# 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 dan bentuk kuadratik, beserta diferensiasinya. Kemudian, dibahas pula mengenai vektor *random* berdistribusi normal. Metode penaksiran parameter menggunakan Metode *Maximum Likelihood* juga dibahas dalam Tugas Akhir ini disertai dengan pendekatan numerik, yaitu *Scoring Algorithm*. Selain itu, dibahas juga mengenai definisi faktor *fixed* dan faktor *random*, definisi *Linear Regression Model* (LRM), *Generalized Linear Regression Model* (GLRM), serta *General Linear Mixed Model* dengan metode penaksiran masing-masing parameternya. Pada akhir Bab II dibahas pula mengenai definisi fungsi genap, fungsi ganjil, dan *translation- 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. $\text{tr}[(k)] = k$
3. $\text{tr}(k\mathbf{A}) = k\,\text{tr}(\mathbf{A})$
4. $\text{tr}(\mathbf{A}+\mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$
5. $\text{tr}(\mathbf{A}^T) = \text{tr}(\mathbf{A})$ **(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

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

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

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

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 $\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA})$.

**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

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

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 ***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 ***W*** dari ***V*** yang terdiri dari semua kombinasi linier dari vektor-vektor dalam ***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 ***W***, sehingga ***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 ***V*** dan $\boldsymbol{S} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ adalah himpunan vektor-vektor dalam ***V***, maka ***S*** disebut basis untuk ***V*** jika dua syarat berikut ini dipenuhi:

1. ***S*** bebas linier
2. ***S*** merentang ***V***.

**Definisi 2.7** Ruang vektor tidak nol ***V*** berdimensi hingga jika ***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 ***V*** berdimensi hingga, yang dinyatakan dengan $\dim(\boldsymbol{V})$, didefinisikan sebagai banyak vektor dalam suatu basis untuk ***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

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

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. $\left|\mathbf{A}^T\right| = |\mathbf{A}|$.

2. $|k\mathbf{A}| = k^n |\mathbf{A}|$.

3. $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$.

4. $\left|\mathbf{A}^T\mathbf{A}\right| = |\mathbf{A}|^2$.

5. $\left|\mathbf{A}^{-1}\right| = 1/|\mathbf{A}|$.

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 $\left|\mathbf{A}_{ij}\right|$ disebut minor dari elemen $a_{ij}$, sedangkan $(-1)^{i+j}\left|\mathbf{A}_{ij}\right|$ disebut kofaktor dari $a_{ij}$. Misalkan $a_{ij}$ menyatakan elemen ke-*ij* dari matriks berukuran $n \times n$, dan misalkan $\alpha_{ij}$ 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 $\mathbf{A} = [a_{ij}]$ 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 $\mathbf{x} = \left(x_1, \ldots, x_n\right)^T$ 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 $\left\{\mathbf{x} \in \Re^{m\times 1} : \left\|\mathbf{x} - \mathbf{c}\right\| < r\right\}$ 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 $\left\{\mathbf{X} \in \Re^{m \times n} : \left\|\mathbf{X} - \mathbf{C}\right\| < r\right\}$.

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 $\varepsilon$, ada *neighborhood* $N_\varepsilon$ dari $\mathbf{c}$ sedemikian sehingga $\left|f(\mathbf{x}) - b\right| < \varepsilon$ di mana $\left\|\mathbf{x} - \mathbf{c}\right\| < \delta$ dengan $\delta$ adalah radius dari *neighborhood* $N_\varepsilon$. 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 $\left(\partial f(\mathbf{x})/\partial x_1, \partial f(\mathbf{x})/\partial x_2, \ldots, \partial f(\mathbf{x})/\partial x_m\right)$ di mana elemennya adalah turunan parsial *first order* dari $f$ di $\mathbf{x}$, sedangkan $\partial f(\mathbf{x})/\partial \mathbf{x}$ menyatakan vektor kolom $\left(\partial f(\mathbf{x})/\partial x_1, \partial f(\mathbf{x})/\partial x_2, \ldots, \partial f(\mathbf{x})/\partial x_m\right)^T$ 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* $\mathbf{c}$ pada $S$ 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 $\mathbf{c}$. 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 $\mathbf{c}$.

Misalkan ada vektor kolom $\mathbf{f} = \left(f_1, f_2, \ldots, f_p\right)^T$ 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. $D_j\mathbf{f}\left(\mathbf{c}\right) = \left[D_j f_1\left(\mathbf{c}\right), D_j f_2\left(\mathbf{c}\right), \ldots, D_j f_p\left(\mathbf{c}\right)\right]^T$ dan
2. $\mathbf{Df}\left(\mathbf{c}\right) = \left[D_1\mathbf{f}\left(\mathbf{c}\right), D_2\mathbf{f}\left(\mathbf{c}\right), \ldots, D_m\mathbf{f}\left(\mathbf{c}\right)\right]$ 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 $\mathbf{x} = \left(x_1, x_2, \ldots, x_m\right)^T$ merupakan vektor dengan $m$ variabel. Kemudian tuliskan $\partial\mathbf{f}\left(\mathbf{x}\right)/\partial\mathbf{x}^T$ (atau $\partial\mathbf{f}/\partial\mathbf{x}^T$) untuk matriks $p\times m$ di mana elemen ke-*sj* adalah $\partial f_s\left(\mathbf{x}\right)/\partial x_j$. Pada konteks ini, $\partial\mathbf{f}\left(\mathbf{x}\right)/\partial\mathbf{x}^T$ disebut turunan dari $\mathbf{f}\left(\mathbf{x}\right)$ 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

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

(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

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

(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

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

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

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

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:

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

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

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

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

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 $\xi_{ij}$ 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 $h(\mathbf{x}) = \det[\mathbf{F}(\mathbf{x})]$. 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 $h(\mathbf{x}) = g[\mathbf{F}(\mathbf{x})]$ telah jelas bahwa $g$ *continuously differentiable* di setiap $\mathbf{Y}$ dengan $\frac{\partial g}{\partial \mathbf{Y}} = [adj(\mathbf{Y})]^T$ .

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 $\det[\mathbf{F}(\mathbf{x})] > 0$. Misalkan $l$ adalah sebuah fungsi dari $\mathbf{x}$ (di $S$) yang didefinisikan sebagai $l(\mathbf{x}) = \log\det[\mathbf{F}(\mathbf{x})]$. 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 $l(\mathbf{x}) = g[h(\mathbf{x})]$. Diketahui dari kalkulus satu variabel bahwa untuk $y > 0$, g *continuously differentiable* di setiap titik di *domain*-nya dengan $\dfrac{\partial \log y}{\partial y} = \dfrac{1}{y}$.

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:

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

$$\begin{aligned} \frac{\partial^2 \mathbf{F}^{-1}}{\partial x_i \partial x_j} &= -\mathbf{F}^{-1} \frac{\partial \mathbf{F}}{\partial x_i \partial x_j} \mathbf{F}^{-1} + \mathbf{F}^{-1} \frac{\partial \mathbf{F}}{\partial x_i} \mathbf{F}^{-1} \frac{\partial \mathbf{F}}{\partial x_j} \mathbf{F}^{-1} \\ &\quad + \mathbf{F}^{-1} \frac{\partial \mathbf{F}}{\partial x_j} \mathbf{F}^{-1} \frac{\partial \mathbf{F}}{\partial x_i} \mathbf{F}^{-1} \end{aligned} \quad \textbf{(2.3.7.b)}$$

(bukti diberikan pada lampiran 5).

Sekarang anggap bahwa $\det[\mathbf{F}(\mathbf{x})] > 0$ 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 Vektor *Random* Normal

Variabel random $X$ dikatakan berdistribusi normal dengan *probability density function* (pdf)-nya adalah sebagai berikut:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], -\infty < x < \infty$$

di mana *mean* $\mu$ dan variansi $\sigma^2$.

Penulisan $f_X(x)$ selanjutnya disederhanakan menjadi $f(x)$. $X$ berdistribusi Normal biasanya direpresentasikan sebagai

$$X \sim N(\mu, \sigma^2)$$

di mana

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$$

$$\sigma^2 = E\left[\left(X-\mu\right)^2\right] = \int_{-\infty}^{\infty}\left(x-\mu\right)^2 f\left(x\right)dx$$

Vektor *random* $X = \left(X_1, X_2, \ldots, X_n\right)^T \in \Re^n$ berdistribusi normal jika pdf-nya adalah sebagai berikut:

$$f_X\left(\boldsymbol{x}\right) = \frac{1}{\left(2\pi\right)^{n/2}\left|\Sigma_{\boldsymbol{x}}\right|^{1/2}}\exp\left\{-\frac{1}{2}\left(\boldsymbol{x}-m_{\boldsymbol{X}}\right)^T \Sigma_{\boldsymbol{x}}^{-1}\left(\boldsymbol{x}-m_{\boldsymbol{X}}\right)\right\} \qquad \textbf{(2.4.a)}$$

di mana

$m_{\boldsymbol{X}} = E\left(\boldsymbol{X}\right)$ adalah vektor *mean* dari vektor *random* $\boldsymbol{X}$,

$\Sigma_{\boldsymbol{X}} = E\left[\left(\boldsymbol{x}-m_{\boldsymbol{X}}\right)\left(\boldsymbol{x}-m_X\right)^T\right]$ adalah matriks kovarians, dan

$n = \dim \boldsymbol{X}$ adalah dimensi dari vektor *random* $\boldsymbol{X}$.

$\boldsymbol{X}$ berdistribusi normal biasanya dinyatakan sebagai

$$\boldsymbol{X} \sim N\left(m_X, \Sigma_{\boldsymbol{X}}\right) \qquad \textbf{(2.4.b)}$$

di mana

$$m_{\boldsymbol{X}} = E\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} m_{X_1} \\ m_{X_2} \\ \vdots \\ m_{X_n} \end{pmatrix}$$

dan

$$\Sigma_{\boldsymbol{X}} = \Sigma_{\boldsymbol{X}}^T = \begin{bmatrix} E\left[\left(X_1 - m_{X_1}\right)^2\right] & E\left[\left(X_1 - m_{X_1}\right)\left(X_2 - m_{X_2}\right)^T\right] & \cdots & E\left[\left(X_1 - m_{X_1}\right)\left(X_n - m_{X_n}\right)^T\right] \\ E\left[\left(X_2 - m_{X_2}\right)\left(X_1 - m_{X_1}\right)^T\right] & E\left[\left(X_2 - m_{X_2}\right)^2\right] & \cdots & E\left[\left(X_2 - m_{X_2}\right)\left(X_n - m_{X_n}\right)^T\right] \\ \vdots & \vdots & \ddots & \vdots \\ E\left[\left(X_n - m_{X_n}\right)\left(X_1 - m_{X_1}\right)^T\right] & E\left[\left(X_n - m_{X_n}\right)\left(X_2 - m_{X_2}\right)^T\right] & \cdots & E\left[\left(X_n - m_{X_n}\right)^2\right] \end{bmatrix}$$

## 2.5 Estimasi Parameter

### 2.5.1 Metode *Maximum Likelihood* (ML)

Misal $\boldsymbol{X} = \left(X_1, X_2, \ldots, X_n\right)^T$ adalah vektor *random* dengan pdf $f\left(\boldsymbol{x};\theta\right)$, $\theta \in \Omega \in \Re^p$ di mana $\theta$ merupakan suatu vektor dari *p*-parameter yang tidak diketahui. Dalam melakukan penaksiran ML ada beberapa tahapan yang harus dilakukan. Pertama, cari *joint* pdf dari $X_1, X_2, \ldots, X_n$, yaitu $f\left(x_1, x_2, \ldots, x_n;\theta\right)$. Oleh karena $\boldsymbol{X}$ adalah vektor random dari $X_1, X_2, \ldots, X_n$ maka $f\left(x_1, x_2, \ldots, x_n;\theta\right) = f\left(\boldsymbol{x};\theta\right)$.

Selanjutnya, cari fungsi *likelihood* yang didefinisikan sebagai *joint* pdf dari $X_1, X_2, \ldots, X_n$ dan dapat dianggap sebagai fungsi dari $\theta$. Misalkan fungsi *likelihood* $L\left(\theta; x_1, x_2, \ldots, x_n\right) = L\left(\theta;\boldsymbol{x}\right)$, maka

$$L\left(\theta;\boldsymbol{x}\right) = f\left(\boldsymbol{x};\theta\right). \qquad \textbf{(2.5.1.a)}$$

Kemudian, cari taksiran dari $\theta$. Dalam Metode penaksiran ML, taksiran dari $\theta$ diperoleh dengan menemukan nilai $\theta$, sebut $\hat{\theta}$, yang memaksimumkan fungsi *likelihood*. Maka $\hat{\theta}$ ini disebut taksiran ML dari $\theta$.

Mencari nilai $\theta$ yang memaksimumkan fungsi $\ln L(\theta; \boldsymbol{x})$, sebut $l(\theta; \boldsymbol{x})$, akan memberikan hasil yang sama dengan mencari nilai $\theta$ yang memaksimumkan $L(\theta; \boldsymbol{x})$ (lihat lampiran 6). Maka baik $L(\theta; \boldsymbol{x})$ atau $l(\theta; x)$ dapat digunakan untuk mencari nilai $\hat{\theta}$.

Nilai $\theta$ yang memaksimumkan $l(\theta; x)$ dapat diperoleh dengan mencari solusi simultan dari persamaan

$$\begin{aligned} S_j(\theta) &= \frac{\partial}{\partial \theta_j} l(\theta; \boldsymbol{x}), \text{ untuk } j = 1, 2, \ldots, p \\ &= \frac{\partial}{\partial \theta_j} \ln L(\theta; \boldsymbol{x}) \qquad \textbf{(2.5.1.b)} \\ &= \frac{1}{L(\theta; \boldsymbol{x})} \frac{\partial}{\partial \theta_j} L(\theta; \boldsymbol{x}) = 0. \end{aligned}$$

Adakalanya sistem persamaan ini dapat diselesaikan secara analitik. Jika tidak, suatu Metode numerik (misal: Metode Newton-Raphson) dapat digunakan (lihat lampiran 7).

### 2.5.2 *Scoring Algorithm*

*Scoring Algorithm* didapat dengan mengganti $\mathbf{H}\left(\hat{\delta}^{(m)}\right)$ pada Metode Newton-Raphson dari lampiran 7, dengan nilai ekspektasinya sehingga menjadi

$$\hat{\delta}^{(m+1)} = \hat{\delta}^{(m)} - \left(E\left(\mathbf{H}\left(\hat{\delta}^{(m)}\right)\right)\right)^{-1} . \mathbf{s}\left(\hat{\delta}^{(m)}\right); m = 0,1,\ldots \qquad \textbf{(2.5.2.a)}$$

## 2.6 Definisi Faktor *Fixed* dan Faktor *Random*

Berikut ini adalah beberapa istilah yang berhubungan dengan faktor *fixed* dan faktor *random*.

- Eksperimen : Sebuah proses untuk mengumpulkan data sampel.
- Variabel respon : Variabel yang diukur dalam eksperimen.
- Faktor : Variabel prediktor (kuantitatif atau kualitatif) yang berhubungan dengan variabel respon.
- *Level* : Isi atau tingkatan dalam suatu faktor (contohnya: nilai yang diasumsikan dalam sebuah faktor dari sebuah eksperimen).
- *Treatment* : Kombinasi khusus dari level-level faktor yang dilakukan dalam sebuah eksperimen.
- Faktor *Fixed* : Faktor-faktor yang level-levelnya dipilih dengan tujuan tertentu yang akan diuji oleh peneliti di mana kesimpulan statistiknya hanya terbatas pada level-level

tersebut saja.

- Faktor *Random* : Faktor-faktor yang level-levelnya dipilih secara acak dari populasi level yang ada dan kesimpulan statistiknya mengenai populasi dari level faktor di mana data tersebut diasumsikan berasal.

Untuk variabel kuantitatif, level berhubungan dengan nilai numerik yang diasumsikan, contohnya: jika jumlah kesalahan dalam suatu produk berkisar antara 0 sampai 3 maka variabel independennya diasumsikan memiliki empat level, yaitu 0, 1, 2, dan 3. Sedangkan untuk variabel kualitatif, level-levelnya dapat didefinisikan hanya dengan menggambarkannya saja, contohnya: variabel independen dari bentuk kemasan yang diobservasi memiliki level A, B, dan C.

Contoh faktor *fixed*:

Pada suatu penelitian kesehatan, terdapat 100 macam merk obat baru untuk penyakit HIV AIDS. Peneliti ingin menganalisa apakah ada pengaruh antara obat merk A, B, C, dan D dalam penyembuhan penyakit HIV AIDS.

Contoh faktor *random*:

Pada suatu penelitian mengenai penyakit HIV AIDS di Jakarta, peneliti ingin menganalisa apakah ada pengaruh antara perawatan rumah sakit dengan penyembuhan penyakit tersebut. Dalam hal ini, peneliti mengambil 15 rumah sakit secara acak dari 100 rumah sakit yang ada di Jakarta.

Contoh faktor *mixed* (*fixed* dan *random*):

Pada suatu penelitian mengenai penyakit HIV AIDS di Jakarta, terdapat 100 macam merk obat baru untuk penyakit ini. Peneliti ingin menganalisa apakah ada pengaruh antara obat merk A, B, C, dan D serta perawatan rumah sakit dengan penyembuhan penyakit tersebut. Dalam hal ini, peneliti mengambil 15 rumah sakit secara acak dari 100 rumah sakit yang ada di Jakarta.

## 2.7 ***Linear Regression Model*, *Generalized Linear Regression Model*, dan *General Linear Mixed Model* beserta Metode-Metode Penaksirannya**

### 2.7.1 ***Linear Regression Model* (LRM)**

Analisis regresi adalah suatu metode yang digunakan dalam menganalisis satu atau lebih variabel prediktor $\mathbf{X}$ dengan satu variabel respon $\mathbf{Y}$. Pola hubungan itu dapat dinyatakan dalam bentuk persamaan regresi. Model regresi linier yang melibatkan lebih dari satu variabel prediktor dengan satu variabel respon disebut model regresi linier berganda. Berikut ini akan diberikan tabel ilustrasi data untuk regresi linier berganda, yaitu:

| Pengamatan | $i$ | $y$ | $x_1$ | $x_2$ | $\ldots$ | $x_k$ |
|---|---|---|---|---|---|---|
| | 1 | $y_1$ | $x_{11}$ | $x_{12}$ | $\ldots$ | $x_{1k}$ |
| | 2 | $y_2$ | $x_{21}$ | $x_{22}$ | $\ldots$ | $x_{2k}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | | $\vdots$ |
| | $n$ | $y_n$ | $x_{n1}$ | $x_{n2}$ | $\ldots$ | $x_{nk}$ |

Model regresi liniernya dapat dituliskan dalam bentuk berikut:

$$\begin{aligned} y_i &= \alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \ldots + \alpha_k x_{ik} + \varepsilon_i \\ &= \alpha_0 + \sum_{j=1}^{k} \alpha_j x_{ij} + \varepsilon_i \end{aligned} \qquad \textbf{(2.7.1.a)}$$

di mana $i = 1, 2, \ldots, n$ ialah banyaknya pengamatan..

Apabila dinyatakan dalam notasi matriks, maka persamaan (2.7.1.a) menjadi:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\alpha} + \boldsymbol{\varepsilon} \qquad \textbf{(2.7.1.b)}$$

di mana

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \quad \boldsymbol{\alpha} = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_k \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

dengan:

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

$\mathbf{X}$ : matriks *full rank* berukuran $n \times p$, di mana $p = k + 1$ dari variabel prediktor yang elemen-elemennya diketahui.

$\alpha$ : vektor parameter bersifat *fixed* yang tidak diketahui dan tidak terobservasi berukuran $p \times 1$ di mana $p = k + 1$.

$\varepsilon$ : vektor *random error* yang tidak terobservasi berukuran $n \times 1$.

Model regresi linier berganda di atas memiliki asumsi sebagai berikut:

1. $E(\varepsilon_i) = 0$ untuk $i = 1, 2, \ldots, n$
2. $E(\varepsilon_i \varepsilon_j) = 0$ untuk $i \neq j$
   $= \sigma^2$ untuk $i = j$
3. $\varepsilon_i \sim N(0, \sigma^2)$

Dalam notasi matriks dinyatakan dengan:

1. $E(\boldsymbol{\varepsilon}) = \mathbf{0}$
2. $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T) = \sigma^2 \mathbf{I}$ $\mathbf{I}$ : matriks identitas
3. $\boldsymbol{\varepsilon}$ berdistribusi normal dengan mean $\mathbf{0}$ dan variansi $\sigma^2 \mathbf{I}$.

Untuk menaksir koefisien regresi atau parameter model regresi dengan menggunakan Metode *Ordinary Least Squares* (OLS) maka asumsi-asumsi di atas harus terpenuhi. Fungsi *least squares* atau *Sum of Squares of Error* (SSE) dinyatakan dengan:

$$S = S(\alpha_0, \alpha_1, \ldots, \alpha_k) = \sum_{i=1}^{n} \varepsilon_i^2$$

$$= \sum_{i=1}^{n} \left( y_i - \alpha_0 - \sum_{j=1}^{k} \alpha_j x_{ij} \right)^2 \qquad \textbf{(2.7.1.c)}$$

Fungsi *S* akan diminimumkan terhadap $\alpha_0, \alpha_1, \ldots, \alpha_k$.

Taksiran *least squares* dari $\alpha_0, \alpha_1, \ldots, \alpha_k$ harus memenuhi

$$\left.\frac{\partial S}{\partial \alpha_0}\right|_{\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_k} = -2\sum_{i=1}^{n} \left( y_i - \hat{\alpha}_0 - \sum_{j=1}^{k} \hat{\alpha}_j x_{ij} \right) = 0 \qquad \textbf{(2.7.1.d)}$$

$$\left.\frac{\partial S}{\partial \alpha_j}\right|_{\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_k} = -2\sum_{i=1}^{n} \left( y_i - \hat{\alpha}_0 - \sum_{j=1}^{k} \hat{\alpha}_j x_{ij} \right) x_{ij} = 0 \qquad (\textbf{2.7.1.e})$$

untuk setiap $j = 1, 2, \ldots, k$.

Dari persamaan (2.7.1.d) dan (2.7.1.e), diperoleh sistem persamaan normal:

$$\begin{array}{ccccccccccc}
n\hat{\alpha}_0 & + & \hat{\alpha}_1 \sum_{i=1}^{n} x_{i1} & + & \hat{\alpha}_2 \sum_{i=1}^{n} x_{i2} & + & \cdots & + & \hat{\alpha}_k \sum_{i=1}^{n} x_{ik} & = & \sum_{i=1}^{n} y_i \\
\hat{\alpha}_0 \sum_{i=1}^{n} x_{i1} & + & \hat{\alpha}_1 \sum_{i=1}^{n} x_{i1}^2 & + & \hat{\alpha}_2 \sum_{i=1}^{n} x_{i1} x_{i2} & + & \cdots & + & \hat{\alpha}_k \sum_{i=1}^{n} x_{i1} x_{ik} & = & \sum_{i=1}^{n} x_{i1} y_i \\
\vdots & & \vdots & & \vdots & & & & \vdots & & \vdots \\
\hat{\alpha}_0 \sum_{i=1}^{n} x_{ik} & + & \hat{\alpha}_1 \sum_{i=1}^{n} x_{ik} x_{i1} & + & \hat{\alpha}_2 \sum_{i=1}^{n} x_{ik} x_{i2} & + & \cdots & + & \hat{\alpha}_k \sum_{i=1}^{n} x_{ik}^2 & = & \sum_{i=1}^{n} x_{ik} y_i
\end{array}$$

Dalam notasi matriks persamaan (2.7.1.c) adalah sebagai berikut:

$$
\begin{aligned}
S(\alpha) &= \sum_{i=1}^{n} \varepsilon_i^2 \\
&= \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} \\
&= (\mathbf{y} - \mathbf{X}\alpha)^T (\mathbf{y} - \mathbf{X}\alpha) \\
&= \mathbf{y}^T\mathbf{y} - \alpha^T \mathbf{X}^T \mathbf{y} - \mathbf{y}^T \mathbf{X}\alpha + \alpha^T \mathbf{X}^T \mathbf{X}\alpha \\
&= \mathbf{y}^T\mathbf{y} - 2\alpha^T \mathbf{X}^T \mathbf{y} + \alpha^T \mathbf{X}^T \mathbf{X}\alpha.
\end{aligned}
$$

Karena $\alpha^T \mathbf{X}^T \mathbf{y}$ merupakan matriks berukuran $1 \times 1$ atau skalar dan

*transpose*-nya $\left(\alpha^T \mathbf{X}^T \mathbf{y}\right)^T = \mathbf{y}^T \mathbf{X}\alpha$ juga skalar yang sama, maka:

$$
\left.\frac{dS}{d\alpha}\right|_{\hat{\alpha}} = -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X}\hat{\alpha} = \mathbf{0} \text{ sehingga, } \mathbf{X}^T \mathbf{X}\hat{\alpha} = \mathbf{X}^T \mathbf{y} \qquad \textbf{(2.7.1.f)}
$$

Dengan mengalikan kedua ruas pada persamaan (2.7.1.f) dengan invers dari matriks $\mathbf{X}^T \mathbf{X}$ yang *nonsingular*, didapat penyelesaian sistem persamaan normal (2.7.1.f) yang memberikan taksiran *least squares* $\alpha$, yaitu:

$$
\hat{\alpha} = \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{y} \qquad \textbf{(2.7.1.g)}
$$

Sifat-sifat dari taksiran *least squares*:

$$
\begin{aligned}
E(\hat{\alpha}) &= E\left[\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{y}\right] \\
&= E\left[\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T (\mathbf{X}\alpha + \boldsymbol{\varepsilon})\right] \\
&= E\left[\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{X}\alpha + \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \boldsymbol{\varepsilon}\right] \\
&= \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{X} E(\alpha) + \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T E(\boldsymbol{\varepsilon}) \\
&= \alpha
\end{aligned}
$$

Karena $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ dan $\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{X} = \mathbf{I}$ Sehingga $\hat{\alpha}$ merupakan taksiran *unbiased* untuk $\alpha$. Sedangkan, matriks kovariansi untuk $\hat{\alpha}$ ialah

$$\begin{aligned} \text{Cov}(\hat{\alpha}) &= E\left[\left(\hat{\alpha} - E(\hat{\alpha})\right)\left(\hat{\alpha} - E(\hat{\alpha})\right)^T\right] \\ &= E\left[(\hat{\alpha} - \alpha)(\hat{\alpha} - \alpha)^T\right] \end{aligned} \qquad \textbf{(2.7.1.h)}$$

Dari persamaan (2.7.1.g), dengan mensubstitusikan $\mathbf{y} = \mathbf{X}\alpha + \boldsymbol{\varepsilon}$ diperoleh

$$\begin{aligned} \hat{\alpha} &= \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T (\mathbf{X}\alpha + \boldsymbol{\varepsilon}) \\ &= \alpha + \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \boldsymbol{\varepsilon} \\ \hat{\alpha} - \alpha &= \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \boldsymbol{\varepsilon} \end{aligned} \qquad \textbf{(2.7.1.i)}$$

Substitusikan persamaan (2.7.1.i) ke (2.7.1.h) sehingga diperoleh:

$$\begin{aligned} \text{Cov}(\hat{\alpha}) &= E\left[\left(\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \boldsymbol{\varepsilon}\right)\left(\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \boldsymbol{\varepsilon}\right)^T\right] \\ &= E\left[\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T \mathbf{X}\left(\mathbf{X}^T \mathbf{X}\right)^{-1}\right] \\ &= \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T E\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\right) \mathbf{X}\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \\ &= \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \sigma^2 \mathbf{I} \mathbf{X}\left(\mathbf{X}^T \mathbf{X}\right)^{-1} \\ &= \sigma^2 \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \end{aligned}$$

### 2.7.2 *Generalized Linear Regression Model* (GLRM)

Dari (2.7.1.b), asumsi-asumsi yang biasa digunakan dalam LRM adalah $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ dan $V(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}$ . Asumsi variansi *error* $\sigma^2\mathbf{I}$ disebut asumsi variansi *error spherical*. Seringkali asumsi-asumsi ini tidak dapat dipenuhi sehingga akan dibahas modifikasi terhadap langkah-langkah OLS yang diperlukan bila $V(\boldsymbol{\varepsilon}) = \boldsymbol{\sigma}^2\Omega,$ dengan $\boldsymbol{\Omega}$ adalah matriks yang diketahui, berukuran $n \times n$ dan diasumsikan *continuously differentiable* dua kali. Asumsi variansi *error* $\boldsymbol{\sigma}^2\Omega$ disebut asumsi variansi *error nonspherical*. Model (2.7.1.b) yang memiliki asumsi variansi *error nonspherical* disebut GLRM.

Oleh karena asumsi dalam Metode OLS tidak terpenuhi, maka taksiran yang didapat sebelumnya (2.7.1.g) tidak dapat digunakan. Dengan demikian, akan dilakukan pendekatan untuk masalah ini dengan melakukan transformasi model untuk kumpulan pengamatan yang baru agar dapat dipenuhi asumsi-asumsi pada Metode OLS sehingga metode tersebut dapat digunakan untuk data yang telah ditransformasi. Oleh karena $\sigma^2\Omega$ adalah matriks kovarians dari *error*, maka $\Omega$ harus definit positif sehingga terdapat matriks **K** yang simetris dan *nonsingular* berukuran $n \times n$ , di mana $\mathbf{K}^T\mathbf{K} = \mathbf{K}\mathbf{K} = \boldsymbol{\Omega}.$ Matriks **K** disebut *square root* dari $\boldsymbol{\Omega}$ .

Akan didefinisikan variabel yang baru, yaitu:

$$\mathbf{z} = \mathbf{K}^{-1}\mathbf{y}, \quad \mathbf{B} = \mathbf{K}^{-1}\mathbf{X}, \quad \mathbf{g} = \mathbf{K}^{-1}\boldsymbol{\varepsilon}$$

sedemikian sehingga model regresi $\mathbf{y} = \mathbf{X}\alpha + \boldsymbol{\varepsilon}$ menjadi

$$\mathbf{K}^{-1}\mathbf{y} = \mathbf{K}^{-1}\mathbf{X}\alpha + \mathbf{K}^{-1}\boldsymbol{\varepsilon},$$

atau dapat ditulis ke dalam bentuk

$$\mathbf{z} = \mathbf{B}\alpha + \mathbf{g}.$$

*Error* dalam model yang ditransformasi ini memiliki *mean* nol, yaitu $E(\mathbf{g}) = \mathbf{K}^{-1}E(\boldsymbol{\varepsilon}) = \mathbf{0}$. Sedangkan, matriks kovariansi untuk **g** adalah:

$$\begin{aligned} V(\mathbf{g}) &= E\left[(\mathbf{g} - E(\mathbf{g}))(\mathbf{g} - E(\mathbf{g}))^T\right] \\ &= E(\mathbf{g}\mathbf{g}^T) \\ &= E(\mathbf{K}^{-1}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\mathbf{K}^{-1}) \\ &= \mathbf{K}^{-1}E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T)\mathbf{K}^{-1} \\ &= \sigma^2\mathbf{K}^{-1}\Omega\mathbf{K}^{-1} \\ &= \sigma^2\mathbf{K}^{-1}\mathbf{K}\mathbf{K}\mathbf{K}^{-1} \\ &= \sigma^2\mathbf{I} \end{aligned}$$

Dengan demikian, **g** memiliki *mean* nol dan variansi konstan serta tidak berkorelasi. Oleh karena error **g** dalam model $\mathbf{z} = \mathbf{B}\alpha + \mathbf{g}$ telah memenuhi asumsi tersebut, maka dapat diterapkan Metode OLS. Fungsi *least squares*-nya adalah sebagai berikut:

$$\begin{aligned} S(\alpha) &= \mathbf{g}^T\mathbf{g} \\ &= \boldsymbol{\varepsilon}^T\boldsymbol{\Omega}^{-1}\boldsymbol{\varepsilon} \\ &= (\mathbf{y} - \mathbf{X}\alpha)^T\boldsymbol{\Omega}^{-1}(\mathbf{y} - \mathbf{X}\alpha) \end{aligned}$$

dan diperoleh persamaan normal *least squares* $(\mathbf{X}^T\boldsymbol{\Omega}^{-1}\mathbf{X})\hat{\alpha} = \mathbf{X}^T\boldsymbol{\Omega}^{-1}\mathbf{y}$

serta penyelesaian untuk persamaan tersebut, yaitu:

$$\hat{\alpha} = \left(\mathbf{X}^T\, \boldsymbol{\Omega}^{-1}\, \mathbf{X}\right)^{-1} \mathbf{X}^T\, \boldsymbol{\Omega}^{-1}\, \mathbf{y} \qquad \textbf{(2.7.2.a)}$$

di mana $\hat{\alpha}$ pada persamaan (2.7.2.a) dikatakan taksiran *Generalized least squares* (GLS) untuk $\alpha$.

### 2.7.3 *General Linear Mixed Model*

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

$$\mathbf{y} = \mathbf{X}\alpha + \mathbf{Z}\mathbf{b} + \mathbf{e} \qquad \textbf{(2.7.3.a)}$$

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.

$\alpha$ : 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

$E(\mathbf{b}) = \mathbf{0}$, $E(\mathbf{e}) = \mathbf{0}$, $Var(\mathbf{b}) = E(\mathbf{bb}^T) = \mathbf{G}$, $Var(\mathbf{e}) = E(\mathbf{ee}^T) = \mathbf{R}$, 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 $\text{cov}(\mathbf{b}, \mathbf{e}) = E(\mathbf{be}^T) = E(\mathbf{eb}^T) = \mathbf{0}$. Berikut ini diberikan bentuk dari matriks **G** dan **R**:

$$\mathbf{G}(\delta) = \begin{pmatrix} g_{11}(\delta) & g_{12}(\delta) & \cdots & g_{1h}(\delta) \\ g_{21}(\delta) & g_{22}(\delta) & \cdots & g_{2h}(\delta) \\ \vdots & \vdots & \ddots & \vdots \\ g_{h1}(\delta) & g_{h2}(\delta) & \cdots & g_{hh}(\delta) \end{pmatrix}$$

$$\mathbf{R}(\delta) = \begin{pmatrix} r_{11}(\delta) & r_{12}(\delta) & \cdots & r_{1n}(\delta) \\ r_{21}(\delta) & r_{22}(\delta) & \cdots & r_{2n}(\delta) \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1}(\delta) & r_{n2}(\delta) & \cdots & r_{nn}(\delta) \end{pmatrix}$$

Prosedur penaksiran parameter dengan menggunakan Metode *Best Linear Unbiasd Prediction* (BLUP) dimulai dengan memisalkan nilai-nilai dari vektor *random* $\mathbf{b}$ yang dapat 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 $h \times 1$. $T(\mathbf{y})$ dapat mengestimasi kombinasi linier tersebut jika $T(\mathbf{y})$ merupakan penaksir yang *unbiased* dan linier di mana definisi masing-masingnya adalah sebagai berikut, $T(\mathbf{y})$ disebut *unbiased* jika $E(T(\mathbf{y})) = E(\boldsymbol{\lambda}^T \boldsymbol{\alpha} + \boldsymbol{\omega}^T \mathbf{b})$, dan disebut linier jika $T(a\mathbf{y}_1 + b\mathbf{y}_2) = aT(\mathbf{y}_1) + bT(\mathbf{y}_2)$.

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 8). **(2.7.3.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 $E\left[\left(T(\mathbf{y}) - \boldsymbol{\lambda}^T \boldsymbol{\alpha} - \boldsymbol{\omega}^T \mathbf{b}\right)^2\right]$ sebagai *Mean Squared Error* (MSE) dari penaksir $T(\mathbf{y})$.

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

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\delta}, \mathbf{y}) = \left(\mathbf{X}^T \boldsymbol{\Omega}^{-1}(\boldsymbol{\delta}) \mathbf{X}\right)^{-1} \mathbf{X}^T \boldsymbol{\Omega}^{-1}(\boldsymbol{\delta}) \mathbf{y} \qquad \textbf{(2.7.3.c)}$$

dan

$$\hat{\boldsymbol{\beta}}(\boldsymbol{\delta}, \mathbf{y}) = \mathbf{G}(\boldsymbol{\delta}) \mathbf{Z}^T \boldsymbol{\Omega}^{-1}(\boldsymbol{\delta}) \left(\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\alpha}}(\boldsymbol{\delta}, \mathbf{y})\right) \qquad \textbf{(2.7.3.d)}$$

di mana

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

Dengan demikian*, Best Linear Unbiased Predictor* (BLUP) dari $\tau$ adalah

$$\hat{\tau}(\delta, \mathbf{y}) = \boldsymbol{\lambda}^T \hat{\boldsymbol{\alpha}}(\delta, \mathbf{y}) + \boldsymbol{\omega}^T \hat{\boldsymbol{\beta}}(\delta, \mathbf{y}) \qquad \textbf{(2.7.3.e)}$$

yang memiliki MSE terkecil dari semua taksiran yang *unbiased* dan linier. Detail penurunan rumus dapat dilihat pada Tugas Akhir Phydelya dengan judul "Penggunaan Metode *Best Linear Unbiased Prediction* pada *Generalized Linear Mixed Model'*.

## 2.8 Definisi Fungsi Genap, Fungsi Ganjil, dan *Translation-Invariant*

### 2.8.1 Fungsi Genap dan Fungsi Ganjil

Sebuah fungsi $f(\mathbf{y})$ dari vektor $\mathbf{y}$ untuk $\mathbf{y} \in \mathrm{W}$ dikatakan fungsi genap jika

$$f(-\mathbf{y}) = f(\mathbf{y}) \text{ untuk } \mathbf{y} \in \mathrm{W}, \qquad \textbf{(2.8.1.a)}$$

dan dikatakan fungsi ganjil jika

$$f(-\mathbf{y}) = -f(\mathbf{y}) \text{ untuk } \mathbf{y} \in \mathrm{W}. \qquad \textbf{(2.8.1.b)}$$

### 2.8.2 *Translation Invariant*

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

$$f(\mathbf{y}-\mathbf{Xa}) = f(\mathbf{y}) \text{ untuk setiap } \mathbf{y} \text{ dan } \mathbf{a}. \qquad \textbf{(2.8.2.a)}$$