text{tr}(CA) \geq 0 \) untuk semua \( C \) semidefinit positif.

Bukti :
\( (\Rightarrow) \) Diketahui \( A \) semidefinit positif sehingga
\[
A = U D U^T
\]
\[
= \sum_{i=1}^{n} \lambda_i u_i u_i^T
\]
dengan \( D = \) diagonal \( (\lambda_1, ..., \lambda_n) \), \( \lambda_i \) adalah nilai eigen dari \( A \) \( (\forall \lambda_i \geq 0) \) dan \( U \) adalah matrik ortogonal.

\[
\text{tr}(CA) = \text{tr} \left( \sum_{i=1}^{n} \lambda_i C u_i u_i^T \right)
\]
\[
= \sum_{i=1}^{n} \lambda_i u_i^T C u_i
\]
\( u_i^T C u_i \) adalah nonnegatif untuk semua \( C \) semidefinit positif sehingga \( \text{tr}(CA) \geq 0 \).

\( (\Leftarrow) \) jika \( \text{tr}(CA) \geq 0 \) untuk semua \( C \) semidefinit positif benar untuk \( C = u_i u_i^T \geq 0 \) maka

\[
\text{tr}(CA) = \text{tr} \left[ u_i u_i^T \left( \sum_{j=1}^{n} \lambda_j u_j u_j^T \right) \right]
\]
\[
= \lambda_i \geq 0
\]

Terbukti \( A \) semidefinit positif.
2.1.5 Penentuan maksimum dan minimum

**Definisi 2.1.15**

Misalkan \( f(x) \) adalah fungsi kontinue dari elemen-elemen vektor \( x = [x_1, x_2, ..., x_p]^T \)
maka derivatif parsial pertama \( f \) terhadap \( x \) didefinisikan

\[
\frac{\partial f}{\partial x} = \left[ \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_p} \right]_{x_1}
\]

dan derivatif parsial kedua didefinisikan

\[
H = \frac{\partial^2 f}{\partial x \partial x^T} = \left[ \frac{\partial^2 f}{\partial x_i \partial x_j} \right]_{i,j}
\]

dan \( H \) disebut matrik Hess.

**Definisi 2.1.16**

Jika \( f(X) \) adalah sebuah fungsi skalar riel dari matrik \( X = [x_{ij}]_{mn} \) maka derivatif
parsial pertama \( f \) terhadap \( X \) didefinisikan

\[
\frac{\partial f}{\partial X} = \left[ \frac{\partial f}{\partial x_{ij}} \right]_{mn} = \left[ \frac{\partial f}{\partial x_{i1}}, \frac{\partial f}{\partial x_{i2}}, ..., \frac{\partial f}{\partial x_{im}} \right]_{i,j}
\]

Suatu syarat perlu untuk maksimum atau minimum dari \( f(x) \) di \( x = x_0 \)
adalah \( \frac{\partial f(x)}{\partial x} = 0 \)
dan \( x_0 \) yang memenuhi persamaan tersebut diatas disebut titik ekstrem.

---

*Material Penunjang*
Syarat cukup untuk nilai maksimum di titik \( x_0 \), yakni bahwa matrik Hess yang dievaluasi di \( x_0 \) adalah definit negatif dan matrik Hess definit positif adalah syarat cukup untuk nilai minimum.

Jika fungsi \( f(x) \) dengan adanya pembatas \( g(x)=0 \) maka untuk mendapatkan nilai maksimum atau minimum digunakan metode Lagrange Multiplier. Secara khusus, kita bentuk suatu fungsi baru

\[
h(x, \lambda) = f(x) - \lambda [g(x) - c]
\]

yang disebut fungsi lagrange.

Syarat perlu \( f(x) \) minimum atau maksimum jika dipenuhi

\[
\frac{\partial h(x, \lambda)}{\partial x} = \frac{\partial f(x)}{\partial x} - \lambda \frac{\partial g(x)}{\partial x} = 0
\]

\[
\frac{\partial h(x, \lambda)}{\partial \lambda} = -g(x) + c = 0
\]

dan \( x_0 \) yang memenuhi persamaan tersebut diatas disebut titik ekstrim.

Syarat cukup untuk nilai maksimum jika \( \frac{\partial h}{\partial xx^T} \) definit negatif dan minimum jika definit positif.

2.2. DISTRIBUSI NORMAL

**Definisi 2.2.1**

Jika vektor acak \( x = [x_1, x_2, ..., x_n]^T \sim \mathcal{N}(\mu, V) \) didefinisikan fungsi densitas

\[
f(x_1, x_2, ..., x_n) = \frac{\exp\left(-\frac{1}{2} (x - \mu)^T V^{-1} (x - \mu)\right)}{(2\pi)^{n/2} |V|^{1/2}}
\]

dengan vektor rata-rata dari \( x \) adalah \( \mu \) dan matrik var-cov\( (x) = V \) definit positif.
**Lemma 2.2.1**

Jika matrik $A_{nm}$ adalah matrik definit positip maka

$$
\int \cdots \int \exp\left(-\frac{1}{2}x^T A x\right) \, dx_1 \cdots dx_n = (2\pi)^{\frac{n}{2}} |A|^{-\frac{1}{2}}
$$

Bukti:

Karena $A$ definit positip, maka ada matrik non-singular $Q$ sedemikian sehingga $Q^T AQ = I_n$. Sehingga dengan sifat determinan $|Q^T AQ| = |Q|^2 |A| = 1$ dan $|Q| = |A|^{\frac{1}{2}}$

Ambil $x = Qy$ sehingga $x^T A x = y^T Q^T A Q y = y^T y$ dengan demikian:

$$
\int \cdots \int \exp\left(-\frac{1}{2}x^T A x\right) \, dx_1 \cdots dx_n = \int \cdots \int \exp\left(-\frac{1}{2}y^T y\right) \, dy_1 \cdots dy_n |Q^2|
$$

$$
= |Q| \int \cdots \int \exp\left(\frac{1}{2} \sum_{i=1}^{n} y_i^2\right) \, dy_1 \cdots dy_n
$$

$$
= |A|^{\frac{1}{2}} \prod_{i=1}^{n} \left\{ \int \exp\left(-\frac{1}{2} y_i^2\right) \, dy_i \right\}
$$

$$
= (2\pi)^{\frac{n}{2}} |A|^{-\frac{1}{2}}
$$

**Teorema 2.2.1**

Jika vektor acak $x = [x_1, ..., x_n]^T \sim N(\mu, V)$ maka fungsi pembangkit momen (m.g.f),

$$
M_x(t) = \exp(t^T \mu - \frac{1}{2} \mu^T V^{-1} \mu)
$$

Bukti:

$$
M_x(t) = E(\exp(t^T x)) = \frac{1}{\sqrt{(2\pi)^n |V|}} \int \cdots \int \exp[t^T x - \frac{1}{2}(x - \mu)^T V^{-1} (x - \mu)] \, dx_1 \cdots dx_n
$$
Distribusi Normal

\[= \frac{1}{\sqrt{(2\pi)^n |V|}} \int \cdots \int \exp\left[-\frac{1}{2}(x - \mu)^T V^{-1}(x - \mu) + t^T \mu + \frac{1}{2} t^T V t\right] dx_1 \cdots dx_n\]

\[= \frac{1}{(2\pi)^{\frac{n}{2}} |V|^{\frac{1}{2}}} \int \cdots \int \exp\left[-\frac{1}{2}(x - \mu - Vt)^T V^{-1}(x - \mu - Vt)\right] dx_1 \cdots dx_n\]

\[= \exp(t^T \mu - \frac{1}{2} \mu^T V^{-1} \mu)\]

**Teorema 2.2.2**

Jika vektor acak \( x = [x_1, x_2, \ldots, x_n]^T \sim N(\mu, V) \) dan matrik \( C_{mx} \) dengan rank \( p \), maka

\( Cx \sim N(C\mu, CVCT) \)

Bukti :

Misal \( y = Cx \), maka semua \( t \) riel, m.g.f \( y \) adalah :

\[M_y(t) = E[\exp(t^Ty)]\]

\[= E[\exp(t^T Cx)]\]

\[= E[\exp(s^T x)] \; \text{dengan} \; s^T = t^T C\]

\[= \exp(s^T \mu + 1/2 s^T V s)\]

\[= \exp(t^T (C\mu) + 1/2 t^T (CVCT)t)\]

Jadi terbukti \( Cx \sim N(C\mu, CVCT) \)

**Teorema 2.2.3**

Nilai ekspektasi dari bentuk kuadratik \( x^T Ax \) dengan \( x \sim N(\mu, V) \) adalah :

\[E(x^T Ax) = \text{tr}(AV) + \mu^T A\mu\]

Materi Penunjang
Bukti:

\[ E(x^T A x) = E(\text{tr}(x^T A x)) \]
\[ = E(\text{tr}(A x x^T)) \]
\[ = \text{tr}(E(A x x^T)) \]
\[ = \text{tr}(A E(x x^T)) \]

dan \( V = E((x - \mu)(x - \mu)^T) = E(x x^T) - \mu \mu^T \). Sehingga didapatkan

\[ E(x^T A x) = \text{tr}(A (V + \mu \mu^T)) \]
\[ = \text{tr}(A V + A \mu \mu^T) \]
\[ = \text{tr}(A V) + \mu^T A \mu \]

**Definisi 2.2.2**

Jika vektor acak \( x \sim N(0, I_n) \) maka \( u = x^T x \) berdistribusi chi-kuadrat central dengan derajat bebas \( n \) dinotasikan \( u \sim \chi^2_n \) dengan fungsi densitas,

\[ f(u) = \frac{u^{\frac{1}{2}n-1} \exp(-\frac{1}{2}u)}{2^{\frac{1}{2}n} \Gamma(\frac{1}{2}n)} \]

dengan \( \Gamma(\frac{1}{2}n) \) adalah fungsi gamma argumen \( \frac{1}{2} n \).

Dan jika \( x \sim N(\mu, I_n) \) maka \( u = x^T x \) berdistribusi chi-kuadrat nonsentral dengan derajat bebas \( n \) dinotasikan \( u \sim \chi^2(n, \lambda) \) dengan fungsi densitas,

\[ f(u) = \exp(-\lambda) \sum_{k=0}^{\infty} \frac{\lambda^k u^{\frac{1}{2}n+k-1} \exp(-\frac{1}{2}u)}{2^{\frac{1}{2}n+k} \Gamma(\frac{1}{2}n + k)} \]

dengan parameter non-central \( \lambda = \frac{1}{2} \mu^T \mu \) dan \( \Gamma(\frac{1}{2}n + k) \) adalah fungsi gamma dengan argumen \( \frac{1}{2} n + k \).
Teorema 2.2.4
Jika \( u \sim \chi^2_n \) maka fungsi pembangkit momen \( M_u(t) = (1-2t)^{\frac{n}{2}} \) dan jika \( u \sim \chi^2(n, \lambda) \) maka fungsi pembangkit momen \( M_u(t) = (1-2t)^{\frac{n}{2}} \exp(-\lambda(1-(1-2t)^{-1})) \).

Bukti:
Diketahui \( u \sim \chi^2_n \) maka
\[
M_u(t) = \int_0^\infty \exp \left( tu \right) \frac{u^{\frac{n}{2}+1}}{2^{\frac{n}{2}} \Gamma \left( \frac{n}{2} + 1 \right)} \, du
\]
\[
= \frac{1}{2^{\frac{n}{2}} \Gamma \left( \frac{n}{2} + 1 \right)} \int_0^\infty \exp \left( -\left( \frac{1}{2} - t \right)u \right) u^{\frac{n}{2}+1} \frac{\left( u \right)^{\frac{n}{2}+1}}{\Gamma \left( \frac{n}{2} + 1 \right)} \, du
\]
\[
= \frac{\Gamma \left( \frac{n}{2} + 1 \right)}{2^{\frac{n}{2}} \Gamma \left( \frac{n}{2} + 1 \right)} \left( \frac{1}{2} - t \right)^{\frac{n}{2}}
\]
\[
= (1-2t)^{\frac{n}{2}}
\]
Diketahui \( u \sim \chi^2(n, \lambda) \) maka
\[
M_u(t) = \exp(-\lambda) \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \quad \text{(mgf. } \chi^2_{n+2k} \text{)}
\]
\[
= \exp(-\lambda) \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} (1-2t)^{-n/2-k}
\]
\[
= \exp(-\lambda(1-(1-2t)^{-1})(1-2t)^{-n/2})
\]

Lemma 2.2.2
Untuk vektor \( g \) dan matrik definit positip \( W \)
\[
(2\pi)^\frac{n}{2} |W|^{-\frac{1}{2}} \exp(\frac{1}{2}g^TWg) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp(-\frac{1}{2}x^TWx + g^Tx) \, dx_1 \cdots dx_n
\]

Materi Penunjang
Bukti:

Diketahui dari lemma 2.2.1 \( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(-\frac{1}{2}x^TAx) \, dx_1 \ldots dx_n = (2\pi)^{\frac{n}{2}} |A|^{\frac{1}{2}} \)

Sehingga \( (2\pi)^{\frac{n}{2}} |W|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[-\frac{1}{2}(x - \mu)^TW^{-1}(x - \mu)] \, dx_1 \ldots dx_n \)

\[= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[-\frac{1}{2} x^TW^{-1}x + \mu^TW^{-1}x - \frac{1}{2} \mu^TW^{-1}\mu] \, dx_1 \ldots dx_n \]

\[= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[-\frac{1}{2} x^TW^{-1}x + g^Tx - \frac{1}{2} g^TWg] \, dx_1 \ldots dx_n \]

\[= \frac{1}{\exp(\frac{1}{2} g^TWg)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[-\frac{1}{2} x^TW^{-1}x + g^Tx] \, dx_1 \ldots dx_n \]

dengan \( g^T = \mu^TW^{-1} \)

**Teorema 2.2.5**

Untuk \( x \sim N(\mu, V) \), jika \( AV \) adalah idempoten, maka

\( x^TAx \sim \chi^2 [r(A), \frac{1}{2} \mu^TA\mu] \).

Bukti:

Fungsi pembangkit momen dari \( x^TAx \) adalah:

\[ M_{x^TAx}(t) = \frac{1}{\sqrt{(2\pi)^n|V|}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[tx^TAx - \frac{1}{2}(x - \mu)^TV^{-1}(x - \mu)] \, dx_1 \ldots dx_n \]

\[= \frac{\exp(-\frac{1}{2} \mu^TV^{-1}\mu)}{(2\pi)^{\frac{n}{2}}|V|^\frac{1}{2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(-\frac{1}{2}x^T(I-2tAV)V^{-1}x + \mu^TV^{-1}x) \, dx_1 \ldots dx_n \]

dari lemma 2.2.1 ambil \( g^T = \mu^TV^{-1} \) dan \( W = [(I-2tAV)^TV^{-1}\mu]_{i} = V(I-2tAV)_{i} \)

didapatkan:

---

**Material Penunjang**
\[ M_{x^T Ax}(t) = \exp(-\frac{1}{2} \mu^T V^{-1} \mu) \left| V \right|^{-\frac{1}{2}} \left| V(I - 2tAV)^{-1} \right|^{-\frac{1}{2}} \exp[\frac{1}{2} \mu^T V^{-1} V(I - 2tAV)^{-1} V^{-1} \mu] \]
\[ = \left| I - 2tAV \right|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \mu^T [I - (I - 2tAV)^T] V^{-1} \mu \right) \]
\[ = \prod_{i=1}^{n}(1 - 2t\lambda_i)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \mu^T \left[-\sum_{k=1}^{n}(2t)^k(AV)^k \right] V^{-1} \mu \right) \]

dengan \( \lambda_i \) untuk \( i = 1, 2, ..., n \) adalah akar karakteristik dari \( AV \). Jika \( AV \) adalah idempoten dengan rank \( r \), maka sebanyak \( r \) akar karakteristiknya bernilai 1 dan sebanyak \( n-r \) bernilai 0 dan \( (AV)^r = AV \). Oleh karena itu
\[ M_{x^T Ax}(t) = \prod_{i=1}^{r}(1 - 2t)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \mu^T [1 - (1 - 2t)^{-1}] A \mu \right) \]
\[ = (1 - 2t)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \mu^T A \mu \left[1 - (1 - 2t)^{-1}\right] \right) \]
 sehingga terlihat bahwa \( x^T Ax \sim \chi^2(r, \frac{1}{2} \mu^T A \mu) \) dengan \( r = r(AV) \).
Karena \( V \) nonsingular, \( r(AV) = r(A) \), jadi \( x^T Ax \sim \chi^2(r(A), \frac{1}{2} \mu^T A \mu) \).

2.3. MODEL REGRESI LINIER DAN PENAKSIRAN PARAMETER

Model regresi linier berbentuk:
\[ y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{12} + \beta_k x_{ik} + \varepsilon_t \]

disebut model regresi linier berganda dengan \( k \) variabel bebas \( x \) dan variabel tak bebas \( y \). Parameter \( \beta_j, j = 1, 2, ..., k \) disebut koefisien regresi dan kesalahan acak \( \varepsilon_t \).

2.3.1. Metode Kuadrat Terkecil Biasa (Ordinary least square)

Metode kuadrat terkecil biasa (metode OLS) dikemukakan oleh Carl Friedrich Gauss seorang ahli matematik bangsa Jerman. Dengan asumsi-asumsi tertentu yang
sering disebut sebagai asumsi standart/klasik, metode OLS mempunyai beberapa sifat statistik yang sangat menarik yang membuatnya menjadi satu metode analisa regresi yang kuat dan popular. Asumsi standart tersebut adalah sebagai berikut:

1. Nilai rata-rata dari unsur kesalahan acak \( \varepsilon_i \) adalah nol atau \( E(\varepsilon_i) = 0 \) untuk tiap \( i \), dalam notasi matrik \( E(\varepsilon) = 0 \) dengan \( \varepsilon \) dan \( 0 \) adalah vektor kolom \( nx1 \), dan 0 merupakan vektor nol.

2. Varian bersyarat dari \( \varepsilon_i \) adalah konstan atau homoskedastik.

\[
\text{Var}(\varepsilon_i) = E(\varepsilon_i - E(\varepsilon_i))^2
\]

\[
= E(\varepsilon_i^2) \text{ karena asumsi 1}
\]

\[
= \sigma^2
\]

3. Tidak ada autokorelasi dalam kesalahan acak.

\[
\text{Cov}(\varepsilon_i, \varepsilon_j) = E((\varepsilon_i - E(\varepsilon_i))(\varepsilon_j - E(\varepsilon_j)))
\]

\[
= E(\varepsilon_i \varepsilon_j) \text{ karena asumsi 1}
\]

\[
= 0 \text{ untuk } i \neq j
\]

Dalam notasi matriks asumsi 2 dan asumsi 3 disajikan \( E(\varepsilon \varepsilon^T) = \sigma^2 I_n \).

4. Variabel bebas \( X \) adalah nonstokastik (yaitu, tetap dalam penyampelan berulang) atau, jika stokastik, didistribusikan secara independen dari kesalahan acak \( \varepsilon_i \).

5. Tidak ada multikolineritas di antara variabel bebas \( X \).

Sehingga matrik \( X \) adalah matrik rank kolom penuh.

---

**Materi Penunjang**
Misal terdapat sebanyak \( n \) observasi, \( n > k \) dan misal \( x_{ij} \) menyatakan observasi ke-\( i \) pada tingkat variabel \( x_j \) dengan asumsi-asumsi diatas, selanjutnya dengan metode OLS untuk mendapatkan penaksiran model persamaan (2.3.1) dinyatakan sebagai berikut:

\[
y_i = \hat{\beta}_0 + \sum_{j=1}^{k} \hat{\beta}_j x_{ij} + e_i
\]

...... (2.3.2)

dengan \( e_i \) adalah unsur sesatan, yang merupakan kesalahan acak untuk sampel.

Selanjutnya meminimumkan jumlah kuadrat sesatan

\[
\sum_{i=1}^{n} e_i^2 = \sum (y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij})^2
\]

...... (2.3.3)

\[\sum_{i=1}^{n} e_i^2 \text{ diminimumkan terhadap } \hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_k .\]

\[
\frac{\partial (\sum e_i^2)}{\partial \hat{\beta}_0} = -2 \sum (y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij}) = 0
\]

...... (2.3.4)

dan

\[
\frac{\partial (\sum e_i^2)}{\partial \hat{\beta}_j} = -2 \sum (y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij}) x_{ij} = 0
\]

...... (2.3.5)

dan diperoleh persamaan-persamaan normal berikut.

\[
n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} + ... + \hat{\beta}_k \sum_{i=1}^{n} x_{ik} = \sum_{i=1}^{n} y_i
\]

\[
\hat{\beta}_0 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1}^2 + ... + \hat{\beta}_k \sum_{i=1}^{n} x_{ik} x_{ij} = \sum_{i=1}^{n} x_{ij} y_i
\]

...... (2.3.6)

\[
\hat{\beta}_0 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_1 \sum_{i=1}^{n} x_{i2} + ... + \hat{\beta}_k \sum_{i=1}^{n} x_{ik}^2 = \sum_{i=1}^{n} x_{ik} y_i
\]

\[
\vdots
\]

\[
\hat{\beta}_0 \sum_{i=1}^{n} x_{in} + \hat{\beta}_1 \sum_{i=1}^{n} x_{in}^2 + ... + \hat{\beta}_k \sum_{i=1}^{n} x_{ink} = \sum_{i=1}^{n} x_{kn} y_i
\]
Diberikan sebuah pengembangan matrik dari persamaan normal di atas. Model persamaan (2.3.2) dapat ditulis dalam bentuk matrik sebagai berikut:

\[
\begin{bmatrix}
  y_1 \\
  y_2 \\
  \vdots \\
  y_n \\
\end{bmatrix} = \begin{bmatrix}
  1 & x_{11} & \cdots & x_{1k} \\
  1 & x_{21} & \cdots & x_{2k} \\
  \vdots & \vdots & \ddots & \vdots \\
  1 & x_{n1} & \cdots & x_{nk} \\
\end{bmatrix} \begin{bmatrix}
  \beta_0 \\
  \beta_1 \\
  \vdots \\
  \beta_k \\
\end{bmatrix} + \begin{bmatrix}
  e_1 \\
  e_2 \\
  \vdots \\
  e_n \\
\end{bmatrix}
\]

atau 
\[ y = X\hat{\beta} + e \] 
\[ \cdots (2.3.7) \]
dengan y adalah vektor (n×1) dari pengamatan variabel tak bebas, X adalah matrik (n×p) dari n pengamatan k variabel bebas (p=k+1) dengan r(X) = p(p≥p), e adalah vektor sesatan (n×1), dan \( \hat{\beta} \) adalah vektor (p×1) dari penaksir koefisien regresi.

Metode Penaksiran OLS yaitu dengan meminimumkan jumlah kuadrat sesatan:

\[ \sum_{i=1}^{n} e_i^2 = e^T e = (y - X\hat{\beta})^T (y - X\hat{\beta}) \]
\[ = y^T y - \hat{\beta}^T X^T y - y^T X \hat{\beta} + \hat{\beta}^T X^T X \hat{\beta} \]
\[ = y^T y - 2\hat{\beta}^T X^T y + \hat{\beta}^T X^T X \hat{\beta} \]

\[ \frac{\partial (e^T e)}{\partial \hat{\beta}} = -2X^T y + 2X^T X \hat{\beta} = 0 \]
\[ X^T X \hat{\beta} = X^T y \]
\[ \cdots (2.3.8) \]

\[ \frac{\partial^2 L}{\partial \hat{\beta}^2} = 2X^T X \]
\[ \cdots (2.3.9) \]
Persamaan (2.3.8) adalah persamaan normal kuadrat terkecil. Persamaan normal tersebut identik dengan persamaan (2.3.6). Menurut teorema 2.1.6 matrik $X^TX$ definit positif sehingga $L$ minimal dan $(X^TX)^{-1}$ ada sehingga

$$\hat{\beta} = (X^TX)^{-1}X^Ty$$

.....(2.3.10)

disebut penaksir kuadrat terkecil biasa (OLS).

2.3.3. Penaksir linier tak bias terbaik (Best linier unbiased estimator)

Suatu penaksir adalah Penaksir linier tak bias terbaik (BLUE) jika penaksir tersebut merupakan penaksir terbaik, linier dan tak bias. Suatu penaksir dikatakan penaksir terbaik jika dalam klas penaksir linier tak bias adalah penaksir yang mempunyai varians minimum.

Definisi 2.3.1

Suatu statistik $\hat{\beta}$ disebut penaksir tak bias dari parameter $\beta$ jika $E(\hat{\beta}) = \beta$.

Dapat ditunjukkan bahwa penaksir OLS (2.3.10) adalah penaksir tak bias untuk $\beta$, sebagai berikut:

$$E(\hat{\beta}) = E((X^TX)^{-1}X^Ty)$$

$$= E((X^TX)^{-1}X^T(X\beta + \varepsilon))$$

$$= E((X^TX)^{-1}X^TX\beta + (X^TX)^{-1}X^T\varepsilon)$$

$$= \beta$$
Dengan mensubstitusikan \( y = X\beta + \varepsilon \) kedalam persamaan (2.3.10) didapatkan
\[
\hat{\beta} = (X^TX)^{-1}X^T(X\beta + \varepsilon)
\]
\[= \beta + (X^TX)^{-1}X^T\varepsilon
\]
Oleh karena itu \( \hat{\beta} - E(\hat{\beta}) = \hat{\beta} - \beta = (X^TX)^{-1}X^T\varepsilon \), sehingga akan didapatkan
\[
\text{var-cov}(\hat{\beta}) = E((\hat{\beta} - E(\hat{\beta}))(\hat{\beta} - E(\hat{\beta}))^T)
\]
\[= E((X^TX)^{-1}X^T\varepsilon)(X^TX)^{-1}X^T\varepsilon^T)
\]
\[= (X^TX)^{-1}X^T \varepsilon \varepsilon^T X(X^TX)^{-1}
\]
\[= (X^TX)^{-1}X^T \sigma^2 I_n X(X^TX)^{-1}
\]
\[= \sigma^2 (X^TX)^{-1}
\]

**Definisi 2.3.2**

\( y \) adalah vektor random. Misal \( C \) dan \( c \) adalah matrik dan vektor dengan ukuran yang sesuai, sehingga vektor
\[
\hat{\beta} = Cy + c
\]
sebagai penaksir linier dari \( \beta \). \( \hat{\beta} \) adalah penaksir linier homogen jika \( c = 0 \).

**Definisi 2.3.3**

Rata-rata sesatan kuadrat (Mean Square Error) suatu penaksir \( \hat{\beta} \) didefinisikan
\[
\text{MSE}(\hat{\beta}) = E(\hat{\beta} - \beta)(\hat{\beta} - \beta)^T
\]

MSE untuk penaksir linier homogen \( \hat{\beta} = Cy \) adalah sebagai berikut:

---

**Materi Penunjang**
\[ 
\text{MSE}(\hat{\beta}) = E(\hat{\beta} - \beta)(\hat{\beta} - \beta)^T \\
= E \left[ ((CX - I)\beta + C\epsilon)((CX - I)\beta + C\epsilon)^T \right] \\
= \sigma^2 CC^T + ((CX-I)\beta)((CX-I)\beta)^T 
\]

dengan \( \hat{\beta} - \beta = (CX - I)\beta + C\epsilon \) ; bias \( = E(\hat{\beta}) - \beta = (CX - I)\beta \) dan \( \text{Var-cov}(\hat{\beta}) = \sigma^2 CC^T \)

**Definisi 2.3.4**

Skalar rata-rata sesatan kuadrat (scalar mean square error) suatu penaksir \( \hat{\beta} \) didefinisikan

\[ 
\text{SMSE}(\hat{\beta}) = E(\hat{\beta} - \beta)^T(\hat{\beta} - \beta) 
\]

dengan operator trace diperoleh \( \text{SMSE}(\hat{\beta}) = \text{tr} E(\hat{\beta} - \beta)(\hat{\beta} - \beta)^T = \text{tr MSE}(\hat{\beta}) \)

**Definisi 2.3.5**

Kerugian sesatan kuadrat penaksir \( \hat{\beta} \) terhadap \( \beta \) adalah

\[ 
L(\hat{\beta}, A) = (\hat{\beta} - \beta)^T A (\hat{\beta} - \beta) 
\]

dengan \( A \) adalah matrik definit positif dengan ukuran \( (p \times p) \).

**Definisi 2.3.2**

Resiko kuadrat dari penaksir \( \hat{\beta} \) adalah

\[ 
R(\hat{\beta}, A) = E L(\hat{\beta}, A) \\
= E (\hat{\beta} - \beta)^T A (\hat{\beta} - \beta) 
\]
Dipergunakan operator trace didapat hubungan antara resiko kuadrat \( \mathbf{R}(\hat{\beta}, \mathbf{A}) \) dan rata-rata sesatan kuadrat MSE sebagai berikut:

\[
\begin{align*}
\mathbf{R}(\hat{\beta}, \mathbf{A}) &= \text{tr} \mathbf{A} \mathbf{E}(\hat{\beta} - \beta)(\hat{\beta} - \beta)^T \\
&= \text{tr} (\mathbf{A} \text{MSE}(\hat{\beta}))
\end{align*}
\]

hal ini memperlihatkan ekuivalennya peminimalan resiko \( \mathbf{R}(\hat{\beta}, \mathbf{A}) \) atau \( \text{MSE}(\hat{\beta}) \) untuk penaksir \( \hat{\beta} \).

**Teorema 2.3.1**

Jika diberikan dua estimator \( \beta_1^* \) dan \( \beta_2^* \) maka dua pernyataan ini ekuivalen

(i) \( \text{MSE}(\beta_1^*) - \text{MSE}(\beta_2^*) \geq 0 \)

(ii) \( \mathbf{R}(\beta_1^*, \mathbf{A}) - \mathbf{R}(\beta_2^*, \mathbf{A}) \geq 0 \)

Bukti:

\[
\begin{align*}
\mathbf{R}(\beta_1^*, \mathbf{A}) - \mathbf{R}(\beta_2^*, \mathbf{A}) &= \text{tr} \mathbf{A} \{ \text{MSE}(\beta_1^*) - \text{MSE}(\beta_2^*) \} \\
&= \text{tr} \mathbf{A} \Delta \\
dengan \( \Delta = \text{MSE}(\beta_1^*) - \text{MSE}(\beta_2^*) \)
\end{align*}
\]

menurut teorema 2.1.7 \( \text{tr} \mathbf{A} \Delta \geq 0 \) untuk semua \( \mathbf{A} \geq 0 \) jika dan hanya jika \( \Delta \geq 0 \).
**Teorema 2.3.2** *(teorema gaus-Markov)*

Model regresi (2.3.7)

\[ y = X\beta + \varepsilon \sim (0, \sigma^2 I_n), \quad r(X) = p \]

penaksir OLS (2.3.10) \( \hat{\beta} = (X^TX)^{-1}Xy \) dengan \( \text{Var-cov}(\hat{\beta}) = \sigma^2 (X^TX)^{-1} \) adalah penaksir BLUE dari \( \beta \).

**Bukti:**

Penaksir OLS (2.3.9) adalah penaksir linier homogen \( \hat{\beta} = (X^TX)^{-1}X^Ty = C_0y \) dengan \( C_0 = (X^TX)^{-1}X^T \) dan merupakan penaksir linier tak bias.

Dalam bentuk umum penaksir linier homogen \( \hat{\beta} = Cy \) dengan \( \hat{\beta} - \beta = (CX-I)\beta + C\varepsilon \) dan bias \( E(\hat{\beta}) - \beta = (CX-I)\beta \). Misalkan penaksir \( \hat{\beta} \) tak bias, \( E(\hat{\beta}) = \beta \) ekuivalen untuk \( CX = I \) dapat ditulis baris demi baris \( CX = e_i \) dan didapatkan fungsi kuadrat resiko

\[ R(\hat{\beta}, I) = \text{tr} \text{MSE}(\hat{\beta}) = \sigma^2 \text{tr}(CC^T). \]

Selanjutnya variansi akan minimal dengan \( \min \{ \sigma^2 \text{tr}(CC^T) \mid CX = I \} \) atau equivalent meminimalkan fungsi lagrange:

\[ h(C, \lambda_i) = R(\hat{\beta}, I) - 2 \sum \lambda_i (e_i - e_i)^T \]

\[ = \sigma^2 \text{tr}(CC^T) - 2 \sum \lambda_i (e_i - e_i)^T \]

---

**Material Pendukung**

---
dengan \( \lambda_1 \) adalah vektor (px1) perkalian lagrange dan dimisalkan

\[
\Lambda = \begin{bmatrix}
\lambda_1^T \\
\vdots \\
\lambda_p^T
\end{bmatrix}
\]

fungsi lagrange \( h(C, \lambda_1) \) diferensialkan terhadap \( C \) dan \( \Lambda \) didapatkan persamaan normal sebagai berikut:

\[
\begin{align*}
\frac{\partial h}{\partial C} &= 2\sigma^2 C - 2\Lambda X^T = 0 \\
\frac{\partial h}{\partial \Lambda} &= CX - I = 0
\end{align*}
\]

\[\ldots \text{(2.3.12)}\]

dan \( \frac{\partial^2 h}{\partial CC^T} = 2\sigma^2 I \)

dengan menyelesaikan persamaan normal (2.3.12) tersebut didapatkan

\[
\Lambda = \sigma^2 (X^T X)^{-1}
\]

\[
C = (X^T X)^{-1} X^T
\]

dan karena \( \frac{\partial^2 h}{\partial CC^T} \) definit positif maka \( h(C, \lambda_1) \) minimal sehingga terbukti variansi minimal untuk \( C = C_0 \). Jadi penaksir OLS memiliki sifat BLUE.