BAB II

LANDASAN TEORI

Pada Bab II berikut akan dibahas mengenai dasar-dasar teori yang digunakan dalam penulisan Tugas Akhir ini, yaitu tentang *trace* matriks, *rank* matriks, determinan matriks, invers matriks, bentuk linier, kuadratik, beserta diferensiasinya. Kemudian, dibahas juga mengenai *General Linear Mixed Model* (*General* LMM) dengan metode penaksiran masing-masing parameternya. Selain itu, didefinisikan pula mengenai fungsi *translation-invariant* dan *maximal-invariant.*

2.1 Matriks

2.1.1 *Trace* Matriks

Misalkan $\mathbf{A} = \{a_{ij}\}$ suatu matriks persegi berukuran $n \times n$, maka *trace* dari **A** didefinisikan sebagai jumlah dari elemen diagonal $\mathbf{A}$ dan dinotasikan dengan $\text{tr}(\mathbf{A})$, yaitu $\text{tr}(\mathbf{A}) = a_{11} + a_{22} + \ldots + a_{nn}$. Beberapa sifat dari *trace* di antaranya adalah sebagai berikut:

1. $\text{tr}(\mathbf{I}_n) = n$

2. $\text{tr}\left[(k)\right] = k$

3. $\mathrm{tr}(k\mathbf{A}) = k\,\mathrm{tr}(\mathbf{A})$

4. $\mathrm{tr}(\mathbf{A}+\mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$

5. $\mathrm{tr}(\mathbf{A}^T) = \mathrm{tr}(\mathbf{A})$ **(2.1.1.a)**

di mana k adalah sembarang skalar, sedangkan $\mathbf{A}$ dan $\mathbf{B}$ adalah matriks berukuran $n \times n$. Untuk $k_1, k_2, \ldots, k_r$ adalah sembarang skalar dan $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_r$ adalah matriks berukuran $n \times n$ maka

$$\mathrm{tr}\left(\sum_{i=1}^{r} k_i \mathbf{A}_i\right) = \sum_{i=1}^{r} k_i\,\mathrm{tr}(\mathbf{A}_i).$$

Misalkan $\mathbf{A} = \{a_{ij}\}$ adalah matriks berukuran $m \times n$ dan $\mathbf{B} = \{b_{ij}\}$ adalah matriks berukuran $n \times m$, maka $\mathrm{tr}(\mathbf{AB}) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} b_{ji}$ di mana $\mathbf{AB} = \sum_{j=1}^{n} a_{ij} b_{ji}$. Berdasarkan (2.1.1.a) maka

$$\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{B}^T \mathbf{A}^T) \qquad \textbf{(2.1.1.b)}$$

sehingga dapat diperluas sebagaimana diberikan dalam lemma berikut ini.

Lemma 2.1 Untuk sembarang matriks $\mathbf{A}$ berukuran $m \times n$ dan matriks $\mathbf{B}$ berukuran $n \times m$ maka $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$.

Bukti:

$$\operatorname{tr}(\mathbf{AB}) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}b_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{m} b_{ji}a_{ij} = \operatorname{tr}(\mathbf{BA})$$

Catatan: Berdasarkan (2.1.1.b) dan lemma 2.1, untuk sembarang matriks $\mathbf{A}$ berukuran $m \times n$ dan matriks $\mathbf{B}$ berukuran $n \times m$, maka

$$\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}\left(\mathbf{B}^T\mathbf{A}^T\right) = \operatorname{tr}(\mathbf{BA}) = \operatorname{tr}\left(\mathbf{A}^T\mathbf{B}^T\right).$$

Bentuk di atas dapat diperluas untuk *trace* dari perkalian matriks $\mathbf{ABC}$ di mana $\mathbf{A}$ matriks berukuran $m \times n$, $\mathbf{B}$ matriks berukuran $n \times p$, dan $\mathbf{C}$ matriks berukuran $p \times m$ sehingga

$$tr(\mathbf{ABC}) = tr(\mathbf{CAB}) = tr(\mathbf{BCA}). \qquad \textbf{(2.1.1.c)}$$

2.1.2 *Rank* Matriks

Definisi 2.2 Himpunan bagian ***W*** dari ruang vektor ***V*** disebut subruang dari ***V*** jika ***W*** merupakan ruang vektor terhadap penjumlahan dan perkalian skalar yang didefinisikan pada ***V***.

Definisi 2.3 Suatu vektor **w** disebut kombinasi linier dari vektor-vektor $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ jika dapat dinyatakan ke dalam bentuk $\mathbf{w} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \ldots + k_r\mathbf{v}_r$ dengan $k_1, k_2, \ldots, k_r$ adalah skalar.

Definisi 2.4 Jika $\boldsymbol{S} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ adalah himpunan vektor-vektor tak kosong, maka persamaan vektor

$$k_1 \mathbf{v}_1 + k_2 \mathbf{v}_2 + \ldots + k_r \mathbf{v}_r = \mathbf{0}$$

memiliki paling tidak satu penyelesaian, yaitu:

$$k_1 = 0, \quad k_2 = 0, \quad \ldots, \quad k_r = 0 .$$

Jika ini adalah satu-satunya penyelesaian, maka $\boldsymbol{S}$ disebut himpunan yang bebas linier.

Definisi 2.5 Jika $\boldsymbol{S} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ adalah himpunan vektor dalam ruang vektor $\boldsymbol{V}$, maka subruang $\boldsymbol{W}$ dari $\boldsymbol{V}$ yang terdiri dari semua kombinasi linier dari vektor-vektor dalam $\boldsymbol{S}$ disebut ruang yang direntang oleh $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$. Vektor-vektor dalam himpunan $\boldsymbol{S} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ merupakan vektor-vektor merentang $\boldsymbol{W}$, sehingga $\boldsymbol{W}$ dapat dituliskan sebagai

$$\boldsymbol{W} = span(\boldsymbol{S}) \text{ atau } \boldsymbol{W} = span\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}.$$

Definisi 2.6 Jika ruang vektor $\boldsymbol{V}$ dan $\boldsymbol{S} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ adalah himpunan vektor-vektor dalam $\boldsymbol{V}$, maka $\boldsymbol{S}$ disebut basis untuk $\boldsymbol{V}$ jika dua syarat berikut ini dipenuhi:

1. $\boldsymbol{S}$ bebas linier
2. $\boldsymbol{S}$ merentang $\boldsymbol{V}$.

Definisi 2.7 Ruang vektor tidak nol $\boldsymbol{V}$ berdimensi hingga jika $\boldsymbol{V}$ memuat himpunan berhingga vektor-vektor $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ yang membentuk basis.

Definisi 2.8 Dimensi ruang vektor $\boldsymbol{V}$ berdimensi hingga, yang dinyatakan dengan $\dim(\boldsymbol{V})$, didefinisikan sebagai banyak vektor dalam suatu basis untuk $\boldsymbol{V}$.

Definisi 2.9 Untuk matriks **A** berukuran $m \times n$,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Vektor-vektor

$$\begin{aligned} \mathbf{r}_1 &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \end{bmatrix} \\ \mathbf{r}_2 &= \begin{bmatrix} a_{21} & a_{22} & \cdots & a_{2n} \end{bmatrix} \\ &\vdots \\ \mathbf{r}_m &= \begin{bmatrix} a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \end{aligned}$$

dalam $\Re^n$ yang dibentuk dari baris-baris **A** disebut vektor-vektor baris dari **A**, dan vektor-vektor

$$\mathbf{c}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \ldots, \quad \mathbf{c}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

dalam $\Re^m$ yang dibentuk dari kolom-kolom **A** disebut vektor-vektor kolom dari **A**.

Definisi 2.10 Jika matriks **A** berukuran $m \times n,$ maka subruang dari $\Re^n$ yang direntang oleh vektor-vektor baris dari **A** disebut ruang baris dari **A**, dan subruang dari $\Re^m$ yang direntang oleh vektor-vektor kolom disebut ruang kolom dari **A**.

Definisi 2.11 Dimensi dari ruang baris dan ruang kolom dari suatu matriks **A** disebut *rank* dari **A**.

Definisi 2.12 Matriks **A** berukuran $m \times n$ dikatakan *full rank* jika

$$r(\mathbf{A}) = min(m,n)$$

2.1.3 Invers Matriks

Misalkan **A** matriks *full rank* berukuran $n \times n$. **A** disebut *nonsingular* jika terdapat matriks $\mathbf{A}^{-1}$ sedemikian sehingga $\mathbf{A}\,\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.

Catatan: Invers dari suatu matriks bujur sangkar adalah tunggal.

2.1.4 Determinan dan *Adjoint* Matriks

Definisi 2.13 Suatu permutasi himpunan bilangan bulat $\{1,2,\ldots,n\}$ adalah suatu susunan bilangan-bilangan bulat ini ke dalam suatu urutan tanpa penghilangan atau pengulangan.

Misalkan $(j_1, j_2, \ldots, j_n)$ menyatakan suatu permutasi umum dari himpunan $\{1, 2, \ldots, n\}$ di mana j_1 adalah bilangan bulat pertama dalam permutasi, j_2 adalah bilangan bulat yang kedua, dan seterusnya.

Definisi 2.14 Suatu pembalikan dikatakan terjadi dalam suatu permutasi $(j_1, j_2, \ldots, j_n)$ bilamana suatu bilangan bulat yang lebih besar mendahului yang lebih kecil.

Total jumlah pembalikan yang terjadi dalam suatu permutasi dapat diperoleh sebagai berikut:

1. Cari jumlah bilangan bulat yang lebih kecil dari j_1 dan yang mengikuti j_1 dalam permutasi tersebut.
2. Cari jumlah bilangan bulat yang lebih kecil dari j_2 dan yang mengikuti j_2 dalam permutasi tersebut.
3. Teruskan proses menghitung ini untuk $j_3, j_4, \ldots, j_{n-1}$.

Total dari jumlah-jumlah ini adalah total jumlah pembalikan dalam permutasi tersebut.

Definisi 2.15 Suatu permutasi disebut genap jika total jumlah pembalikan merupakan suatu bilangan bulat genap dan disebut ganjil jika total jumlah pembalikan merupakan suatu bilangan bulat ganjil.

Misalkan $\mathbf{A}$ adalah suatu matriks berukuran $n \times n$, determinan dari matriks $\mathbf{A}$, dinotasikan dengan $|\mathbf{A}|$ atau $\det(\mathbf{A})$, didefinisikan sebagai

$$|\mathbf{A}| = \sum \pm\, a_{1j_1} a_{2j_2} \ldots a_{nj_n}$$

di mana $\sum$ menunjukkan bahwa suku-suku dijumlahkan atas semua permutasi $(j_1, j_2, \ldots, j_n)$, dan + serta – dipilih pada setiap suku, tergantung pada apakah permutasi tersebut genap (+) atau ganjil (-).

Beberapa sifat determinan, di antaranya:

1. $|\mathbf{A}^T| = |\mathbf{A}|$.
2. $|k\mathbf{A}| = k^n |\mathbf{A}|$.
3. $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$.
4. $|\mathbf{A}^T\mathbf{A}| = |\mathbf{A}|^2$.
5. $|\mathbf{A}^{-1}| = 1/|\mathbf{A}|$.

untuk sembarang matriks $\mathbf{A}$ dan $\mathbf{B}$ berukuran $n \times n$ serta skalar k.

Misalkan $\mathbf{A} = \{a_{ij}\}$ menyatakan matriks berukuran $n \times n$ dan $\mathbf{A}_{ij}$ menyatakan submatriks dari $\mathbf{A}$ berukuran $(n-1)\times(n-1)$ yang didapat dengan menghilangkan baris serta kolom yang mengandung elemen a_{ij}, yakni baris ke-*i* dan kolom ke-*j*. Determinan dari submatriks $|\mathbf{A}_{ij}|$ disebut minor dari elemen a_{ij}, sedangkan $(-1)^{i+j}|\mathbf{A}_{ij}|$ disebut kofaktor dari a_{ij}. Misalkan a_{ij} menyatakan elemen ke-*ij* dari matriks berukuran $n\times n$, dan misalkan α_{ij} menyatakan kofaktor dari a_{ij} $(i, j=1,2,\ldots,n)$. Maka, untuk $i=1,2,\ldots,n$,

$$|\mathbf{A}|=\sum_{j=1}^{n}a_{ij}\alpha_{ij}=a_{i1}\alpha_{i1}+\ldots+a_{in}\alpha_{in}$$

dan untuk $j=1,2,\ldots,n$,

$$|\mathbf{A}|=\sum_{i=1}^{n}a_{ij}\alpha_{ij}=a_{1j}\alpha_{1j}+\ldots+a_{nj}\alpha_{nj}.$$

Catatan: *Transpose* matriks kofaktor dari $\mathbf{A}$ disebut matriks *adjoint* dan dinotasikan dengan $adj(\mathbf{A})$, yaitu:

$$adj\,\mathbf{A}=\begin{pmatrix}\alpha_{11} & \cdots & \alpha_{1n}\\ \vdots & \ddots & \vdots\\ \alpha_{n1} & \cdots & \alpha_{nn}\end{pmatrix}^{T}=\begin{pmatrix}\alpha_{11} & \cdots & \alpha_{n1}\\ \vdots & \ddots & \vdots\\ \alpha_{1n} & \cdots & \alpha_{nn}\end{pmatrix}$$

di mana memenuhi sifat:

1. $\mathbf{A}\,adj(\mathbf{A})=(adj\,\mathbf{A})\mathbf{A}=|\mathbf{A}|\mathbf{I}_n$.

2. $adj(\mathbf{A})=|\mathbf{A}|\mathbf{A}^{-1}$ atau ekivalen dengan $\mathbf{A}^{-1}=(1/|\mathbf{A}|)adj(\mathbf{A})$. **(2.1.4.a)**

2.2 Notasi Bentuk Linier, Bentuk Kuadratik, dan Definit Positif

Misalkan $\mathbf{a} = (a_1, \ldots, a_n)^T$ adalah vektor kolom berdimensi *n.* Pandang fungsi $\mathbf{a}^T\mathbf{x} = \sum_i a_i x_i$, dengan vektor $\mathbf{x} = (x_1, \ldots, x_n)^T$ di $\Re^n$. Fungsi tersebut merupakan fungsi bentuk linier.

Misalkan matriks $\mathbf{A} = \left[a_{ij}\right]$ berukuran $n \times n$. Pandang fungsi

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{i,j} a_{ij} x_i x_j = \sum_i a_i x_i^2 + \sum_{i,j \neq i} a_{ij} x_i x_j$$

dengan vektor $\mathbf{x} = (x_1, \ldots, x_n)^T$ di $\Re^n$. Fungsi tersebut merupakan bentuk kuadratik.

Definisi 2.16 Misalkan $\mathbf{x}^T\mathbf{A}\mathbf{x}$ merupakan bentuk kuadratik dengan $\mathbf{A} = \left[a_{ij}\right]$ adalah matriks berukuran $n \times n$ dan vektor $\mathbf{x} = (x_1, \ldots, x_n)^T$ di $\Re^n$.

Bentuk kuadratik dikatakan definit positif jika

$$\mathbf{x}^T\mathbf{A}\,\mathbf{x} > 0, \text{ untuk setiap } \mathbf{x} \neq 0.$$

Definisi 2.17 Matriks simetris $\mathbf{A}$ disebut matriks definit positif jika $\mathbf{x}^T\mathbf{A}\mathbf{x}$ adalah bentuk kuadratik definit positif.

Berikut ini adalah lemma mengenai matriks definit positif yang merupakan matriks *nonsingular* (bukti diberikan pada lampiran 1).

Lemma 2.18 Sembarang matriks definit positif adalah *nonsingular*.

2.3 Diferensiasi Matriks

2.3.1 Definisi dan Notasi Diferensiasi Matriks

Sebelum membahas lebih jauh mengenai turunan atau diferensiasi pada matriks, ada beberapa hal yang perlu diketahui, di antaranya mengenai definisi, notasi, dan beberapa pendahuluan lainnya. Misalkan $\mathbf{c}$ menyatakan sembarang vektor kolom berdimensi m. *Neighborhood* dari $\mathbf{c}$ didefinisikan sebagai sebuah himpunan dengan bentuk umumnya $\left\{\mathbf{x} \in \Re^{m\times 1} : \|\mathbf{x}-\mathbf{c}\| < r\right\}$ di mana r adalah suatu bilangan positif yang disebut radius dari *neighborhood*. Oleh karena itu, *neighborhood* dari $\mathbf{c}$ dengan radius r adalah himpunan semua vektor kolom berdimensi m yang memiliki jarak terhadap $\mathbf{c}$ kurang dari r. Definisi tersebut dapat diperluas untuk vektor baris dan matriks. *Neighborhood* dari sembarang matriks berukuran $m \times n$ adalah sebuah himpunan dengan bentuk umumnya adalah $\left\{\mathbf{X} \in \Re^{m\times n} : \|\mathbf{X}-\mathbf{C}\| < r\right\}$.

Misalkan S menyatakan sembarang himpunan vektor kolom berdimensi m. Maka, vektor $\mathbf{x}$ di S disebut *interior point* dari S jika ada *neighborhood* dari $\mathbf{x}$ yang semua titiknya berada di S. Hubungan tersebut berlaku pula untuk sembarang matriks berukuran $m \times n$.

Misalkan suatu fungsi f bernilai skalar memiliki domain di sebuah himpunan $\Re^{m\times n}$. Nilai fungsi f dari vektor $\mathbf{x}$ atau matriks $\mathbf{X}$ dinotasikan dengan $f(\mathbf{x})$ atau $f(\mathbf{X})$. Misalkan $\mathbf{f}=\{f_s\}$ adalah vektor berukuran $p\times 1$ dan $\mathbf{F}=\{f_{st}\}$ adalah matriks berukuran $p\times q$ di mana setiap elemennya adalah fungsi yang terdefinisi di himpunan $\Re^{m\times n}$. Maka untuk sembarang matriks $\mathbf{X}$, vektor $\left[f_1(\mathbf{X}), f_2(\mathbf{X}),\ldots, f_p(\mathbf{X})\right]^T$ berukuran $p\times 1$ dan matriks dengan elemen ke-st adalah $f_{st}(\mathbf{X})$ dapat dinotasikan sebagai $\mathbf{f}(\mathbf{X})$ atau $\mathbf{F}(\mathbf{X})$ serta dapat dinyatakan sebagai nilai dari $\mathbf{f}$ atau $\mathbf{F}$ di $\mathbf{X}$.

Misalkan f menyatakan fungsi yang domainnya adalah himpunan S di $\Re^{m\times 1}$. Maka berdasarkan definisi, f dikatakan kontinu di sebuah *interior point* $\mathbf{c}$ dari S, jika $\lim_{\mathbf{x}\to\mathbf{c}} f(\mathbf{x}) = f(\mathbf{c})$ di mana $\lim_{\mathbf{x}\to\mathbf{c}} f(\mathbf{x})$ berupa skalar (misalkan: b). Dengan demikian, untuk setiap skalar ε, ada *neighborhood* N_ε dari $\mathbf{c}$ sedemikian sehingga $\left|f(\mathbf{x})-b\right|<\varepsilon$ di mana $\|\mathbf{x}-\boldsymbol{c}\|<\delta$ dengan δ adalah radius dari *neighborhood* N_ε. Sebuah matriks $\mathbf{F}$ berukuran $p\times q$ yang setiap *entri*-nya memiliki domain yang sama pada himpunan S di $\Re^{m\times 1}$ dapat dikatakan kontinu di sebuah *interior point* $\mathbf{c}$ dari S jika semua elemen dari $\mathbf{F}$ kontinu di $\mathbf{c}$.

Misalkan f menyatakan sebuah fungsi yang terdefinisi di sebuah himpunan S dari vektor $\mathbf{x}=(x_1, x_2,\ldots, x_m)^T$ dengan m variabel. Anggap bahwa

S mengandung beberapa *interior point* dan misalkan $\mathbf{c} = (c_1, c_2, \ldots, c_m)^T$ menyatakan salah satu dari titik-titik tersebut. Misalkan $\mathbf{u}_j$ menyatakan kolom ke- j dari $\mathbf{I}_m$. Pandang limit berikut

$$\lim_{t \to 0} \frac{f(\mathbf{c} + t\mathbf{u}_j) - f(\mathbf{c})}{t}.$$

Ketika limit tersebut ada, ini disebut turunan parsial *first order* ke- j dari f di $\mathbf{c}$ dan dinotasikan dengan $D_j f(\mathbf{c})$ atau $D_j f$. Notasi $\mathbf{D}f$ menyatakan vektor baris $(D_1 f, D_2 f, \ldots, D_m f)$ di mana elemennya adalah turunan parsial *first order* dari f, sedangkan notasi $\mathbf{D}f(\mathbf{c})$ berarti vektor baris $(D_1 f(\mathbf{c}), D_2 f(\mathbf{c}), \ldots, D_m f(\mathbf{c}))$ di mana elemennya adalah turunan parsial *first order* dari f di $\mathbf{c}$. Notasi lainnya adalah $\partial f(\mathbf{x}) / \partial x_j$ yang merupakan turunan parsial ke- j dari f di $\mathbf{x}$. Selain itu, $\partial f(\mathbf{x}) / \partial \mathbf{x}^T$ menyatakan vektor baris $(\partial f(\mathbf{x}) / \partial x_1, \partial f(\mathbf{x}) / \partial x_2, \ldots, \partial f(\mathbf{x}) / \partial x_m)$ di mana elemennya adalah turunan parsial *first order* dari f di $\mathbf{x}$, sedangkan $\partial f(\mathbf{x}) / \partial \mathbf{x}$ menyatakan vektor kolom $(\partial f(\mathbf{x}) / \partial x_1, \partial f(\mathbf{x}) / \partial x_2, \ldots, \partial f(\mathbf{x}) / \partial x_m)^T$ di mana elemennya juga merupakan turunan parsial *first order* dari f di $\mathbf{x}$.

Fungsi f (dengan domain S di $\Re^{m \times 1}$) dikatakan *continuously differentiable* di *interior point* $\mathbf{c}$ (di S) jika $D_1 f(\mathbf{x}), D_2 f(\mathbf{x}), \ldots, D_m f(\mathbf{x})$ ada dan kontinu di setiap titik $\mathbf{x}$ dalam *neighborhood* dari $\mathbf{c}$. Oleh karena itu, jika f *continuously differentiable* di $\mathbf{c}$ maka

$$\lim_{\mathbf{x}\to\mathbf{c}} \frac{f(\mathbf{x})-\left(f(\mathbf{c})+\mathbf{D}f(\mathbf{c})(\mathbf{x}-\mathbf{c})\right)}{\|\mathbf{x}-\mathbf{c}\|}=\mathbf{0}.$$

Limit di atas mengindikasikan bahwa untuk $\mathbf{x}$ yang cukup dekat dengan $\mathbf{c}$, maka formula Taylor order pertama dapat dinyatakan sebagai berikut:

$$f(\mathbf{x})=f(\mathbf{c})+Df(\mathbf{c})(\mathbf{x}-\mathbf{c})+r(\mathbf{x}-\mathbf{c}),$$

di mana

$$\lim_{\mathbf{x}\to\mathbf{c}} \frac{r(\mathbf{x}-\mathbf{c})}{\|\mathbf{x}-\mathbf{c}\|}=0.$$

Dengan demikian, jika sebuah fungsi f, dengan domain S di $\Re^{m\times 1}$ *continuously differentiable* di *interior point* $\mathbf{c}$ dari S maka f kontinu di $\mathbf{c}$.

Anggap bahwa $D_1 f(\mathbf{x}), D_2 f(\mathbf{x}),\ldots, D_m f(\mathbf{x})$ ada untuk setiap $\mathbf{x}$ dalam *neighborhood* dari $\mathbf{c}$. Ketika turunan parsial *first order* ke-i dari $D_j f$ di $\mathbf{c}$ ada, ini disebut turunan parsial *second order* ke-ij dari f di $\mathbf{c}$ dan dinotasikan dengan $D_{ij}^2 f(\mathbf{c})$ atau dalam bentuk matiks, yaitu $\mathbf{H}f(\mathbf{c})$. Selain itu, notasi tersebut dapat juga ditulis sebagai $\partial^2 f(\mathbf{x})/\partial x_i \partial x_j$ atau pada kasus khusus $i = j, \partial^2 f(\mathbf{x}) / \partial x_i^2$. Notasi $\partial^2 f(\mathbf{x}) / \partial x_i \partial x_j$ dan $\partial^2 f(\mathbf{x}) / \partial x_i^2$ biasanya disingkat menjadi $\partial^2 f / \partial x_i \partial x_j$ dan $\partial^2 f / \partial x_i^2$.

Fungsi f dikatakan *continuously differentiable* dua kali di $\mathbf{c}$ jika f dan semua turunan parsial *first order continuously differentiable* di $\mathbf{c}$. Jika f *continuously differentiable* dua kali di $\mathbf{c}$ maka

$$\lim_{\mathbf{x} \to \mathbf{c}} \frac{f(\mathbf{x}) - \left(f(\mathbf{c}) + \mathbf{D}f(\mathbf{c})(\mathbf{x}-\mathbf{c}) + \frac{1}{2}(\mathbf{x}-\mathbf{c})^T \mathbf{H}f(\mathbf{c})(\mathbf{x}-\mathbf{c}) \right)}{\|\mathbf{x}-\mathbf{c}\|^2} = 0.$$

Limit di atas mengindikasikan bahwa untuk $\mathbf{x}$ yang cukup dekat dengan $\mathbf{c}$, maka formula Taylor order kedua dapat dinyatakan sebagai berikut:

$$f(\mathbf{x}) = f(\mathbf{c}) + Df(\mathbf{c})(\mathbf{x}-\mathbf{c}) + \frac{1}{2}(\mathbf{x}-\mathbf{c})' \mathbf{H}f(\mathbf{c})(\mathbf{x}-\mathbf{c}) + r(\mathbf{x}-\mathbf{c}),$$

di mana

$$\lim_{\mathbf{x} \to \mathbf{c}} \frac{r(\mathbf{x}-\mathbf{c})}{\|\mathbf{x}-\mathbf{c}\|^2} = 0.$$

Misalkan ada vektor kolom $\mathbf{f} = (f_1, f_2, \ldots, f_p)^T$ dengan p fungsi, kemudian diturunkan. Domain dari semua fungsi ini adalah himpunan S di $\Re^{m \times 1}$. Misalkan S memiliki beberapa *interior point* dan $\mathbf{c}$ menyatakan salah satu titik tersebut. Simbol $D_j\mathbf{f}$ digunakan untuk menyatakan vektor kolom berdimensi-*p* di mana elemen ke-*s* adalah turunan parsial ke-*j* dari f_s dan dinotasikan dengan $D_j f_s$. Notasi $\mathbf{Df}$ digunakan untuk menyatakan matriks $p \times m$ di mana elemen ke-*sj* adalah $D_j f_s$, atau ekivalen dengan matriks $p \times m$ di mana kolom ke-*j* adalah $D_j \mathbf{f}$. Bentuk-bentuknya adalah sebagai berikut:

1. $D_j \mathbf{f}(\mathbf{c}) = \left[D_j f_1(\mathbf{c}), D_j f_2(\mathbf{c}), \ldots, D_j f_p(\mathbf{c}) \right]^T$ dan
2. $\mathbf{Df}(\mathbf{c}) = \left[D_1 \mathbf{f}(\mathbf{c}), D_2 \mathbf{f}(\mathbf{c}), \ldots, D_m \mathbf{f}(\mathbf{c}) \right]$ di mana setiap m turunan parsial dari setiap p fungsi $f_1, f_2, \ldots, f_p$ di $\mathbf{c}$ ada.

Selain itu, bentuk di atas dapat pula ditulis sebagai berikut. Misalkan $\mathbf{x} = (x_1, x_2, \ldots, x_m)^T$ merupakan vektor dengan m variabel. Kemudian tuliskan $\partial \mathbf{f}(\mathbf{x}) / \partial \mathbf{x}^T$ (atau $\partial \mathbf{f} / \partial \mathbf{x}^T$) untuk matriks $p \times m$ di mana elemen ke-*sj* adalah $\partial f_s(\mathbf{x}) / \partial x_j$. Pada konteks ini, $\partial \mathbf{f}(\mathbf{x}) / \partial \mathbf{x}^T$ disebut turunan dari $\mathbf{f}(\mathbf{x})$ terhadap $\mathbf{x}^T$ di mana notasi $\partial \mathbf{f}(\mathbf{x}) / \partial \mathbf{x}^T$ memiliki interpretasi yang sama dengan $\mathbf{Df}(\mathbf{x})$ (atau biasa disingkat menjadi $\mathbf{Df}$).

Anggap bahwa ada matriks $\mathbf{F} = \{ f_{st} \}$ berukuran $p \times q$ dengan pq fungsi, diturunkan. Domain dari fungsi-fungsi tersebut adalah himpunan S di $\Re^{m \times 1}$. Matriks $\mathbf{F}$ disebut *continuously differentiable* di *interior point* $\mathbf{c}$ (pada S) jika semua *entri*-nya *continuously differentiable* di $\mathbf{c}$, dan *continuously differentiable* dua kali di $\mathbf{c}$ jika semua *entri*-nya *continuously differentiable* dua kali di $\mathbf{c}$.

2.3.2 Diferensiasi Bentuk Linier dan Bentuk Kuadratik

Berikut ini diberikan diferensiasi dari bentuk linier $\mathbf{a}^T\mathbf{x}$ dan bentuk kuadratik $\mathbf{x}^T\mathbf{A}\mathbf{x}$ yang buktinya dapat dilihat pada lampiran 2.

$$\frac{\partial\left(\mathbf{a}^T\mathbf{x}\right)}{\partial\mathbf{x}} = \mathbf{a} \qquad \textbf{(2.3.2.a)}$$

$$\frac{\partial\left(\mathbf{a}^T\mathbf{x}\right)}{\partial\mathbf{x}^T} = \mathbf{a}^T \qquad \textbf{(2.3.2.b)}$$

$$\frac{\partial\left(\mathbf{x}^T\mathbf{A}\mathbf{x}\right)}{\partial\mathbf{x}} = \left(\mathbf{A}+\mathbf{A}^T\right)\mathbf{x} \qquad \textbf{(2.3.2.c)}$$

2.3.3 Diferensiasi dari Penjumlahan dan Perkalian Matriks

Misalkan $\mathbf{F} = \{f_{is}\}$ menyatakan matriks berukuran $p \times q$ dari fungsi-fungsi yang terdefinisi di himpunan S, oleh vektor $\mathbf{x} = (x_1, x_2, \ldots, x_m)^T$ dengan m variabel. Jika $\mathbf{F} = \mathbf{K}$ di mana $\mathbf{K}$ adalah suatu matriks konstanta berukuran $p \times q$, maka di sembarang *interior point* dari S, $\partial\mathbf{F}/\partial x_j = \mathbf{0}$.

Misalkan $\mathbf{F} = \{f_{is}\}$ dan $\mathbf{G} = \{g_{is}\}$ menyatakan matriks-matriks berukuran $p \times q$ dari fungsi-fungsi yang terdefinisi di himpunan S, oleh vektor $\mathbf{x} = (x_1, x_2, \ldots, x_m)^T$ dengan m variabel. Misalkan pula *a* dan *b* suatu fungsi konstanta yang kontinu di sembarang *interior point* dari S. Jika $\mathbf{F}$ dan $\mathbf{G}$ *continuously differentiable* di sembarang *interior point* dari S maka $a\mathbf{F} + b\mathbf{G}$ juga *continuously differentiable* di sembarang *interior point* tersebut dengan

$$\frac{\partial(a\mathbf{F} + b\mathbf{G})}{\partial x_j} = a\frac{\partial \mathbf{F}}{\partial x_j} + b\frac{\partial \mathbf{G}}{\partial x_j} \quad \textbf{(2.3.3.a)}$$

(bukti dapat dilihat pada lampiran 3)

sehingga dapat disimpulkan sebagai berikut:

$$\frac{\partial a\mathbf{F}}{\partial x_j} = a\frac{\partial \mathbf{F}}{\partial x_j}, \frac{\partial(\mathbf{F} + \mathbf{G})}{\partial x_j} = \frac{\partial \mathbf{F}}{\partial x_j} + \frac{\partial \mathbf{G}}{\partial x_j}, \frac{\partial(\mathbf{F} - \mathbf{G})}{\partial x_j} = \frac{\partial \mathbf{F}}{\partial x_j} - \frac{\partial \mathbf{G}}{\partial x_j}.$$

Misalkan $\mathbf{F} = \{f_{is}\}$ dan $\mathbf{G} = \{g_{is}\}$ menyatakan matriks-matriks berukuran $p \times q$ dan $q \times r$ dari fungsi-fungsi yang terdefinisi di himpunan S, oleh vektor $\mathbf{x} = (x_1, x_2, \ldots, x_m)^T$ dengan m variabel. Jika $\mathbf{F}$ dan $\mathbf{G}$ *continuously differentiable* di sembarang *interior point* dari S maka $\mathbf{FG}$ juga *continuously differentiable* di sembarang *interior point* tersebut dengan

$$\frac{\partial \mathbf{FG}}{\partial x_j} = \mathbf{F}\frac{\partial \mathbf{G}}{\partial x_j} + \frac{\partial \mathbf{F}}{\partial x_j}\mathbf{G}. \quad \textbf{(2.3.3.b)}$$

(bukti dapat dilihat pada lampiran 4).

Bentuk di atas dapat diperluas untuk $\mathbf{F}$, $\mathbf{G}$, dan $\mathbf{H}$, yaitu matriks-matriks berukuran $p \times q$, $q \times r$, dan $r \times v$ dari fungsi-fungsi yang terdefinisi di himpunan S , oleh vektor $\mathbf{x} = (x_1, x_2, \ldots, x_m)^T$ dengan m variabel. Jika $\mathbf{F}$, $\mathbf{G}$, dan $\mathbf{H}$ *continuously differentiable* di sembarang *interior point* dari S maka $\mathbf{FGH}$ juga *continuously differentiable* di sembarang *interior point* tersebut dengan

$$\frac{\partial \mathbf{FGH}}{\partial x_j} = \mathbf{FG}\frac{\partial \mathbf{H}}{\partial x_j} + \mathbf{F}\frac{\partial \mathbf{G}}{\partial x_j}\mathbf{H} + \frac{\partial \mathbf{F}}{\partial x_j}\mathbf{GH} \quad . \qquad \textbf{(2.3.3.c)}$$

2.3.4 Diferensiasi dari *Trace* Matriks

Misalkan $\mathbf{F} = \{f_{is}\}$ menyatakan matriks berukuran $p \times p$ dari fungsi-fungsi yang terdefinisi di himpunan S , oleh vektor $\mathbf{x} = (x_1, x_2, \ldots, x_m)^T$ dengan m variabel. Jika $\mathbf{F}$ *continuously differentiable* di sembarang *interior point* dari S maka $tr(\mathbf{F})$ juga *continuously differentiable* di sembarang *interior point* tersebut dengan

$$\frac{\partial tr(\mathbf{F})}{\partial x_j} = \frac{\partial f_{11}}{\partial x_j} + \frac{\partial f_{22}}{\partial x_j} + \ldots + \frac{\partial f_{pp}}{\partial x_j} = tr\left(\frac{\partial \mathbf{F}}{\partial x_j}\right)$$

di mana $tr(\mathbf{F}) = f_{11} + f_{22} + \ldots + f_{pp}$.

Misalkan $\mathbf{F}=\{f_{is}\}$ dan $\mathbf{G}=\{g_{is}\}$ menyatakan matriks-matriks berukuran $p\times q$ dan $q\times p$ dari fungsi-fungsi yang terdefinisi di himpunan S, oleh vektor $\mathbf{x}=(x_1,x_2,\ldots,x_m)^T$ dengan m variabel. Jika $\mathbf{F}$ dan $\mathbf{G}$ *continuously differentiable* di sembarang *interior point* dari S maka $tr(\mathbf{FG})$ atau $tr(\mathbf{GF})$ juga *continuously differentiable* di sembarang *interior point* tersebut dengan

$$\frac{\partial tr(\mathbf{FG})}{\partial x_j}=\frac{\partial tr(\mathbf{GF})}{\partial x_j}=tr\left(\mathbf{F}\frac{\partial \mathbf{G}}{\partial x_j}\right)+tr\left(\mathbf{G}\frac{\partial \mathbf{F}}{\partial x_j}\right). \qquad \textbf{(2.3.4.a)}$$

2.3.5 Aturan Rantai

Aturan rantai dapat membantu dalam melakukan penurunan parsial dari fungsi f yang merupakan komposit dari fungsi g dengan vektor fungsi $\mathbf{h}$, di mana f memiliki nilai di sembarang titik $\mathbf{x}$ yang diberikan oleh formula

$$f(\mathbf{x})=g\left[\mathbf{h}(\mathbf{x})\right].$$

Misalkan $\mathbf{h}=\{h_i\}$ menyatakan vektor berukuran $n\times 1$ dari fungsi-fungsi yang terdefinisi di himpunan S, oleh vektor $\mathbf{x}=(x_1,x_2,\ldots,x_m)^T$ dengan m variabel. Misalkan pula g menyatakan sebuah fungsi yang terdefinisi di T oleh vektor $\mathbf{y}=(y_1,y_2,\ldots,y_n)^T$ dengan n variabel. Anggap bahwa $\mathbf{h}(\mathbf{x})\in T$ untuk setiap $\mathbf{x}$ di S. Kemudian, ambil f sebagai fungsi komposit yang terdefinisi di himpunan S di mana $f(\mathbf{x}) = g[\mathbf{h}(\mathbf{x})]$. Jika $\mathbf{h}$ *continuously differentiable* di sembarang *interior point* $\mathbf{c}$ dari S dan jika g *continuously differentiable* di $\mathbf{h}(\mathbf{c})$ (asumsikan bahwa $\mathbf{h}(\mathbf{c})$ adalah *interior point* dari T) maka f *continuously differentiable* di $\mathbf{c}$ dengan

$$D_j f(\mathbf{c}) = \sum_{i=1}^{n} D_i g[\mathbf{h}(\mathbf{c})] D_j h_i(\mathbf{c}) = \mathbf{D}g[\mathbf{h}(\mathbf{c})] D_j \mathbf{h}(\mathbf{c}). \qquad \textbf{(2.3.5.a)}$$

Bentuk di atas dapat pula ditulis sebagai berikut:

$$\frac{\partial f}{\partial x_j} = \sum_{i=1}^{n} \frac{\partial g}{\partial y_i} \frac{\partial h_i}{\partial x_j} = \frac{\partial g}{\partial \mathbf{y}^T} \frac{\partial \mathbf{h}}{\partial x_j}, \qquad \textbf{(2.3.5.b)}$$

di mana $\partial g / \partial y_i$ dan $\partial g / \partial \mathbf{y}^T$ diinterpretasikan memiliki nilai di $\mathbf{y} = \mathbf{h}(\mathbf{x})$.

Persamaan (2.3.5.a) dan (2.3.5.b) dapat dinyatakan sebagai berikut:

$$\mathbf{D}f(\mathbf{c}) = \sum_{i=1}^{n} D_i g[\mathbf{h}(\mathbf{c})] \mathbf{D}h_i(\mathbf{c}) = \mathbf{D}g[\mathbf{h}(\mathbf{c})] \mathbf{D}\mathbf{h}(\mathbf{c})$$

atau

$$\frac{\partial f}{\partial \mathbf{x}^T} = \sum_{i=1}^{n} \frac{\partial g}{\partial y_i} \frac{\partial h_i}{\partial \mathbf{x}^T} = \frac{\partial g}{\partial \mathbf{y}^T} \frac{\partial \mathbf{h}}{\partial \mathbf{x}^T}.$$

Bentuk-bentuk di atas dapat diperluas dengan mengambil $\mathbf{g} = \{g_s\}$ sebagai vektor berukuran $p \times 1$ dari fungsi-fungsi yang terdefinisi di $\mathbf{T}$ oleh $\mathbf{y}$, sedangkan $\mathbf{f} = \{f_s\}$ sebagai vektor berukuran $p \times 1$ dari fungsi komposit yang terdefinisi di himpunan S dengan $f_s(\mathbf{x}) = g_s\left[\mathbf{h}(\mathbf{x})\right]$ $(s = 1, 2, \ldots, p)$ atau $\mathbf{f}(\mathbf{x}) = \mathbf{g}\left[\mathbf{h}(\mathbf{x})\right]$. Jika $\mathbf{h}$ *continuously differentiable* di sembarang *interior point* $\mathbf{c}$ dari S dan $\mathbf{g}$ *continuously differentiable* di $\mathbf{h}(\mathbf{c})$, maka $\mathbf{f}$ *continuously differentiable* di $\mathbf{c}$ dengan

$$D_j\mathbf{f}(\mathbf{c}) = \sum_{i=1}^{n} D_i\mathbf{g}\left[\mathbf{h}(\mathbf{c})\right] D_j h_i(\mathbf{c}) = \mathbf{Dg}\left[\mathbf{h}(\mathbf{c})\right] D_j\mathbf{h}(\mathbf{c}) \qquad \textbf{(2.3.5.c)}$$

atau

$$\frac{\partial \mathbf{f}}{\partial x_j} = \sum_{i=1}^{n} \frac{\partial \mathbf{g}}{\partial y_i}\frac{\partial h_i}{\partial x_j} = \frac{\partial \mathbf{g}}{\partial \mathbf{y}^T}\frac{\partial \mathbf{h}}{\partial x_j} \qquad \textbf{(2.3.5.d)}$$

di mana $\partial\mathbf{g}/\partial y_i$ dan $\partial\mathbf{g}/\partial\mathbf{y}^T$ diinterpretasikan memiliki nilai di $\mathbf{y} = \mathbf{h}(\mathbf{x})$.

Persamaan (2.3.5.c) dan (2.3.5.d) dapat dinyatakan sebagai berikut:

$$\mathbf{Df}(\mathbf{c}) = \sum_{i=1}^{n} D_i\mathbf{g}\left[\mathbf{h}(\mathbf{c})\right]\mathbf{D}h_i(\mathbf{c}) = \mathbf{Dg}\left[\mathbf{h}(\mathbf{c})\right]\mathbf{Dh}(\mathbf{c})$$

atau

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}^T} = \sum_{i=1}^{n} \frac{\partial \mathbf{g}}{\partial y_i}\frac{\partial h_i}{\partial \mathbf{x}^T} = \frac{\partial \mathbf{g}}{\partial \mathbf{y}^T}\frac{\partial \mathbf{h}}{\partial \mathbf{x}^T}.$$

Berikut ini akan disajikan generalisasi dari aturan rantai pada matriks. Misalkan $\mathbf{H} = \{h_{is}\}$ adalah matriks berukuran $n \times r$ dari fungsi-fungsi yang terdefinisi di S oleh $\mathbf{x}$ di mana g adalah sebuah fungsi yang terdefinisi di himpunan T oleh sebuah matriks $\mathbf{Y} = \{y_{is}\}$ berukuran $n \times r$ dengan nr variabel. Misalkan pula $\mathbf{H}(\mathbf{x}) \in T$ untuk setiap $\mathbf{x}$ di S dan f merupakan fungsi komposit terdefinisi di S dengan $f(\mathbf{x}) = g[\mathbf{H}(\mathbf{x})]$.

Anggap bahwa elemen-elemen dari $\mathbf{H}$ dan $\mathbf{Y}$ masing-masing disusun ke dalam bentuk vektor kolom $\mathbf{h}$ dan $\mathbf{y}$. Kemudian, untuk tujuan penurunan, g diinterpretasikan kembali sebagai fungsi dari $\mathbf{y}$. Jika $\mathbf{h}$ atau $\mathbf{H}$ *continuously differentiable* di sembarang *interior point* $\mathbf{c}$ dari S dan jika g *continuously differentiable* di $\mathbf{h}(\mathbf{c})$ atau $\mathbf{H}(\mathbf{c})$ (asumsikan bahwa $\mathbf{H}(\mathbf{c})$ adalah *interior point* dari T) maka f *continuously differentiable* di $\mathbf{x} = \mathbf{c}$, yakni:

$$\frac{\partial f}{\partial x_j} = \sum_{i=1}^{n} \sum_{s=1}^{r} \frac{\partial g}{\partial y_{is}} \frac{\partial h_{is}}{\partial x_j}$$

atau dapat ditulis kembali sebagai

$$\frac{\partial f}{\partial x_j} = tr\left[\left(\frac{\partial g}{\partial \mathbf{Y}}\right)^T \frac{\partial \mathbf{H}}{\partial x_j}\right] \qquad \textbf{(2.3.5.e)}$$

di mana $\partial g / \partial y_{is}$ dan $\partial g / \partial \mathbf{Y}$ diinterpretasikan memiliki nilai di $\mathbf{Y} = \mathbf{H}(\mathbf{x})$.

2.3.6 Diferensiasi Parsial Pertama dari Determinan, Invers, dan *Adjoint* Matriks

Misalkan $\mathbf{X} = \{x_{ij}\}$ menyatakan sebuah matriks berukuran $m \times m$ dengan m^2 variabel "independen" (di mana $m \geq 2$), kemudian anggap bahwa nilai dari $\mathbf{X}$ terdiri dari bilangan-bilangan di $\Re^{m\times m}$. Misalkan ξ_{ij} menyatakan kofaktor dari x_{ij}. Maka $f(\mathbf{X}) = \det(\mathbf{X})$ di $\Re^{m\times m}$ *continuously differentiable* di setiap $\mathbf{X}$ dengan

$$\frac{\partial \det(\mathbf{X})}{\partial x_{ij}} = \xi_{ij} . \qquad \textbf{(2.3.6.a)}$$

Untuk melihat hal tersebut, susun kembali elemen-elemen dari $\mathbf{X}$ ke dalam bentuk vektor kolom $\mathbf{x}$ berdimensi m^2, dan misalkan f sebagai fungsi dari $\mathbf{x}$. Oleh karena penjumlahan serta perkalian dari fungsi yang *continuously differentiable* adalah *continuously differentiable*, berdasarkan definisi determinan maka f *continuously differentiable*. Bentuk $\det(\mathbf{X})$ dapat dijabarkan sebagai $\det(\mathbf{X}) = \sum_{t=1}^{m} x_{it}\xi_{it}$ di mana $\xi_{i1}, \xi_{i2}, \ldots, \xi_{im}$ berupa konstanta sehingga didapat

$$\frac{\partial \det(\mathbf{X})}{\partial x_{ij}} = \sum_{t=1}^{m} \xi_{it} \frac{\partial x_{it}}{\partial x_{ij}} = \xi_{ij} .$$

Hasil (2.3.6.a) mengindikasikan bahwa turunan dari $\det(\mathbf{X})$ terhadap $\mathbf{X}$ adalah matriks kofaktor dari $\mathbf{X}$ atau ekivalen dengan

$$\frac{\partial \det(\mathbf{X})}{\partial \mathbf{X}} = \left[adj(\mathbf{X})\right]^T .$$

Selanjutnya, akan diturunkan determinan dari matriks $\mathbf{F} = \{f_{is}\}$ berukuran $p \times p$ dari fungsi-fungsi yang terdefinisi di himpunan S, oleh vektor $\mathbf{x} = (x_1, x_2, \ldots, x_m)^T$ dengan m variabel. Maka, fungsi yang diturunkan adalah fungsi h dari $\mathbf{x}$ yang terdefinisi di himpunan S dengan $h(\mathbf{x}) = \det\left[\mathbf{F}(\mathbf{x})\right]$. Untuk menurunkan h, diperkenalkan fungsi g dari matriks $\mathbf{Y} = \{y_{is}\}$ berukuran $p \times p$ dengan p^2 variabel, terdefinisi di $\Re^{p \times p}$, dengan $g(\mathbf{Y}) = \det(\mathbf{Y})$. Untuk menyatakan h sebagai komposit dari g dan $\mathbf{F}$, yaitu $h(\mathbf{x}) = g\left[\mathbf{F}(\mathbf{x})\right]$ telah jelas bahwa g *continuously differentiable* di setiap $\mathbf{Y}$ dengan $\frac{\partial g}{\partial \mathbf{Y}} = \left[adj(\mathbf{Y})\right]^T$.

Berdasarkan aturan rantai (2.3.5.e), jika $\mathbf{F}$ *continuously differentiable* di *interior point* $\mathbf{c}$ dari S, maka h *continuously differentiable* di $\mathbf{x} = \mathbf{c}$, yaitu:

$$\frac{\partial \det(\mathbf{F})}{\partial x_j} = tr\left[adj(\mathbf{F})\frac{\partial \mathbf{F}}{\partial x_j}\right]. \qquad \textbf{(2.3.6.b)}$$

Oleh karena itu, jika $\mathbf{F}$ *nonsingular* dan *continuously differentiable* di $\mathbf{c}$, maka berdasarkan (2.1.4.a) didapat

$$\frac{\partial \det(\mathbf{F})}{\partial x_j} = |\mathbf{F}|\, tr\left(\mathbf{F}^{-1}\frac{\partial \mathbf{F}}{\partial x_j}\right). \qquad \textbf{(2.3.6.c)}$$

Sekarang anggap bahwa S adalah himpunan semua nilai $\mathbf{x}$ di mana $\det\left[\mathbf{F}(\mathbf{x})\right] > 0$. Misalkan l adalah sebuah fungsi dari $\mathbf{x}$ (di S) yang didefinisikan sebagai $l(\mathbf{x}) = \log\det\left[\mathbf{F}(\mathbf{x})\right]$. Untuk tujuan penurunan, misalkan g menyatakan fungsi dari sebuah variabel y yang didefinisikan sebagai $g(y) = \log y$ (untuk $y > 0$). Kemudian, nyatakan l sebagai komposit dari g dan h sedemikian sehingga $l(\mathbf{x}) = g\left[h(\mathbf{x})\right]$. Diketahui dari kalkulus satu variabel bahwa untuk $y > 0$, g *continuously differentiable* di setiap titik di *domain*-nya dengan $\frac{\partial \log y}{\partial y} = \frac{1}{y}$.

Menggunakan aturan rantai dan berdasarkan (2.3.6.b) serta (2.3.6.c) didapat bahwa jika $\mathbf{F}$ *continuously differentiable* di *interior point* $\mathbf{c}$ dari S (di mana h *continuously differentiable* di $\mathbf{c}$), maka l *continuously differentiable* di $\mathbf{x} = \mathbf{c}$ dengan

$$\frac{\partial \log\det(\mathbf{F})}{\partial x_j} = \frac{1}{|\mathbf{F}|}\frac{\partial \det(\mathbf{F})}{\partial x_j} = \frac{1}{|\mathbf{F}|} tr\left(adj(\mathbf{F})\frac{\partial \mathbf{F}}{\partial x_j}\right) = tr\left(\mathbf{F}^{-1}\frac{\partial \mathbf{F}}{\partial x_j}\right). \qquad \textbf{(2.3.6.d)}$$

Selanjutnya, diketahui bahwa $\mathbf{F}\mathbf{F}^{-1} = \mathbf{I}$. Maka berdasarkan (2.3.3.b), di $\mathbf{x} = \mathbf{c}$ didapat

$$\mathbf{F}\frac{\partial \mathbf{F}^{-1}}{\partial x_j} + \frac{\partial \mathbf{F}}{\partial x_j}\mathbf{F}^{-1} = \frac{\partial \mathbf{I}}{\partial x_j} = \mathbf{0}$$

dengan demikian

$$\mathbf{F}\frac{\partial \mathbf{F}^{-1}}{\partial x_j} = -\frac{\partial \mathbf{F}}{\partial x_j}\mathbf{F}^{-1}.$$

Dengan mengalikan kedua ruas dengan $\mathbf{F}^{-1}$ didapat

$$\frac{\partial \mathbf{F}^{-1}}{\partial x_j} = -\mathbf{F}^{-1}\frac{\partial \mathbf{F}}{\partial x_j}\mathbf{F}^{-1}. \qquad \textbf{(2.3.6.e)}$$

Oleh karena $adj(\mathbf{F}) = |\mathbf{F}|\mathbf{F}^{-1}$, berdasarkan (2.3.6.c) didapat pula

$$\frac{\partial adj(\mathbf{F})}{\partial x_j} = \frac{\partial |\mathbf{F}|}{\partial x_j}\mathbf{F}^{-1} + |\mathbf{F}|\frac{\partial \mathbf{F}^{-1}}{\partial x_j}$$
$$= |\mathbf{F}|\, tr\left(\mathbf{F}^{-1}\frac{\partial \mathbf{F}}{\partial x_j}\right)\mathbf{F}^{-1} + |\mathbf{F}|\left(-\mathbf{F}^{-1}\frac{\partial \mathbf{F}}{\partial x_j}\mathbf{F}^{-1}\right) = |\mathbf{F}|\left[tr\left(\mathbf{F}^{-1}\frac{\partial \mathbf{F}}{\partial x_j}\right)\mathbf{F}^{-1} - \mathbf{F}^{-1}\frac{\partial \mathbf{F}}{\partial x_j}\mathbf{F}^{-1}\right].$$

2.3.7 Diferensiasi Parsial Kedua dari Determinan dan Invers Matriks

Misalkan $\mathbf{F} = \{f_{is}\}$ menyatakan matriks berukuran $p \times p$ dari fungsi-fungsi yang terdefinisi di himpunan S , oleh vektor $\mathbf{x} = (x_1, x_2, \ldots, x_m)^T$ dengan

m variabel. Anggap bahwa $\mathbf{F}(\mathbf{x})$ *nonsingular* untuk setiap $\mathbf{x}$ di S, dan notasikan $\mathbf{c}$ sembarang *interior point* dari S di mana $\mathbf{F}$ *continuously differentiable* dua kali sehingga $\mathbf{F}$ *continuously differentiable* di $\mathbf{c}$. Pada kenyataannya $\mathbf{F}$ *continuously differentiable* di setiap titik dalam *neighborhood* *N* dari $\mathbf{c}$.

Dari SubBab 2.3.6 diketahui bahwa $\det(\mathbf{F})$ dan $\mathbf{F}^{-1}$ *continuously differentiable* di setiap titik di *N* di mana untuk $\mathbf{x} \in N$

$$\frac{\partial \det(\mathbf{F})}{\partial x_j} = |\mathbf{F}|\, tr\left(\mathbf{F}^{-1}\frac{\partial \mathbf{F}}{\partial x_j}\right) \text{ dan } \frac{\partial \mathbf{F}^{-1}}{\partial x_j} = -\mathbf{F}^{-1}\frac{\partial \mathbf{F}}{\partial x_j}\mathbf{F}^{-1}$$

di mana $\partial \mathbf{F}/\partial x_j$ *continuously differentiable* di $\mathbf{x} = \mathbf{c}$ sehingga $\partial \det(\mathbf{F})/\partial x_j$ dan $\partial \mathbf{F}^{-1}/\partial x_j$ *continuously differentiable* di $\mathbf{c}$.

Dengan demikian, $\det(\mathbf{F})$ dan $\mathbf{F}^{-1}$ *continuously differentiable* dua kali di $\mathbf{x} = \mathbf{c}$, yaitu sebagai berikut:

$$\frac{\partial^2 \det(\mathbf{F})}{\partial x_i \partial x_j} = |\mathbf{F}| \begin{bmatrix} tr\left(\mathbf{F}^{-1}\dfrac{\partial \mathbf{F}}{\partial x_i \partial x_j}\right) + |\mathbf{F}|\, tr\left(\mathbf{F}^{-1}\dfrac{\partial \mathbf{F}}{\partial x_i}\right) tr\left(\mathbf{F}^{-1}\dfrac{\partial \mathbf{F}}{\partial x_j}\right) \\ -tr\left(\mathbf{F}^{-1}\dfrac{\partial \mathbf{F}}{\partial x_i}\mathbf{F}^{-1}\dfrac{\partial \mathbf{F}}{\partial x_j}\right) \end{bmatrix} \quad \textbf{(2.3.7.a)}$$

$$\frac{\partial^2 \mathbf{F}^{-1}}{\partial x_i \partial x_j} = -\mathbf{F}^{-1} \frac{\partial \mathbf{F}}{\partial x_i \partial x_j} \mathbf{F}^{-1} + \mathbf{F}^{-1} \frac{\partial \mathbf{F}}{\partial x_i} \mathbf{F}^{-1} \frac{\partial \mathbf{F}}{\partial x_j} \mathbf{F}^{-1} + \mathbf{F}^{-1} \frac{\partial \mathbf{F}}{\partial x_j} \mathbf{F}^{-1} \frac{\partial \mathbf{F}}{\partial x_i} \mathbf{F}^{-1} \qquad \textbf{(2.3.7.b)}$$

(bukti diberikan pada lampiran 5).

Sekarang anggap bahwa $\det\left[\mathbf{F}(\mathbf{x})\right] > 0$ untuk setiap $\mathbf{x}$ di S.

Berdasarkan hasil sebelumnya, $\log\det(\mathbf{F})$ *continuously differentiable* di setiap titik di *N* dan (untuk $\mathbf{x} \in N$)

$$\frac{\partial \log\det(\mathbf{F})}{\partial x_j} = tr\left(\mathbf{F}^{-1} \frac{\partial \mathbf{F}}{\partial x_j}\right)$$

Oleh karena $\mathbf{F}^{-1}$ dan $\partial\mathbf{F}/\partial x_j$ *continuously differentiable* di $\mathbf{c}$, maka $\partial \log\det(\mathbf{F})/\partial x_j$ juga *continuously differentiable* di $\mathbf{c}$. Dengan demikian, $\log\det(\mathbf{F})$ *continuously differentiable* dua kali di $\mathbf{x} = \mathbf{c}$, yakni:

$$\frac{\partial^2 \log\det(\mathbf{F})}{\partial x_i \partial x_j} = tr\left(\mathbf{F}^{-1} \frac{\partial \mathbf{F}}{\partial x_i \partial x_j}\right) - tr\left(\mathbf{F}^{-1} \frac{\partial \mathbf{F}}{\partial x_i} \mathbf{F}^{-1} \frac{\partial \mathbf{F}}{\partial x_j}\right) \qquad \textbf{(2.3.7.c)}$$

(bukti diberikan pada lampiran 5).

2.4 Penaksiran Parameter pada *General Linear Mixed Model* dengan Menggunakan Metode BLUP dan EBLUP

Berikutnya diketahui *General Linear Mixed Model* adalah sebagai berikut:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{Z}\mathbf{b} + \mathbf{e} \qquad \textbf{(2.4.a)}$$

di mana

$\mathbf{y}$: vektor *random* dari variabel respon yang terobservasi berukuran $n \times 1$ di mana nilai observasinya disebut vektor data.

$\mathbf{X}$: matriks *full rank* berukuran $n \times k$ dari variabel prediktor yang elemen elemennya diketahui.

$\boldsymbol{\alpha}$: vektor parameter bersifat *fixed* berukuran $k \times 1$ yang tidak diketahui dan tidak terobservasi.

$\mathbf{Z}$: matriks *full rank* berukuran $n \times h$ dari variabel prediktor yang elemen elemennya diketahui.

$\mathbf{b}$: vektor *random* parameter yang tidak diketahui dan tidak terobservasi berukuran $h \times 1$.

$\mathbf{e}$: vektor *random error* yang tidak terobservasi berukuran $n \times 1$.

dengan asumsi

$E(\mathbf{b}) = \mathbf{0}$, $E(\mathbf{e}) = \mathbf{0}$, $Var(\mathbf{b}) = E\left(\mathbf{b}\mathbf{b}^T\right) = \mathbf{G}$, $Var(\mathbf{e}) = E\left(\mathbf{e}\mathbf{e}^T\right) = \mathbf{R}$, di mana **b** dan **e** *independently distributed* sedangkan **G** dan **R** adalah matriks varians kovarians yang tergantung pada vektor parameter dari variansi efek *random*, yaitu $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_q)^T$. Oleh karena **b** dan **e** independen, maka $corr(\mathbf{b}, \mathbf{e}) = \mathbf{0}$ sehingga $\operatorname{cov}(\mathbf{b}, \mathbf{e}) = E(\mathbf{b}\mathbf{e}^T) = E(\mathbf{e}\mathbf{b}^T) = \mathbf{0}$.

Misalkan nilai-nilai dari vektor *random* $\mathbf{b}$ dinotasikan dengan $\boldsymbol{\beta}$ di mana $\boldsymbol{\beta}$ berbentuk vektor. Ingin diketahui pengaruh dari efek *fixed* dan efek *random* (tetapi bukan *error*) yang merupakan kombinasi linier dari $\boldsymbol{\alpha}$ dan $\boldsymbol{\beta}$. Misalkan $T(\mathbf{y}) = c + \boldsymbol{l}^T\mathbf{y}$ adalah sembarang penaksir dari kombinasi linier $\boldsymbol{\lambda}^T\boldsymbol{\alpha} + \boldsymbol{\omega}^T\boldsymbol{\beta}$ di mana c suatu nilai konstanta dan $\boldsymbol{l}$ suatu vektor konstanta sedangkan $\boldsymbol{\lambda}$ berukuran $k \times 1$ dan $\boldsymbol{\omega}$ berukuran $h \times 1$. $T(\mathbf{y})$ disebut *unbiased* jika $E(T(\mathbf{y})) = E(\boldsymbol{\lambda}^T\boldsymbol{\alpha} + \boldsymbol{\omega}^T\mathbf{b})$ dan disebut linier jika $T(a\mathbf{y}_1 + b\mathbf{y}_2) = aT(\mathbf{y}_1) + bT(\mathbf{y}_2)$. $T(\mathbf{y})$ dapat mengestimasi kombinasi linier tersebut jika $T(\mathbf{y})$ merupakan penaksir yang *unbiased* dan linier.

Diketahui bahwa $\boldsymbol{\lambda}^T\boldsymbol{\alpha} + \boldsymbol{\omega}^T\boldsymbol{\beta}$ dapat diestimasi oleh $T(\mathbf{y})$ jika dan hanya jika $c = 0$ dan $\boldsymbol{\lambda}^T$ merupakan kombinasi linier dari baris-baris $\mathbf{X}$ (bukti diberikan pada lampiran 6). **(2.4.b)**

Selanjutnya, ambil sembarang $\tau = \boldsymbol{\lambda}^T\boldsymbol{\alpha} + \boldsymbol{\omega}^T\boldsymbol{\beta}$ yang merupakan kombinasi linier yang dapat diestimasi kemudian definisikan $\mathrm{t} = \boldsymbol{\lambda}^T\boldsymbol{\alpha} + \boldsymbol{\omega}^T\mathbf{b}$. Selain itu, definisikan pula *Mean Squared Error* (MSE) dari penaksir $T(\mathbf{y})$ adalah sebagai berikut:

$$E\left[\left(T(\mathbf{y}) - \boldsymbol{\lambda}^T\boldsymbol{\alpha} - \boldsymbol{\omega}^T\mathbf{b}\right)^2\right] \qquad \textbf{(2.4.c)}$$

Dengan menggunakan Metode *Best Linear Unbiased Prediction* (BLUP) di mana diasumsikan $\boldsymbol{\delta}$ diketahui, didapat $\hat{\boldsymbol{\alpha}}(\boldsymbol{\delta},\mathbf{y})$ dan $\hat{\boldsymbol{\beta}}(\boldsymbol{\delta},\mathbf{y})$ sedemikian sehingga

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\delta},\mathbf{y}) = \left(\mathbf{X}^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\delta})\mathbf{X}\right)^{-1}\mathbf{X}^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\delta})\mathbf{y} \qquad \textbf{(2.4.d)}$$

dan

$$\hat{\boldsymbol{\beta}}(\boldsymbol{\delta},\mathbf{y}) = \mathbf{G}(\boldsymbol{\delta})\mathbf{Z}^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\delta})\left(\mathbf{y}-\mathbf{X}\hat{\boldsymbol{\alpha}}(\boldsymbol{\delta},\mathbf{y})\right) \qquad \textbf{(2.4.e)}$$

di mana

$$\boldsymbol{\Omega}(\boldsymbol{\delta}) = \boldsymbol{\Omega} = \mathbf{R} + \mathbf{ZGZ}^T.$$

Dengan demikian*, Best Linear Unbiased Predictor* (BLUP) dari τ adalah

$$\hat{\tau}(\boldsymbol{\delta},\mathbf{y}) = \boldsymbol{\lambda}^T\hat{\boldsymbol{\alpha}}(\boldsymbol{\delta},\mathbf{y}) + \boldsymbol{\omega}^T\hat{\boldsymbol{\beta}}(\boldsymbol{\delta},\mathbf{y}) \qquad \textbf{(2.4.f)}$$

yang memiliki MSE terkecil dari semua taksiran yang *unbiased* dan linier.

Namun, pada kenyataannya $\boldsymbol{\delta}$ tidak diketahui nilainya sehingga harus ditaksir. Ada banyak metode untuk menaksir $\boldsymbol{\delta}$, salah satunya adalah metode *Maximum Likelihood*. Oleh karena digunakan metode ML maka dibutuhkan fungsi *likelihood* yang merupakan *pdf* dari $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T$ di mana **b** dan **e** diasumsikan berdistribusi normal $\left(\mathbf{b} \sim N(0,\mathbf{G})\ \ dan\ \mathbf{e} \sim N(0,\mathbf{R})\right)$. Oleh karena **b** dan **e** berdistribusi normal, maka **y** pun berdistribusi normal atau ditulis sebagai $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\alpha},\boldsymbol{\Omega})$ sehingga fungsi *likelihood*-nya adalah sebagai berikut:

$$L(\boldsymbol{\alpha},\boldsymbol{\delta};\mathbf{y}) = f(\mathbf{y}) = \frac{1}{(2\pi)^{n/2}\left|\boldsymbol{\Omega}(\boldsymbol{\delta})\right|^{1/2}}\exp\left(-\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\alpha})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\delta})(\mathbf{y}-\mathbf{X}\boldsymbol{\alpha})\right). \quad \textbf{(2.4.g)}$$

(Bukti diberikan pada lampiran 7).

Kemudian Metode ML menggunakan fungsi *log-likelihood*, menjadi:

$$\ln L(\boldsymbol{\alpha},\boldsymbol{\delta};\mathbf{y}) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln\left(\left|\boldsymbol{\Omega}(\boldsymbol{\delta})\right|\right) - \frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\alpha})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\delta})(\mathbf{y}-\mathbf{X}\boldsymbol{\alpha})$$

$$= c - \frac{1}{2}\left(\ln\left(\left|\boldsymbol{\Omega}(\boldsymbol{\delta})\right|\right) + (\mathbf{y}-\mathbf{X}\boldsymbol{\alpha})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\delta})(\mathbf{y}-\mathbf{X}\boldsymbol{\alpha})\right) \quad \textbf{(2.4.h)}$$

dengan c suatu konstanta. Selanjutnya dengan menggunakan (2.3.6.d) dan (2.3.6.e) fungsi *log-likelihood* tersebut diturunkan terhadap $\boldsymbol{\delta}$, dinotasikan dengan $\mathbf{s}(\boldsymbol{\alpha},\boldsymbol{\delta})$, di mana untuk elemen ke-*j* adalah sebagai berikut:

$$s_j(\boldsymbol{\alpha},\boldsymbol{\delta}) = -\frac{1}{2}tr\left(\boldsymbol{\Omega}^{-1}\boldsymbol{\Omega}_{(j)}\right) + \frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\alpha})^T\left(\boldsymbol{\Omega}^{-1}\boldsymbol{\Omega}_{(j)}\boldsymbol{\Omega}^{-1}\right)(\mathbf{y}-\mathbf{X}\boldsymbol{\alpha}) \quad \textbf{(2.4.i)}$$

dengan $\boldsymbol{\Omega}_{(j)} = \dfrac{\partial\boldsymbol{\Omega}(\boldsymbol{\delta})}{\partial\delta_j}$ (bukti dapat dilihat pada lampiran 8).

Sesuai dengan prosedur ML, berikutnya (2.4.i) dicari solusinya menjadi

$$tr\left(\boldsymbol{\Omega}^{-1}(\boldsymbol{\delta})\boldsymbol{\Omega}_{(j)}\right) = (\mathbf{y}-\mathbf{X}\tilde{\boldsymbol{\alpha}})^T\left(\boldsymbol{\Omega}^{-1}(\boldsymbol{\delta})\boldsymbol{\Omega}_{(j)}\boldsymbol{\Omega}^{-1}(\boldsymbol{\delta})\right)(\mathbf{y}-\mathbf{X}\tilde{\boldsymbol{\alpha}}) \quad \textbf{(2.4.j)}$$

Terlihat dari (2.4.j) $\boldsymbol{\delta}$ tidak dapat dicari solusinya. Oleh karena itu, taksiran $\boldsymbol{\delta}$ dengan Metode ML diselesaikan secara iteratif yaitu dengan *Scoring Algorithm*:

$$\boldsymbol{\delta}^{(a+1)} = \boldsymbol{\delta}^{(a)} + \left(\boldsymbol{\Upsilon}\left(\boldsymbol{\delta}^{(a)}\right)\right)^{-1}\mathbf{s}\left(\tilde{\boldsymbol{\beta}}\left(\boldsymbol{\delta}^{(a)}\right),\boldsymbol{\delta}^{(a)}\right)$$

dengan

$$\mathbf{s}(\boldsymbol{\alpha},\boldsymbol{\delta}) = \left(s_1(\boldsymbol{\alpha},\boldsymbol{\delta}), s_2(\boldsymbol{\alpha},\boldsymbol{\delta}),\ldots, s_q(\boldsymbol{\alpha},\boldsymbol{\delta})\right)^T \text{ dan } \Upsilon(\boldsymbol{\delta}) = \left(\Upsilon_1^T(\boldsymbol{\delta}), \Upsilon_2^T(\boldsymbol{\delta}),\ldots, \Upsilon_q^T(\boldsymbol{\delta})\right)^T$$

di mana

$$\Upsilon(\boldsymbol{\delta}) = -E\left(\frac{\partial^2 \ln L(\boldsymbol{\alpha},\boldsymbol{\delta};\mathbf{y})}{\partial\boldsymbol{\delta}\partial\boldsymbol{\delta}^T}\right) = \frac{1}{2} tr\left(\boldsymbol{\Omega}^{-1}\boldsymbol{\Omega}_{(j)}\boldsymbol{\Omega}^{-1}\boldsymbol{\Omega}_{(k)}\right) \qquad \textbf{(2.4.k)}$$

(dibuktikan pada lampiran 9)

Setelah didapatkan taksiran $\boldsymbol{\delta}$, dinotasikan dengan $\hat{\boldsymbol{\delta}}(\mathbf{y})$, maka didapat taksiran parameter yang baru (*two-stage estimator*) yang didapat dengan mensubstitusikan $\hat{\boldsymbol{\delta}}$ ke $\hat{\boldsymbol{\alpha}}$ dan $\hat{\boldsymbol{\beta}}$ pada persamaan (2.4.d) dan (2.4.e) sehingga menjadi

$$\hat{\boldsymbol{\alpha}}\left(\hat{\boldsymbol{\delta}}(\mathbf{y}),\mathbf{y}\right) = \left(\mathbf{X}^T\boldsymbol{\Omega}^{-1}\left(\hat{\boldsymbol{\delta}}\right)\mathbf{X}\right)^{-1}\mathbf{X}^T\boldsymbol{\Omega}^{-1}\left(\hat{\boldsymbol{\delta}}\right)\mathbf{y} \text{ dan}$$

$$\hat{\boldsymbol{\beta}}\left(\hat{\boldsymbol{\delta}}(\mathbf{y}),\mathbf{y}\right) = \mathbf{G}\left(\hat{\boldsymbol{\delta}}\right)\mathbf{Z}^T\boldsymbol{\Omega}^{-1}\left(\hat{\boldsymbol{\delta}}\right)\left(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\alpha}}\left(\hat{\boldsymbol{\delta}}(\mathbf{y}),\mathbf{y}\right)\right)$$

Berdasarkan (2.4.f) maka taksiran kombinasi linier yang baru atau dikenal dengan taksiran EBLUP adalah sebagai berikut:

$$\hat{\tau}\left(\hat{\boldsymbol{\delta}}(\mathbf{y}),\mathbf{y}\right) = \boldsymbol{\lambda}^T\hat{\boldsymbol{\alpha}}\left(\hat{\boldsymbol{\delta}}(\mathbf{y}),\mathbf{y}\right) + \boldsymbol{\omega}^T\hat{\boldsymbol{\beta}}\left(\hat{\boldsymbol{\delta}}(\mathbf{y}),\mathbf{y}\right). \qquad \textbf{(2.4.l)}$$

2.5 Definisi *Translation-Invariant* dan *maximal-Invariant*

2.5.1 Definisi *Translation-Invariant*

Sebuah fungsi $f(\mathbf{y})$ dikatakan *translation-invariant* jika

$$f(\mathbf{y}-\mathbf{Xa}) = f(\mathbf{y}) \text{ untuk setiap } \mathbf{y} \text{ dan } \mathbf{a}. \qquad \textbf{(2.5.a)}$$

2.5.2 Definisi *maximal-Invariant*

Misalkan $\mathcal{X}$ suatu ruang dan $\mathcal{G}$ adalah grup dari transformasi $\mathcal{X}$. Fungsi $T(x)$ pada $\mathcal{X}$ dikatakan maximal-invariant pada $\mathcal{G}$ jika:

a. (invariance) $T(g(x)) = T(x)$, untuk setiap $x \in \mathcal{X}$ dan $g \in \mathcal{G}$;

b. (maximality) $T(x_1) = T(x_2) \rightarrow x_1 = g(x_2)$, untuk setiap $g \in \mathcal{G}$.

***2.6 Mean Squared Error* (MSE)**

2.6.1 Definisi MSE

Misalkan θ merupakan suatu parameter dan $\hat{\theta}$ merupakan taksiran dari parameter θ, sehingga *Mean Squared Error* (MSE) dari suatu taksiran parameter θ didefinisikan sebagai berikut:

$$\begin{aligned} \text{MSE}\left[\hat{\theta}\right] &= E\left[\left(\hat{\theta}-\theta\right)^2\right] \\ &= \text{var}\left[\hat{\theta}\right] + \left[\text{bias}\left(\hat{\theta}\right)\right]^2 \end{aligned}$$

Bukti:

misalkan $E\left[\hat{\theta}\right]=m$ sehingga:

$$\begin{aligned}
\mathrm{MSE}\left[\hat{\theta}\right] &= E\left[\left(\hat{\theta}-\theta\right)^2\right] \\
&= E\left[\left(\hat{\theta}-m+m-\theta\right)^2\right] \\
&= E\left[\left(\hat{\theta}-m\right)^2+2\left(\hat{\theta}-m\right)(m-\theta)+(m-\theta)^2\right] \\
&= E\left[\left(\hat{\theta}-m\right)^2\right]+2E\left[\left(\hat{\theta}-m\right)(m-\theta)\right]+E\left[(m-\theta)^2\right] \\
&= E\left[\left(\hat{\theta}-m\right)^2\right]+2\underbrace{E\left[\left(\hat{\theta}-m\right)\right]}_{=0}(m-\theta)+(m-\theta)^2 \\
&= \mathrm{var}\left[\hat{\theta}\right]+\left[\mathrm{bias}\left(\hat{\theta}\right)\right]^2
\end{aligned}$$

Berdasarkan definisi MSE di atas, jika $\hat{\theta}$ yang diperoleh *unbiased*, maka MSE dari taksiran parameter θ akan sama dengan variansi dari taksiran parameter θ.

Bukti:

$\hat{\theta}$ dikatakan unbiased jika:

$E\left(\hat{\theta}\right)=\theta$, sehingga

$$\begin{aligned}
\mathrm{MSE}\left[\hat{\theta}\right] &= E\left[\left(\hat{\theta}-\theta\right)^2\right] \\
&= E\left[\left(\hat{\theta}-m\right)^2\right]+(m-\theta)^2, \;\; E\left[\hat{\theta}\right]=m=\theta \\
&= \mathrm{var}\left[\hat{\theta}\right]+0 \\
&= \mathrm{var}\left[\hat{\theta}\right]
\end{aligned}$$

Dengan demikian, terbukti bahwa MSE dari taksiran parameter θ akan sama dengan variansi dari taksiran parameter θ jika $\hat{\theta}$ yang diperoleh *unbiased*.

2.6.2 Tujuan MSE

MSE ini mempunyai beberapa tujuan, diantaranya adalah sebagai berikut:

1. Mengukur seberapa baik taksiran parameter yang diperoleh.
2. Pembentukan *confidence interval*.
3. Pembentukan statistik uji.

2.6.3 MSE pada *Simple Linear Regression* (Regresi Linier Sederhana)

Bentuk dari regresi linier sederhana adalah sebagai berikut:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \qquad \text{i=1,2,}\cdots\text{,n}$$

di mana

y = variabel dependen (variable respon)

x = variabel independen (variable prediktor)

$\beta_0 =$ titik potong dengan sumbu y

$\beta_1 =$ kemiringan garis

$\varepsilon =$ komponen random error

dengan asumsi

$\varepsilon \sim NIID\left(0, \sigma^2\right).$

Dengan menggunakan metode *Ordinary Least Square* (OLS) didapatkan taksiran parameter sebagai berikut:

$$\hat{\beta}_1 = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{(x_i - \overline{x})^2}$$

$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x}$$

di mana

$$\overline{y} = \frac{1}{n} \sum y_i; \text{ dan}$$

$$\overline{x} = \frac{1}{n} \sum x_i$$

Taksiran parameter di atas atau dalam hal ini $\hat{\beta}_0$ dan $\hat{\beta}_1$ telah memenuhi sifat-sifat *Best, Linier, dan Unbiased.* Oleh karena itu, taksiran ini merupakan taksiran yang paling baik untuk model regresi linier sederhana. Selanjutnya, akan diperiksa seberapa baik taksiran parameter yang diperoleh tersebut. Salah satu alat yang digunakan untuk melihat seberapa baik taksiran parameter yang didapat adalah dengan menghitung *Mean Squared Error* (MSE). Berdasarkan definisi pada subbab 2.6.1 diketahui bahwa MSE adalah variansi ditambah dengan bias kuadrat. Oleh karena taksiran yang didapat *unbiased*, maka diperoleh MSE sebagai berikut:

$$MSE = \sigma^2$$

Karena σ^2 merupakan penyimpangan yang terjadi pada populasi, yang nilainya tidak diketahui, maka σ^2 biasanya ditaksir berdasarkan data sampel. Adapun taksirannya sebagai berikut:

$$s^2 = \left[\frac{\left(y_i - \hat{y}_i \right)^2}{n-2} \right].$$

Selanjutnya, MSE ini akan digunakan dalam pembentukan *confidence interval* dan statistik uji untuk parameter $\beta_j, j = 0,1$. Adapun bentuk dari *confidence interval* dan statistik uji untuk parameter β_j adalah sebagai berikut:

- *Confidence interval* untuk parameter β_j

$$\hat{\beta}_j \pm t_{\alpha/2} s_{\hat{\beta}_j}$$

di mana

$$s_{\hat{\beta}_1} = \left\{ \frac{s^2}{\sum \left(x_i - \bar{x} \right)^2} \right\}^{1/2}$$

$$s_{\hat{\beta}_0} = \left\{ \frac{\sum x_i^2}{n \sum \left(x_i - \bar{x} \right)^2} \right\}^{1/2} s$$

Nilai MSE sangat berpengaruh pada *confidence interval* di atas. Semakin kecil nilai MSE yang diperoleh maka semakin sempit *confidence*

Interval yang diperoleh. Dengan demikian, taksiran parameter yang didapat pun akan semakin baik.

❖ Statistik uji untuk hipotesis parameter β_j (uji t)

Uji t merupakan suatu pengujian yang bertujuan untuk mengetahui apakah β_j signifikan atau tidak. Sebelum dilakukan pengujian, biasanya dibuat hipotesis terlebih dahulu, yaitu:

$$H_0 : \beta_j = 0$$
$$H_1 : \text{ tidak demikian}$$

Artinya, berdasarkan data yang tersedia, akan dilakukan pengujian terhadap β_j, apakah sama dengan nol, yang berarti tidak mempunyai pengaruh signifikan terhadap variabel terikat, atau tidak sama dengan nol, yang berarti mempunyai pengaruh signifikan.

Uji t didefinisikan sebagai berikut:

$$t = \frac{\hat{\beta}_j - \beta_j}{s_{\hat{\beta}_j}}$$

Kemudian berdasarkan hipotesis, uji t menjadi:

$$t = \frac{\hat{\beta}_j}{s_{\hat{\beta}_j}}$$

Bila ternyata setelah dihitung $|t| > t_{\alpha/2}$, maka hipotesis nol ditolak. Dalam hal ini dapat dikatakan bahwa β_j signifikan.

Jadi, berdasarkan statistik uji di atas terlihat bahwa apabila nilai dari MSE yang diperoleh semakin kecil maka nilai t yang didapat akan semakin besar, sehingga hipotesis nol akan cenderung ditolak. Hal ini sesuai dengan yang diharapkan, yaitu parameternya mempunyai pengaruh yang signifikan.