BAB II

TEORI PENUNJANG

2.1. Vektor

Definisi 2.1.1:

Vektor adalah suatu kuantitas (besaran) yang mempunyai besar dan arah.

Contoh: kecepatan, percepatan, gaya, berat, pergeseran titik.

Definisi 2.1.2 :

Penjumlahan dua vektor $\mathbf{a} = [a_1, a_2, \ldots, a_n]$ dan $\mathbf{b} = [b_1, b_2, \ldots, b_n]$ adalah vektor, berlaku :

$$\mathbf{a} + \mathbf{b} = [a_1, a_2, \ldots, a_n] + [b_1, b_2, \ldots, b_n]$$

$$= [a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n]$$

Definisi 2.1.3 :

Perkalian vektor $\mathbf{a} = [a_1, a_2, \ldots, a_n]$ dengan skalar k adalah vektor.

$$k\mathbf{a} = k[a_1, a_2, \ldots, a_n]$$

$$= [ka_1, ka_2, \ldots, ka_n]$$

2.2. Matriks

Definisi 2.2.1:

Matriks adalah kumpulan bilangan yang disusun secara empat persegi panjang (menurut baris dan kolom).

Bilangan tersebut disebut elemen-elemen matriks. Secara umum sebuah matriks **A** = (a_{ij}), i = 1, 2, ..., m dan j = 1, 2, ..., n. Berarti banyaknya baris m dan banyaknya kolom n.

Contoh:

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

Definisi 2.2.2 :

Pandang **A** = (a_{ij}) dan **B** = (b_{ij}) adalah matriks yang berukuran sama. Jumlah dari matriks **A** dan **B** adalah matriks **C** = (c_{ij}) yang berukuran sama, dengan $c_{ij} = a_{ij} + b_{ij}$ untuk setiap i dan j, atau **A** + **B** = $(a_{ij} + b_{ij})$.

Definisi 2.2.3 :

A = (a_{ij}) adalah suatu matriks dan k adalah suatu skalar, hasil suatu k**A** adalah suatu matriks yang diperoleh dengan mengalikan semua elemen matriks **A** dengan k.

Contoh :

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \text{ maka } 2\mathbf{A} = 2\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 4 & 8 \end{bmatrix}$$

Definisi 2.2.4:

Pandang **A** = (a_{ij}) berukuran p x q dan **B** = (b_{ij}) berukuran q x r . Perkalian matriks **AB** adalah suatu matriks **C** = (c_{ij}) berukuran p x r dengan $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{iq}b_{qj}$, untuk setiap i = 1, 2 , . . . , p dan j = 1, 2 , . . . , r

Syarat perkalian matriks adalah banyaknya kolom matriks pertama sama dengan banyaknya baris matriks ke dua.

Contoh :

$$\mathbf{A} = [\,3\ 2\ 1\,] \text{ dan } \mathbf{B} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$$

Karena banyaknya kolom matriks **A** = 3 dan baris matriks **B** = 3 maka **A** x **B** ada dan berukuran 1 x 1.

$$\mathbf{A} \times \mathbf{B} = [\,3\ 2\ 1\,] \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$$

$$= [\,3.3 + 2.1 + 1.0\,]$$

$$= [\,11\,]$$

Definisi 2.2.5 :

Pandang $\mathbf{A} = (a_{ij})$ berukuran m x n . Tranpose dari **A** adalah $\mathbf{A}^T = (a_{ji})$ berukuran n x m, untuk i = 1, 2 , . . . , m dan j = 1, 2 , . . . , n

Definisi 2.2.6 :

A adalah matriks yang mempunyai jumlah baris dan kolom sama maka **A** disebut matriks bujursangkar.

Definsi 2.2.7 :

A matriks bujursangkar yang mempunyai ukuran n x n, dikatakan simetri jika $\mathbf{A} = \mathbf{A}^T$, dengan $a_{ij} = a_{ji}$, untuk i = 1, 2 , . . . , n dan j = 1, 2 , . . . , n.

2.3. Determinan

2.3.1. Permutasi

Definisi 2.3.1.1 :

Barisan bilangan-bilangan ($j_1, j_2, \ldots, j_n$) dengan $j_i \neq j_k$, untuk $i \neq k$ (i dan k = 1, 2, . . . n) dan j_i salah satu dari bilangan asli (1, 2, . . . , n) disebut suatu permutasi.

Definisi 2.3.1.2 :

Yang dimaksud sebuah inversi pada suatu permutasi ($j_1, j_2, \ldots, j_n$) adalah adanya j_k yang mendahului j_i padahal $j_i < j_k$, (i dan k = 1, 2, . . . ,n).

Contoh :

(2, 1, 4, 3) maka terdapat dua inversi :

1. 2 mendahului 1
2. 4 mendahului 3

Definisi 2.3.1.3 :

Misalkan ($j_1, j_2, \ldots, j_n$) suatu permutasi maka tanda dari permutasi tersebut ditulis dengan $b(j_1, j_2, \ldots, j_n) = (+1)$ jika genap dan (-1) jika ganjil.

Catatan :

dalam pemakaiannya (+1) sering ditulis dengan (+) dan (-1) ditulis (-)

Definsi 2.3.1.4 :

Determinan dari matriks bujursangkar **A** yang berukuran n x n adalah jumlah dari semua n! hasil kali bertanda dari elemen-elemen matriks **A** itu .

Det (**A**) = $|A| = \sum b(j_1, j_2, \dots, j_n)\, a_{1j_1} a_{2j_2} \dots a_{nj_n}$.

Terdapat suatu sifat determinan yang penting yaitu misalkan **A** adalah matriks bujursangkar berukuran n x n dan k adalah sembarang skalar, maka:

det(k**A**) = k^n det(**A**).

dengan menggunakan definisi-definisi dari 2.3.1 dapat dijelaskan sifat determinan tersebut.

Misal **A** = $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$

|**A**| $= \sum b(j_1, j_2, \dots, j_n)\, a_{1j_1} a_{2j_2} \dots a_{nj_n}$

dan misalkan |**B**| adalah |k**A**| maka

|**B**| $= \sum b(j_1, j_2, \dots, j_n)\, ka_{1j_1} ka_{2j_2} \dots ka_{nj_n}$

$= k^n \sum b(j_1, j_2, \dots, j_n)\, a_{1j_1} a_{2j_2} \dots a_{nj_n}$

$= k^n$ |**A**|

2.4. Distribusi Gaussian.

Fungsi densitas distribusi Gaussian multivariat dibentuk dari Gaussian univariat dengan dimensi p≥2. Distribusi Gaussian univariat dengan rataan μ dan variansi σ^2 mempunyai fungsi densitas:

$$f(x) = \frac{1}{(2\pi\sigma^2)^{0,5}} e^{-0,5\left(\frac{x-\mu}{\sigma}\right)^2} \quad , -\infty \leq x \leq \infty \qquad (2.4.1)$$

plot dari fungsi densitas ini berupa kurva dengan bentuk bel yang mempunyai luasan satu.

Fungsi densitas Gaussian dengan rataan μ dan variansi σ^2 sering ditulis dengan $N(\mu,\sigma^2)$. Jadi N(10,4) menunjukkan fungsi densitas Gaussian dengan rataan 10 dan σ = 2. Notasi ini kemudian akan diperluas untuk kasus multivariat.

Bentuk persamaan di bawah ini:

$$\left(\frac{x-\mu}{\sigma}\right)^2 = (x-\mu)(\sigma^2)(x-\mu) \qquad (2.4.2)$$

merupakan bagian dalam eksponensial dari Gaussian univariat. Persamaan (2.5.2) dapat dibentuk untuk vektor **x** yang mempunyai dimensi p x 1

$$(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu) \qquad (2.4.3)$$

vektor μ (p x 1) menunjukkan nilai harapan vektor acak **x** dan matriks Σ (p x p) menunjukkan kovariansi matriks, dan diasumsikan bahwa Σ adalah definit posistif. Densitas Gaussian multivariat dihasilkan dengan mengganti persamaan (2.4.2) dengan (2.4.3) dan memasukkannya ke dalam fungsi (2.4.1). Jika penggantian ini telah dibuat, maka konstanta normal univariat $(2\pi)^{-0,5}(\sigma^2)^{-0,5}$ harus di ubah ke dalam bentuk yang lebih umum, dengan membuat volume di bawah luasan

$$e^{-0,5(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)} = 1$$

maka diperlukan konstanta $(2\pi)^{\frac{-p}{2}}|\Sigma|^{\frac{1}{2}}$. Dengan demikian densitas Gaussian multivariat dengan p variat mempunyai bentuk

$$f(x) = \frac{1}{(2\pi)^{\frac{p}{2}}|\Sigma|^{0,5}} \mathrm{ekp}\left(-0,5(x-\mu)^T(\Sigma^{-1})(x-\mu)\right)$$

$-\infty \le x \le \infty$, i = 1,2, . . .,p.

Yang dinyatakan dengan $N_p(\mu, \Sigma)$ dengan

$$(\mathbf{x} - \mu) = \begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \\ \vdots \\ x_p-\mu_p \end{bmatrix} \text{ dan kovariansi matriks } \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1p} & \sigma_{2p} & \cdots & \sigma_{pp} \end{bmatrix}$$

2.5. Taksiran Maksimum Likelihood

Prosedur taksiran maksimum likelihood, menguji apakah taksiran parameter yang tak diketahui dari fungsi likelihood suatu sampel nilainya sudah memaksimumkan fungsi itu.

Definisi 2.5.1:

Asumsikan $X_1, X_2, \ldots, X_n$ adalah sampel acak independen dari peubah acak X yang mempunyai probabilitas P(x) atau fungsi densitas f(x) dan tergantung pada parameter yang tak diketahui, misalnya θ, Maka fungsi likelihoodnya adalah

$$L(\theta|X) = P_x(x_1)\,P_x(x_2)\,...\,P_x(x_n) \quad \text{, jika X diskrit}$$

$$= f_x(x_1)\,f_x(x_2)\,...\,f_x(x_n) \quad \text{, jika X kontinu.}$$

Dengan $P_x(x)$ dan $f_x(x)$ secara berturut-turut menyatakan fungsi probabilitas dan fungsi densitas.

Pandang X adalah peubah acak diskrit, dan $X_1, X_2, ..., X_n$ adalah sampel acak dari X; $x_1, x_2, ..., x_n$ adalah nilai-nilai pengamatan pada sampel. Maka probabilitas nilai pengamatan yang terjadi dalam sampel adalah:

$$P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n) = P(X_1 = x_1), P(X_2 = x_2), ..., P(X_n = x_n)$$

$$P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n) = P_x(x_1).P_x(x_2)\,...\,P_x(x_n)$$

$$= L(\theta|X)$$

Fungsi likelihood yaitu $L(\theta|X)$ menyatakan probabilitas nilai pengamatan yang dihasilkan sebagai fungsi θ. Jika X peubah acak kontinu dan sampel acaknya memuat nilai-nailai yang terletak dalam interval:

$$x_1 \le X_1 \le x + \Delta x, x_2 \le X_2 \le x_2 + \Delta x, ..., x_n \le X_n \le x_n + \Delta x$$

maka probabilitas nilai pengamatan yang terjadi dalam sampel adalah:

$$P(x_1 \le X_1 \le x_1 + \Delta x, x_2 \le X_2 \le x_2 + \Delta x, ..., x_n \le X_n \le x_n + \Delta x)$$

$$= f_x(x_1)\,\Delta x f_x(x_2) \Delta x\,...\,f_x(x_n)\,\Delta x$$

$$= f_x(x_1)\,f_x(x_2)\,...\,f_x(x_n)(\Delta x)^n$$

$$= L(\theta|X).(\Delta x)^n$$

Ide dasar dari taksiran maksimum likelihood adalah memaksimumkan fungsi likelihood $L(\theta|X)$. Jika fungsi likelihood $L(\theta|X)$ terdeferensialkan ke θ maka untuk memaksimumkan fungsi likelihood $L(\theta|X)$ adalah dengan menurunkan pertama terhadap θ dan nilainya dipandang sama dengan nol.

$$\frac{d\,L(\theta|X)}{d\,\theta} = 0.$$

Jika terdapat beberapa parameter dalam distribusi, misalnya $(\theta_1, \theta_2, \ldots, \theta_n)$ maka fungsi likelihoodnya adalah:

$$\frac{\partial\, L(\theta_1, \ldots, \theta_n)}{\theta_i} = 0, \qquad i = 1, 2, \ldots, n.$$

Karena $L(\theta|X)$ adalah probabilitas nilai pengamatan sampel, maka nilainya akan selalu non negatif, untuk semua kemungkinan nilai-nilai θ. Diperbolehkan menguji likelihood dengan menggunakan logaritma natural dari fungsi likelihood, yaitu :

$K(\theta|X) = \ln L(\theta|X)$.

Karena logaritma natural fungsinya adalah monoton naik untuk nilai θ, sehingga memaksimumkan nilai $K(\theta|X)$ identik dengan memaksimumkan nilai $L(\theta|X)$. Memaksimumkan nilai $K(\theta|X)$ ini lebih sering digunakan karena lebih sederhana dalam perhitungannya.

Contoh:

Pandang kasus taksiran untuk parameter λ yang mempunyai probabilitas eksponensial, jika diberikan sampel acak nilai pengamatan dengan ukuran n dari suatu populasi. Misalkan nilai-nilai yang terambil, n = 5 adalah 30,4; 7,8; 1,4; 13,1; 67,3. Maka fungsi likelhood untuk sampel adalah:

$$L(\lambda|X) = \left(\lambda e^{-30,4\lambda}\right)\ldots\left(\lambda e^{-67,3\lambda}\right)$$

$$= \lambda^5 e^{-120\lambda} \quad \text{,untuk } \lambda > 0$$

Logaritma natural dari $L(\lambda|X)$ adalah

$$K(\lambda|X) = \ln\left(\lambda^5 e^{-120\lambda}\right)$$

$$= 5 \ln \lambda - 120\,\lambda$$

Untuk mencari nilai maksimum, maka:

$$\frac{d\,K(\lambda|X)}{d\,\lambda} = \frac{5}{\hat{\lambda}} - 120$$

Dengan memandang nilainya sama dengan nol maka

$$\frac{5}{\hat{\lambda}} = 120$$

$\hat{\lambda} = 0.042$,Karena

$$\frac{d^2\,K(\lambda|X)}{d^2\,\lambda} = -\frac{5}{\hat{\lambda}^2} < 0,$$ $\hat{\lambda}$ adalah nilai maksimum untuk $K(\lambda|X)$.

Secara umum untuk nilai pengamatan $x_1, x_2, \ldots, x_n$ fungsi likelihoodnya:

$$L(\lambda|X) = \prod_{i=1}^{n}\left(\lambda e^{-\lambda x_i}\right)$$

$$= \lambda^n e^{-\lambda \Sigma x_i}$$

$$K(\lambda|X) = \ln L(\lambda|X) = n \ln\lambda - \lambda\,\Sigma x_i$$

$$\frac{d\,K(\lambda|X)}{d\lambda} = \frac{n}{\lambda} - \Sigma x_i$$

$$\frac{n}{\hat{\lambda}} = \Sigma x_i$$

$$\hat{\lambda} = \frac{n}{\Sigma x_i} = \frac{1}{\bar{x}}$$

2.6. Taksiran Kuadrat Rata-rata .

2.6.1 Prinsip Ortogonalitas.

Definisi 2.6.1.1:

Dua vektor **x** dan **y** dikatakan ortogonal jika nilai harapan dari perkalian kedua vektor tersebut sama dengan nol.

E [**x**.**y**] = 0

Jika diberikan n data pengamatan **y**(1), **y**(2), . . . , **y**(n) dan akan ditentukan n konstanta a_1, a_2, . . . , a_n sedemikian hingga jika ditaksir $\hat{\mathbf{x}}(n+1|\mathcal{Y}_n)$ sebagai kombinasi linier dari data.

$\hat{\mathbf{x}}(n+1|\mathcal{Y}_n) = a_1\mathbf{y}(1) + a_2\mathbf{y}(2) + \ldots + a_n\mathbf{y}(n)$.

dengan $\hat{\mathbf{x}}(n+1|\mathcal{Y}_n)$ disebut taksiran kuadrat rata-rata proses **x**(n+1) jika diberikan data pengamatan **y**(1), **y**(2), . . . , **y**(n). Harga momen ke dua perkiraan galat prosesnya adalah :

$P = \{(\mathbf{x}(n+1) - \hat{\mathbf{x}}(n+1|\mathcal{Y}_n))^2\}$

perkiraan galat proses yang dihasilkan

$\varepsilon(n+1,n) = \mathbf{x}(n+1) - \hat{\mathbf{x}}(n+1|\mathcal{Y}_n)$

adalah minimum.

Untuk menunjukkan bahwa perkiraan galat proses yang dihasilkan adalah minimum dengan menunjukkan momen ke dua perkiraan galat proses adalah minimum, yaitu dengan menggunakan teorema proyeksi.

Teorema proyeksi

Momen ke dua perkiraan galat proses P adalah minimum jika terdapat konstanta a_k sembarang, (k = 1, 2, 3,. . . , n) , sedemikian hingga perkiraan galat prosesnya ortogonal dengan vektor data pengamatannya.

$$E\{(\mathbf{x}(n+1) - \hat{\mathbf{x}}(n+1|\mathcal{Y}_n))\,\mathbf{y}(k)\}, \quad k = 1, 2, \ldots, n \tag{2.6.1.1}$$

Bukti:

Pembuktian dibagi dua bagian, pertama jika P minimum maka perkiraan galat prosesnya ortogonal dengan vektor data. Ke dua akan ditunjukkan jika a_k memenuhi persamaan (2.6.1.1), maka momen ke dua perkiraan galat proses, P , akan minimum.

1. Bukti pertama.

Karena momen ke dua perkiraan galat proses merupakan fungsi dari konstanta a_k, maka

$$\frac{\partial P}{\partial a_k} = E\{\, 2(\mathbf{x}(n+1) - \hat{\mathbf{x}}(n+1|\mathcal{Y}_n))\,(-\mathbf{y}(k))\,\}$$

$$\frac{\partial P}{\partial a_k} = -2\, E\{ (\mathbf{x}(n+1) - \hat{\mathbf{x}}(n+1|\mathcal{Y}_n))\, \mathbf{y}(k) \}$$

$$0 = E\{ (\mathbf{x}(n+1) - \hat{\mathbf{x}}(n+1|\mathcal{Y}_n))\, \mathbf{y}(k) \}$$

$$\frac{\partial^2 P}{\partial a_k} = -2\, E\{ -\mathbf{y}(k)\mathbf{y}(k) \}$$

$$= 2\, (\mathbf{y}(k))^2$$

$$> 0.$$

Terbukti bahwa jika P adalah minimum maka perkiraan galat prosesnya ortogonal dengan data.

$$E\{ (\mathbf{x}(n+1) - \hat{\mathbf{x}}(n+1|\mathcal{Y}_n))\, \mathbf{y}(k) \} = 0$$

2. Akan ditunjukkan bahwa jika a_k memenuhi persamaan (2.6.1.1) maka momen ke dua perkiraan galat proses P adalah minimum. Dari persamaan (2.6.1.1)

$$E\{(\mathbf{x}(n+1) - \hat{\mathbf{x}}(n+1|\mathcal{Y}_n))\, \mathbf{y}(k)\, (c_1\mathbf{y}(1) + c_2\mathbf{y}(2) + \ldots + c_n\mathbf{y}(n))\} = 0 \qquad (2.6.1.2)$$

Untuk sembarang c_k. Jika $\bar{\bar{a}}_k$ himpunan konstanta sembarang, maka

$$E\{(\mathbf{x}(n+1) - (\bar{\bar{a}}_1\mathbf{y}(1) + \bar{\bar{a}}_2\mathbf{y}(2) \ldots + \bar{\bar{a}}_n\mathbf{y}(n)))^2\}$$

$$= E\{((\mathbf{x}(n+1) - (a_1\mathbf{y}(1) + \ldots + a_n\mathbf{y}(n))) + ((a_1 - \bar{\bar{a}}_1)\mathbf{y}(1) + \ldots + (a_n - \bar{\bar{a}}_n)\mathbf{y}(n))^2\}$$

$$= E\{((\mathbf{x}(n+1) - (a_1\mathbf{y}(1) + \ldots + a_n\mathbf{y}(n)))^2\} + 2\, E\{((\mathbf{x}(n+1) - (a_1\mathbf{y}(1) + \ldots + a_n\mathbf{y}(n)))\, ((a_1 - \bar{\bar{a}}_1)\mathbf{y}(1) + \ldots + (a_n - \bar{\bar{a}}_n)\mathbf{y}(n)))\} + E\{((a_1 - \bar{\bar{a}}_1)\mathbf{y}(1) + \ldots + (a_n - \bar{\bar{a}}_n)\mathbf{y}(n))^2\}.$$

Sesuai dengan prinsip ortogonalitas, yaitu pada persamaan (2.6.1.2) maka nilai harapan yang ke dua sama dengan nol.

$$2\,E\{((\mathbf{x}(n+1) - (a_1\mathbf{y}(1)+ \ldots +a_n\mathbf{y}(n))) ((a_1-\bar{\bar{a}}_1)\mathbf{y}(1) + \ldots +(a_n-\bar{\bar{a}}_n)\mathbf{y}(n)))\} = 0$$

Sehingga persamaannya menjadi

$$E\{(\mathbf{x}(n+1) - (\bar{\bar{a}}_1\mathbf{y}(1) + \bar{\bar{a}}_2\mathbf{y}(2) \ldots +\bar{\bar{a}}_n\mathbf{y}(n)))^2\}$$

$$= E\{((\mathbf{x}(n+1)-(a_1\mathbf{y}(1)+ \ldots +a_n\mathbf{y}(n)))^2\}+E\{((a_1-\bar{\bar{a}}_1)\mathbf{y}(1) + \ldots + (a_n- \bar{\bar{a}}_n)\mathbf{y}(n))^2\}$$

nilainya akan minimum jika $a_K = \bar{\bar{a}}_k$. □

Bukti selesai.

Dari persamaan (2.6.1.2) mengakibatkan

$(\mathbf{x}(n+1) - \hat{\mathbf{x}}(n+1|\mathcal{Y}_n)) \perp (c_1\mathbf{y}(1) + c_2\mathbf{y}(2) + \ldots + c_n\mathbf{y}(n))$ untuk sembarang c_k. Sehingga $(\mathbf{x}(n+1) - \hat{\mathbf{x}}(n+1|\mathcal{Y}_n)) \perp \hat{\mathbf{x}}(n+1|\mathcal{Y}_n)$.

Taksiran kuadrat rata-rata yang digunakan dalam teori tapis kalman adalah taksiran kuadrat rata-rata dengan menggunakan proses inovasi. Pada bagian ini diberikan hubungan antara data pengamatan dan proses inovasi.

Terdapat ekivalensi secara linier antara proses inovasi $\alpha(k)$ dengan data pengamatan $\mathcal{Y}_n$, artinya setiap $\alpha(k)$ adalah fungsi linier dari elemen-elemen data $\mathcal{Y}_n$. Untuk membuktikan sifat ini digunakan metode Gram-Schmidt. Metode ini mengasumsikan bahwa $\alpha(k)$ hanya tergantung pada k (k = 1, 2, . . ., n) data pertama, $\mathbf{y}(1)$, $\mathbf{y}(2)$, . . ., $\mathbf{y}(n)$.

$\alpha(1) = \mathbf{y}(1)$ kemudian

$\alpha(2) = \mathbf{y}(2) + a_{1,1}\mathbf{y}(1)$

koefisien $a_{1,1}$ ini dipilih sedemikian hingga inovasi $\alpha(1)$ ortogonal pada $\alpha(2)$

$E[\alpha(2)\alpha^T(1)] = 0$

$E[(\mathbf{y}(2) + a_{1,1}\mathbf{y}(1))\,\alpha^T(1)] = 0$

$E[\mathbf{y}(2)\,\alpha(1)] + a_{1,1}\,E[\mathbf{y}(1)\,\alpha^T(1)] = 0$

$$a_{1,1} = -\frac{E[y(1)\alpha^T(1)]}{E[y(2)\alpha^T(1)]}$$

$$a_{1,1} = -\frac{E[y(1)y^T(1)]}{E[y(2)y^T(1)]}$$

kemudian

$\alpha(3) = \mathbf{y}(3) + a_{2,1}\mathbf{y}(2) + a_{2,2}\mathbf{y}(1)$

dengan koefisien $a_{2,1}$ dan $a_{2,2}$ dipilih sedemikian hingga inovasi $\alpha(3)$ ortogonal dengan $\alpha(1)$ dan $\alpha(2)$ dan seterusnya. Secara umum hubungan antara $\alpha(1)$, $\alpha(2)$, . . . , $\alpha(n)$ dengan data pengamatan $\mathbf{y}(1)$, $\mathbf{y}(2)$, . . . , $\mathbf{y}(n)$ dapat ditulis :

$$\begin{bmatrix} \alpha(1) \\ \alpha(2) \\ \vdots \\ \alpha(n) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_{1,1} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,n-1} & a_{n-1,n-2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(n) \end{bmatrix}$$

sebagai catatan bahwa $a_{k,o} = 1$, untuk semua k.

dari ekivalensi linier antara proses-proses inovasi $\alpha(n)$ dan pengamatan **y**(n), maka taksiran kuadrat rata-rata proses **x**(n+1) jika diberikan data **y**(1), **y**(2), . . . , **y**(n) dapat dinyatakan sebagai taksiran kuadrat rata-rata proses **x**(n+1) jika diberikan inovasi-inovasi $\alpha(1), \alpha(2), \ldots, \alpha(n)$.

$$\hat{\mathbf{x}}(n|\mathcal{Y}_n) = b_1\alpha(1) + b_2\alpha(2) + \ldots + b_n\alpha(n)$$

koefisien b_k (k = 1, 2, . . ., n) ini dipilih sedemikian sehingga

$$(\mathbf{x}(n) - \hat{\mathbf{x}}(n|\mathcal{Y}_n)) \perp \alpha(n) \;, \quad 1 \leq k \leq n.$$