# Konsistensi dan Asimtotik Normalitas Model Multivariate Adaptive Regression Spline (Mars) Respon Biner

## *Consistency and Asymptotic Normality of Maximum Likelihood Estimator in MARS Binary Response Model*

Bambang Widjanarko Otok
*Jurusan Statistika FMIPA Institut Teknologi Sepuluh Nopember*

**ABSTRACT**

A good estimator has to fulfill some propertries such as unbiased, efficient and consistent. This research aims to study consistency and asymptotic normality of maximum likelihood estimator in MARS binary response model of Friedman, for predicting the continue response variable, $\mathbf{y}$, at $p$ predictor variables, $\mathbf{x} = (x_{1i}, \ldots, x_{pi})$ with linear combination of spline truncated. Let response variable, $y \in (0,1)$, be MARS binary response. Using penalized log-likelihood method to look for MARS binary response estimator, obtained the following result $\nabla_{\hat{\beta}}(L)(\hat{\beta}) = \mathbf{B}^T[\mathbf{y} - \sigma_L(\mathbf{B}\hat{\beta})]$ with $\mathbf{B} = [1, \{(x_{v,(k,m)} - t_{km})\}_1^K]$. Asymptotic properties for consistency is absolute minimum $\hat{\beta}$ from $\varepsilon_{y,\sigma_L}$ (from $\ln L$ absolute maximum), this is not only unique but also convergence in probability to $\hat{\beta}$, $n \to \infty$, while asymptotic normality is $\left\|\hat{\beta} - \beta\right\| = O(p/n)(1/(a_{pn}^{\varsigma}(\mathbf{B}, \hat{\beta}))^2)$ and $s_i^2 = \sigma_L^{'}(B_i^T \hat{\beta}) \leq \frac{1}{4}$.

Keywords : Asymptotic normality, maximum likelihood estimator, MARS binary response

### PENDAHULUAN

Analisis regresi memperlihatkan hubungan dan pengaruh variabel prediktor terhadap variabel respon. Misalnya y adalah variabel respon dan $\mathbf{x}_i = (x_{1i}, \ldots, x_{pi})$ adalah variabel prediktor, untuk n buah pengamatan, secara umum hubungan antara y dan $\mathbf{x}_i = (x_{1i}, \ldots, x_{pi})$ dapat ditulis sebagai berikut :

$$y_i = f(\mathbf{x}_i) + \varepsilon_i \qquad , i = 1, \ldots, n$$

dengan $\varepsilon_i$ adalah kesalahan random dan $f(\mathbf{x}_i)$ merupakan kurva regresi.

Untuk mengestimasi kurva atau parameter regresi, ada dua pendekatan yaitu parametrik dan nonparametrik. Pendekatan parametrik digunakan jika ada informasi sebelumnya tentang bentuk kurva, yang diperoleh berdasarkan teori atau pengalaman masa lalu, sehingga dapat dikatakan bahwa mengestimasi kurva ekuivalen dengan mengestimasi parameter, dimana hasil estimasi mengikuti model tertentu. Sedang pendekatan nonparametrik digunakan jika tidak ada informasi tentang bentuk kurva $f(\mathbf{x}_i)$, tidak tergantung pada asumsi bentuk kurva tertentu, sehingga memberikan fleksibilitas yang lebih besar (Eubank 1988, Emond & Steven 1997, Hardle 1990), jika kurva regresi merupakan model parametrik maka disebut sebagai regresi parametrik dan apabila model yang diasumsikan ini benar, maka pendugaan parametrik sangat efisien, tetapi jika tidak, menyebabkan interpretasi data yang menyesatkan.

Ada beberapa teknik estimasi dalam regresi nonparametrik antara lain pendekatan histogram, estimator spline, estimator kernel, estimator deret orthogonal, analisis wavelet dan lain-lain. Pendekatan estimator spline ada bermacam-macam antara lain spline original, spline type M, spline relaxed, spline terbobot dan lain-lain. Pendekatan spline mempunyai suatu basis fungsi. Basis fungsi yang biasa dipakai antara lain spline truncated dan B-spline. (Friedman 1990). Model MARS Friedman (1990), bertujuan untuk prediksi, yang merupakan kombinasi *recursive partition* regression dengan spline truncated yang telah dimodifikasi. Holmes & Denisson (2002), secara khusus membahas penerapan statistika Bayesian pada pembentukan model MARS klasifikasi.

Green (1995) dalam Holmes & Denisson (2002), mengembangkan penggunaan MCMC dalam statistika inferensia Bayesian MARS. Selanjutnya Xiang & Meullenet (2002), membandingkan regresi logistik dan MARS pada efek aktivitas air, pH dan potassium.

Tujuan paper ini mencari estimator kurva dari model MARS respon biner, dan selanjutnya mengkaji sifat-sifat asimtotik estimatornya.

## HASIL DAN PEMBAHASAN

**MARS respon biner**

Pertimbangkan $y$ suatu variabel random berdistribusi Bernoulli, $y \sim Ber(1, \pi(\mathbf{x}))$ dengan y $\in$ (0, 1) dan $\mathbf{x} \in \Re^p$, maka $P(Y_i = 1) = \pi(\mathbf{x})$ dan $P(Y_i = 0) = 1 - \pi(\mathbf{x})$. $\mathbf{x} \in \Re^p$ adalah vektor dari p variabel prediktor dan $P(Y = 1 | \mathbf{x}) = \pi(\mathbf{x})$. Misalkan model ditulis sebagai,

$$y = \pi(\mathbf{x}) + r \text{ maka } r = y - \pi(\mathbf{x})$$

sehingga,

$$\begin{aligned} E(r) &= E(y - \pi(\mathbf{x})) \\ &= E(y) - E(\pi(\mathbf{x})) = \pi(\mathbf{x}) - \pi(\mathbf{x}) = 0 \end{aligned} \quad (1)$$

$$Var(r) = Var(y - \pi(\mathbf{x})) = \pi(\mathbf{x})(1 - \pi(\mathbf{x})) \quad (2)$$

Lemma 1: Jika hubungan dengan model logistik, $\sigma_L : R \to (0,1)$, $\sigma_L = \pi(x) = [e^z / (1 + e^z)]$ maka inver dari dari $\sigma_L$ dapat dikatakan sebagai transformasi logit, $\text{logit } \pi(x) = \ln[\pi(x)/(1 - \pi(x)] = z$

Bukti:

$$\sigma_L = \pi(x) = \frac{e^z}{1+e^z} = \frac{e^z / e^z}{(1+e^z)/e^z} = \frac{1}{1+e^{-z}}$$

Turunan pertama dari $\sigma_L$ yang dinotasikan $\sigma'_L$ adalah:

$$\begin{aligned} \sigma'_L(z) &= \frac{\partial}{\partial z}\left(\frac{e^z}{1+e^z}\right) = \frac{e^z(1+e^z) - (e^z)^2}{(1+e^z)^2} \\ &= \frac{e^z}{(1+e^z)^2} > 0, \quad \forall z \in \Re \end{aligned} \quad (3)$$

Invers dari $\sigma_L$ yaitu:

$$\begin{aligned} \pi(x) = \frac{e^z}{1+e^z} &\Leftrightarrow \pi(x)[1+e^z] = e^z \\ &\Leftrightarrow \pi(x) + \pi(x)e^z = e^z \\ &\Leftrightarrow \pi(x) = e^z - \pi(x)e^z \\ &\Leftrightarrow \pi(x) = [1 - \pi(x)]e^z \\ &\Leftrightarrow \frac{\pi(x)}{1-\pi(x)} = e^z \Leftrightarrow \ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = z \end{aligned}$$

Jadi invers dari $\sigma_L$ dapat dikatakan sebagai transformasi logit, yaitu:

$$\text{logit } \pi(x) = \ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = z$$

Selanjutnya jika Lemma 1 terpenuhi dan $z = \hat{f}(x)$, yaitu

$$z = \hat{f}(x) = \beta_0 + \sum_{m=1}^{M} \beta_m \prod_{k=1}^{K_m} [s_{km}.(x_{v(k,m)} - t_{km})]$$

maka dapat ditulis dalam model,

$$\text{logit } \pi(x) = \ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \sum_{m=1}^{M} \beta_m \prod_{k=1}^{K_m} [s_{km}.(x_{v(k,m)} - t_{km})]$$

dan dalam bentuk matriks,

$$\text{logit } \pi(x) = \mathbf{B}\beta \quad (4)$$

dengan,

$$\beta = (\beta_0, \beta_1, \ldots, \beta_M)^T$$

$$\mathbf{B} = \begin{pmatrix} 1 & \prod_{k=1}^{K_1} s_{1m}(x_{1(1,m)} - t_{1m}) & \cdots & \prod_{k=1}^{K_M} s_{Mm}(x_{1(M,m)} - t_{Mm}) \\ 1 & \prod_{k=1}^{K_1} s_{1m}(x_{2(1,m)} - t_{1m}) & \cdots & \prod_{k=1}^{K_M} s_{Mm}(x_{2(M,m)} - t_{Mm}) \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ 1 & \prod_{k=1}^{K_1} s_{1m}(x_{n(1,m)} - t_{1m}) & \cdots & \prod_{k=1}^{K_M} s_{Mm}(x_{n(M,m)} - t_{Mm}) \end{pmatrix}$$

Persamaan (4) dikatakan sebagai model MARS respon biner, dan dalam bentuk fungsi probabilitas dapat dinyatakan sebagai,

$$\pi(x) = \sigma_L(\mathbf{B}^T\beta)$$

Untuk mencari estimasi $\beta$ digunakan metode maksimum likelihood.

**Estimator MARS respon biner**

Misalkan variabel random Y, $y \sim Ber(1, \sigma_L(\mathbf{B}_i\beta))$ dan dinyatakan dalam model, $y = \sigma_L(\mathbf{B}\beta) + r$ maka $r = y - \sigma_L(\mathbf{B}\beta)$. Jika fungsi $\sigma_L$ diterapkan pada vektor $z = (z_1, \cdots, z_n)^T \in R^n$ mempunyai interpretasi sebagai $\sigma_L(z) = (\sigma_L(z_1), \cdots, \sigma_L(z_n))^T$. Analog Persamaan (1) dan (2) diperoleh,

$E(r) = 0$, dan

$$s_i^2 = Var(r_i) = Var(y - \sigma_L(\mathbf{B}_i^T\beta)) = \sigma_L(\mathbf{B}_i^T\beta)(1-\sigma_L(\mathbf{B}_i^T\beta))$$

$$= \frac{e^{\mathbf{B}_i^T\beta}}{1+e^{\mathbf{B}_i^T\beta}}\left[1-\frac{e^{\mathbf{B}_i^T\beta}}{1+e^{\mathbf{B}_i^T\beta}}\right] = \frac{e^{\mathbf{B}_i^T\beta}}{1+e^{\mathbf{B}_i^T\beta}}\left[\frac{1+e^{\mathbf{B}_i^T\beta}-e^{\mathbf{B}_i^T\beta}}{1+e^{\mathbf{B}_i^T\beta}}\right]$$

$$= \frac{e^{\mathbf{B}_i^T\beta}}{(1+e^{\mathbf{B}_i^T\beta})^2} = \sigma_L^{'}(\mathbf{B}_i^T\beta)$$

Sehingga fungsi likelihood dapat dinyatakan sebagai

$$L(\beta) = \prod_{i=1}^{n} P(\mathbf{y} = y_i)$$

$$= \prod_{i=1}^{n}\binom{1}{y_i}[\sigma_L(\mathbf{B}_i^T\beta)]^{y_i}[(1-\sigma_L(\mathbf{B}_i^T\beta)]^{1-y_i}$$

$$= \prod_{i=1}^{n}\frac{1}{y_i!(1-y_i)!}[\sigma_L(\mathbf{B}_i^T\beta)]^{y_i}[(1-\sigma_L(\mathbf{B}_i^T\beta)]^{1-y_i}$$

$$= \prod_{i=1}^{n}\left(\frac{\sigma_L(\mathbf{B}_i^T\beta)}{(1-\sigma_L(\mathbf{B}_i^T\beta)}\right)^{y_i}(1-\sigma_L(\mathbf{B}_i^T\beta)$$

$$= \prod_{i=1}^{n}\left(\frac{\dfrac{e^{\mathbf{B}_i^T\beta}}{1+e^{\mathbf{B}_i^T\beta}}}{\dfrac{1+e^{\mathbf{B}_i^T\beta}-e^{\mathbf{B}_i^T\beta}}{1+e^{\mathbf{B}_i^T\beta}}}\right)^{y_i}\left(\frac{1+e^{\mathbf{B}_i^T\beta}-e^{\mathbf{B}_i^T\beta}}{1+e^{\mathbf{B}_i^T\beta}}\right)$$

$$= \prod_{i=1}^{n}(e^{\mathbf{B}_i^T\beta})^{y_i}\left(\frac{1}{1+e^{\mathbf{B}_i^T\beta}}\right)$$

$$= \prod_{i=1}^{n}\frac{e^{\mathbf{B}_i^T\beta y_i}}{1+e^{\mathbf{B}_i^T\beta}}$$

dan fungsi log-likelihoodnya adalah

$$\ln L(\beta) = \ln\left(\prod_{i=1}^{n}\frac{e^{\mathbf{B}_i^T\beta y_i}}{1+e^{\mathbf{B}_i^T\beta}}\right)$$

$$= \sum_{i=1}^{n}\ln(\exp(\mathbf{B}_i^T\beta y_i)) - \sum_{i=1}^{n}\ln(1+\exp(\mathbf{B}_i^T\beta))$$

$$= \sum_{i=1}^{n}\mathbf{B}_i^T\beta y_i - \sum_{i=1}^{n}\ln(1+\exp(\mathbf{B}_i^T\beta))$$

$$= \mathbf{y}^T\mathbf{B}\beta - \sum_{i=1}^{n}\ln(1+\exp(\mathbf{B}_i^T\beta)) \tag{5}$$

Untuk menunjukkan fungsi maksimum, diturunkan fungsi log-likelihood terhadap $\beta$, yaitu sebagai berikut:

$$\frac{\partial}{\partial\beta}\ln L(\beta) = \frac{\partial}{\partial\tilde{a}}\left(\mathbf{y}^T\mathbf{B}\beta - \sum_{i=1}^{n}\ln(1+\exp(\mathbf{B}_i^T\beta))\right)$$

$$= \mathbf{B}^T\mathbf{y} - \sum_{i=1}^{n}\frac{\exp(\mathbf{B}_i^T\beta)}{(1+\exp(\mathbf{B}_i^T\beta)}\mathbf{B}_i$$

$$= \mathbf{B}^T\mathbf{y} - \mathbf{B}^T\sigma_L(\mathbf{B}\beta)$$

$$= \mathbf{B}^T[\mathbf{y} - \sigma_L(\mathbf{B}\beta)]$$

Jika turunan pertama disamadenganakan nol, maka $\hat{\beta}$ adalah estimator maksimum likelihood dari $\beta$, yaitu:

$$\mathbf{B}^T[\mathbf{y} - \sigma_L(\mathbf{B}\hat{\beta})] = 0 \tag{6}$$

Untuk membuktikan $\hat{\beta}$ adalah estimator yang memaksimumkan Persamaan (4) dicari turunan kedua dari fungsi log-likelihood, yaitu:

$$\frac{\partial^2}{\partial\beta_s\partial\beta_r}\ln L(\beta) = \frac{\partial}{\partial\beta_s}\left(\mathbf{B}^T[\mathbf{y} - \sigma_L(\mathbf{B}\beta)]\right)$$

$$= \frac{\partial}{\partial\beta_s}\left(\sum_{i=1}^{n}\mathbf{B}_{ir}\mathbf{y}_i - \sum_{i=1}^{n}\mathbf{B}_{ir}\sigma_L(\mathbf{B}_i^T\beta)]\right)$$

$$= -\sum_{i=1}^{n}\mathbf{B}_{ir}\frac{\partial}{\partial\beta_s}\sigma_L(\mathbf{B}_i^T\beta)$$

$$= -\sum_{i=1}^{n}\mathbf{B}_{ir}\mathbf{B}_{is}\sigma_L^{'}(\mathbf{B}_i^T\beta)$$

$$= -\mathbf{B}^T D(\beta)\mathbf{B}$$

dan diperoleh matriks Hessian, $H_{\ln L}(\beta) \in \Re^p$, yaitu:

$$H_{\ln L}(\beta) = -\mathbf{B}^T D(\beta)\mathbf{B}$$

dimana $D(\beta) = d_{ij} \in \Re^n$ matrik diagonal yang didefinisikan sebagai,

$$d_{ij} = \begin{cases}\sigma_L^{'}(\mathbf{B}_i^T\beta) & \text{, jika } i = j\\ 0 & \text{, jika } i \neq j\end{cases} \tag{7}$$

dan $H_{\ln L}(\beta)$ semi definit negatif untuk setiap $\beta \in \Re^p$. Sehingga diperoleh:

$$\mathbf{u}^T H_{\ln L}(\beta)\mathbf{u} = -\mathbf{u}^T\mathbf{B}^T D(\beta)\mathbf{B}\mathbf{u} = -\sum_{i=1}^{n}(\mathbf{B}_i^T\mathbf{u})^2\sigma_L^{'}(\mathbf{B}_i^T\beta) \tag{8}$$

Jadi turunan pertama selalu positip, sedangkan dari Persamaan (8) bahwa $\mathbf{u}^T H_{\ln L}(\beta)\mathbf{u} \le 0$ untuk semua $\mathbf{u} \in \Re^p$ dan $\beta \in \Re^p$. Jadi dapat dikatakan bahwa $\hat{\beta}$ adalah estimator yang memaksimumkan.

**Sifat asimtotik estimator MARS respon biner**

Sifat konsistensi pada $\hat{\beta}$ dapat ditunjukkan bahwa $\beta$ konvergen pada nilai $\hat{\beta}$ untuk n observasi pada suatu model. Pandang fungsi error,

$$\varepsilon_{y,\sigma}(\beta) = 1_n^T H_\sigma(\mathbf{B}\beta) - \mathbf{y}^T\mathbf{B}\beta \qquad (9)$$

dimana, $1_n = (1,\cdots,1)^T \in \Re^n$ dan $H_\sigma$ adalah matrik Hat dari $\sigma : \Re \to \Re^n$ . Turunan pertama dari fungsi error,

$$\frac{\partial}{\partial\beta}(\varepsilon_{y,\sigma}(\beta)) = \mathbf{B}^T\sigma(\mathbf{B}\beta) - \mathbf{B}^T\mathbf{y} = \mathbf{B}^T(\sigma(\mathbf{B}\beta) - \mathbf{y})$$

hal ini sama dengan,

$$\frac{\partial}{\partial\beta}\ln L(\beta) = -\frac{\partial}{\partial\beta}(\varepsilon_{y,\sigma_L}(\beta))$$

yang berarti estimator $\hat{\beta}$ yang memaksimumkan fungsi log-likelihood $\ln L$ meminimumkan fungsi error $\varepsilon_{y,\sigma_L}(\beta)$.

Pandang $H_\sigma(z) = \ln(1+e^z)$, dan dari Persamaan (5) dan (9) menunjukkan bahwa $\varepsilon_{y,\sigma_L}(\beta) = -\ln L(\beta)$.

Selanjutnya untuk menyelidiki kekonsistensian, diperlukan asumsi dan teorema berikut:

Asumsi 1. $\sum_{j=1}^{p}\sum_{i=1}^{n} B_{ij}^2 \le Knp$ untuk beberapa konstanta $K > 0$.

Asumsi 2. Jika $\lambda_*$ sebagai eigen value terkecil dari matrik simetris $\mathbf{B}^T\mathbf{B} \in \Re^p$ , ada konstanta k positif $k > 0$ sedemikian hingga $\lambda_* > kn$ untuk semua n.

Teorema 1. Jika $\beta \in \Re^p$ , dan $\mathbf{y} = \sigma_L(\mathbf{B}\hat{\beta}) + \mathbf{r}$ vektor random, dengan $\mathbf{r} = (r_1,\cdots,r_n)^T$ dan saling independen pada setiap $r_i$ sedemikian hingga $E(r_i) = 0$ dan $E(r_i^2) = s_i^2 > 0$ untuk semua $i \in \{1,\cdots,n\}$. Misalkan $Ber(\hat{\beta},\varsigma) \subset \Re^p$ adalah bola terbuka dengan radius $\varsigma$ terpusat pada $\hat{\beta}$, dan $a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta}) = \inf_{\zeta \in \mathbf{B}Ber(\hat{\beta},\varsigma)} \min_{i=1,\cdots,n} \sigma_L'(\zeta) > 0$. Selanjutnya, diasumsikan bahwa $p = p(n)$ tergantung pada n sedemikian hingga,

$$\sqrt{\frac{p}{n}}\,\frac{1}{a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})} \xrightarrow{n\to\infty} 0 \qquad (10)$$

untuk semua $\varsigma > 0$, maka minimum absolut $\hat{\beta}$ dari $\varepsilon_{y,\sigma_L}$ konvergen dalam probabilitas ke-1, $n \to \infty$ dan akar unik dari $\frac{\partial}{\partial\beta}(\varepsilon_{y,\sigma_L}(\beta)) = -\frac{\partial}{\partial\beta}\ln L(\beta)$ adalah konvergen dalam probabilitas ke nilai $\hat{\beta}$.

Bukti:

Pertimbangkan fungsi $G(\beta) = -\frac{\partial}{\partial\beta}(\varepsilon_{y,\sigma_L}(\beta))$, akan ditunjukkan keberadaan bola $Ber(\hat{\beta},\varsigma) \subset \Re^p$ , yang mana konvergen dalam probabilitas ke 1 untuk $n \to \infty$, yang terdiri dari akar $\hat{\beta}$ dari $G$ untuk perubahan kecil $\varsigma > 0$.

Misalkan $G(\beta)_j$ adalah komponen ke-j dari vektor $G(\beta)$. Maka,

$$\begin{aligned} G(\beta)_j &= (\mathbf{B}^T(\mathbf{y} - \sigma(\mathbf{B}\beta)))_j \\ &= \sum_{i=1}^{n} B_{ij}(y_i - \sigma_L(B_i^T\beta)) \\ &= \sum_{i=1}^{n} B_{ij}(y_i - \sigma_L(B_i^T\hat{\beta} + B_i^T\gamma)) \end{aligned}$$

dimana $\gamma = \beta - \hat{\beta}$ dan didefinisikan $\boldsymbol{\eta} = (\eta_1,\cdots,\eta_n)^T = \mathbf{B}\hat{\beta}$, dengan mempertimbangkan Y variabel random, $y \sim Ber(1,\sigma_L(\mathbf{B}_i\beta))\ (i = 1,\cdots,n)$, maka

$$G(\beta)_j = \sum_{i=1}^{n} B_{ij}(\sigma_L(\eta_i)) + r_i - (\sigma_L(\eta_i + B_i^T\gamma))$$

sedemikian hingga sebagai hasil teorema, bahwa keberadaan $\zeta_i = \eta_i + \alpha_i B_i^T\gamma$ dengan $\|\alpha_i\| \in (0,1)$

$$G(\beta)_j = \sum_{i=1}^{n} B_{ij}(r_i - \sigma_L'(\eta_i)B_i^T\gamma)$$

Selanjutnya, akan ditunjukkan bahwa $\gamma^T G(\beta) \le 0$ untuk semua $\gamma = \beta - \hat{\beta}$ dengan $\|\gamma\| = \varsigma$.

Misalkan $Ber^* = Ber(\hat{\beta},\varsigma)$ adalah bola terbuka dalam $\Re^p$ dengan pusat pada $\hat{\beta}$ dan radius $\varsigma > 0$. Diasumsikan bahwa $G : \overline{Ber} \subset \Re^p \to \Re^p$ adalah kontinu dan $(\beta - \hat{\beta})^T G(\beta) \le 0$ untuk semua $\beta \in \partial Ber$ , maka $G$ mempunyai akar dari $\overline{Ber}$ .

Hal ini dapat ditunjukkan dengan mempertimbangkan $Ber_o = Ber(0,\varsigma)$ dan didefinisikan $G_o : \overline{Ber_o} \subset \Re^n$ yaitu $G_o(\gamma) = \gamma + G(\gamma + \hat{\beta})$ . Jika G kontinu maka $G_o$

juga kontinu. Misalkan $\gamma \in \partial Ber_o$ yaitu $\|\gamma\|^2 = \varsigma^2$. Kemudian $\gamma + \hat{\beta} \in \partial Ber$ dan ini untuk setiap $\lambda > 1$,

$$\begin{aligned}\gamma^T(\lambda\gamma - G_o(\gamma)) &= \gamma^T((\lambda-1)\gamma - G(\gamma+\hat{\beta}))\\ &= (\lambda-1)\gamma^T\gamma - \gamma^T G(\gamma+\hat{\beta})\\ &= (\lambda-1)\|\gamma\|^2 - \gamma^T G(\gamma+\hat{\beta}) > 0\end{aligned} \tag{11}$$

Sekarang ditunjukkan bahwa $G_o$ mempunyai titik tetap $\hat{\gamma} \in \overline{Ber_o}$ yaitu $G_o(\hat{\gamma}) = \hat{\gamma}$. Hasil ini mempunyai arti bahwa $G(\hat{\gamma} + \hat{\beta}) = 0$ dan $\beta = \hat{\gamma} + \hat{\beta} \in \overline{Ber}$ adalah akar dari G.

Diasumsikan bahwa $G_o$ mempunyai titik tidak tetap dalam $\overline{Ber_o}$. Letak $\hat{G}(\gamma) = \dfrac{\varsigma(G_o(\gamma) - \gamma)}{\|G_o(\gamma) - \gamma\|}$ adalah kontinu dalam $\overline{Ber_o}$, dan $\|\hat{G}(\gamma)\| = \varsigma$ untuk semua $\gamma \in \overline{Ber_o}$. $\hat{G}$ mempunyai titik tetap $\gamma^*$ dalam $\overline{Ber_o}$ dan $\|\gamma^*\| = \|\hat{G}(\gamma^*)\| = \varsigma$. Jadi

$$\gamma^* = \hat{G}(\gamma^*) = \varsigma \frac{(G_o(\gamma^*) - \gamma^*)}{\|G_o(\gamma^*) - \gamma^*\|}$$

dengan,

$$G_o(\gamma^*) = \frac{\gamma^*}{\varsigma}\|G_o(\gamma^*) - \gamma^*\| + \gamma^* = (1 + \tfrac{1}{\varsigma}\|G_o(\gamma^*) - \gamma^*\|)\gamma^* = \lambda^*\gamma^*$$

.

Hasil ini, $\gamma^{*T}(\lambda^*\gamma^* - G_o(\gamma^*)) = 0$ kontradiksi dengan Persamaan (11).

Sehingga $\gamma^T G(\beta)$ dapat diekspresikan sebagai,

$$\begin{aligned}\gamma^T G(\beta) &= \sum_{j=1}^{p} \gamma_j G(\beta)_j\\ &= \sum_{j=1}^{p} \gamma_j \left( \sum_{i=1}^{n} B_{ij}(r_i - \sigma_L'(\xi_i) B_i^T \gamma) \right)\\ &= \sum_{i=1}^{n} r_i \left( \sum_{j=1}^{p} \gamma_j B_{ij} \right) - \sum_{i=1}^{n} \left( \sum_{j=1}^{p} \gamma_j B_{ij} \right) \sigma_L'(\xi_i) B_i^T \gamma\\ &= \sum_{i=1}^{n} r_i B_i^T \gamma - \sum_{i=1}^{n} \left( B_i^T \gamma \right)^2 \sigma_L'(\xi_i)\\ &= A_1 - A_2\end{aligned}$$

Dengan pertidaksamaan Cauchy-Schwarz, diperoleh batas atas $|A_1|$ sebagai berikut:

$$\begin{aligned}|A_1| &= \left| \sum_{i=1}^{n} r_i B_i^T \gamma \right|\\ &= \left| \left( \sum_{i=1}^{n} r_i B_i^T \right) \gamma \right| \le \left\| \sum_{i=1}^{n} r_i B_i^T \right\| \|\gamma\| \le \left\| \sum_{i=1}^{n} r_i B_i^T \right\| \varsigma\end{aligned}$$

(12)

dan moment kedua dari $\left\| \sum_{i=1}^{n} r_i B_i^T \right\|$ diperoleh,

$$\begin{aligned}E\left( \left\| \sum_{i=1}^{n} r_i B_i \right\|^2 \right) &= E\left( \sum_{j=1}^{p} \left( \sum_{i=1}^{n} r_i B_{ij} \right)^2 \right)\\ &= E\left( \sum_{j=1}^{p} \left( \sum_{i=1}^{n} r_i^2 B_{ij}^2 + \sum_{i=1}^{n} \sum_{\substack{k=1\\ k \ne i}}^{n} r_i r_k B_{ij} B_{kj} \right) \right)\\ &= \sum_{j=1}^{p} \left( \sum_{i=1}^{n} E(r_i^2) B_{ij}^2 + \sum_{i=1}^{n} \sum_{\substack{k=1\\ k \ne i}}^{n} E(r_i r_k) B_{ij} B_{kj} \right)\end{aligned}$$

dimana, variabel random $r_i$ adalah saling independen dan ekspektasinya sama dengan nol, $E(r_i)E(r_k) = 0$ dan $s_i^2 = \sigma_L'(\mathbf{B}_i^T \hat{\beta}) \le \frac{1}{4}$ untuk semua $i$

$$E\left( \left\| \sum_{i=1}^{n} r_i B_i \right\|^2 \right) = \sum_{i=1}^{n} E(r_i^2) \sum_{j=1}^{p} B_{ij}^2 = \sum_{i=1}^{n} s_i^2 \sum_{j=1}^{p} B_{ij}^2$$

dengan menggunakan Asumsi (1), diperoleh

$$E\left( \left\| \sum_{i=1}^{n} r_i B_i \right\|^2 \right) \le \frac{1}{4} \sum_{i=1}^{n} \sum_{j=1}^{p} B_{ij}^2 \le \frac{Knp}{4} \tag{13}$$

untuk beberapa konstanta positip $K > 0$. Dari Persamaan (12) dan (13) bahwa $E(A_1) \le \frac{1}{4}\varsigma^2 Knp$. Jika digunakan Ketidaksamaan Tchbychev diperoleh:

$$P(|A_1| \ge t) \le t^{-2} E(A_1^2) \le t^{-2} \tfrac{1}{4}\varsigma^2 Knp \qquad \forall t > 0$$

dan ekuivalen dengan,

$$\begin{aligned}P(|A_1| < t) &\ge 1 - t^{-2} \tfrac{1}{4}\varsigma^2 Knp && \forall t > 0\\ &\ge 1 - \varepsilon && \forall t > 0\end{aligned}$$

$$P\left( |A_1| < \frac{\varsigma\sqrt{Knp}}{2\sqrt{\varepsilon}} \right) \ge 1 - \varepsilon \qquad \forall t > 0$$

Misalkan $K^* = \sqrt{K}/2$ maka diperoleh

$$P\left( |A_1| < \varsigma K^* \sqrt{\frac{np}{\varepsilon}} \right) \ge P\left( |A_1| < \frac{\varsigma K^* \sqrt{np}}{\sqrt{\varepsilon}} \right)$$

Jika $n_\varepsilon = n/\varepsilon$ untuk setiap $\varepsilon > 0$ maka

$$P(A_1 \le \varsigma K^* \sqrt{n_\varepsilon p}) \ge 1 - \varepsilon$$

untuk semua $\varepsilon > 0$, ada sejumlah $n_\varepsilon \in N$ sedemikian hingga $n \ge n_\varepsilon$ dan diberikan $K^* > 0$, maka diperoleh

$$P(A_1 \le \varsigma K^* \sqrt{np}) \ge 1-\varepsilon \tag{14}$$

Sedangkan untuk $A_2$, misal

$Z = \{\xi_i \in \Re^n \mid \xi_i = B_i^T \hat{\beta} + \alpha_i B_i^T \gamma, \alpha_i \in (0,1)\}$.

Pandang, setiap vektor $\xi \in Z$ diperoleh,

$\sigma_L^{'}(\xi_i) \ge a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta}) = \inf_{\xi \in \mathbf{B}Ber(\hat{\beta},\varsigma)} \min_{i=1,\cdots,n} \sigma_L^{'}(\xi_i) \quad \forall_i \in \{1,\cdots,n\}$

Diasumsikan bahwa $\xi^* \in Z$ sedemikian hingga untuk $i^* \in \{1,\cdots,n\}$ diperoleh, $\sigma_L^{'}(\xi_{i^*}^*) < \inf_{\xi \in \mathbf{B}Ber(\hat{\beta},\varsigma)} \min_{i=1,\cdots,n} \sigma_L^{'}(\xi_i)$, dengan $\xi_{i^*}^* = B_i^T \hat{\beta} + \alpha_i B_i^T \gamma$. Misalkan ditentukan $\zeta^* = \mathbf{B}\hat{\beta} + \alpha_{i^*}\mathbf{B}\gamma = \mathbf{B}(\hat{\beta} + \alpha_{i^*}\gamma)$ dan $\|\alpha_{i^*}\gamma\| < \|\gamma\| = \varsigma$ maka diperoleh $\zeta^* = \mathbf{B}Ber(\hat{\beta},\varsigma)$.

Hasil

$\min_{i=1,\cdots,n} \sigma_L^{'}(\zeta_i^*) \le \sigma_L^{'}(\zeta_i^*) = \sigma_L^{'}(B_{i^*}^T\hat{\beta} + \alpha_{i^*}B_{i^*}^T\gamma) = \sigma_L^{'}(\xi_i^*)$

adalah kontradiksi.

Sekarang pandang,

$$A_2 = \sum_{i=1}^{n}\left(B_i^T\gamma\right)^2 \sigma_L^{'}(\xi_i) \ge a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})\sum_{i=1}^{n}\left(B_i^T\gamma\right)^2 = a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})\|\mathbf{B}\gamma\|^2 \tag{15}$$

dimana $\|\mathbf{B}\gamma\|^2 = (\mathbf{B}\gamma)^T(\mathbf{B}\gamma) = \gamma^T\mathbf{B}^T\mathbf{B}\gamma$, dan $\mathbf{B}^T\mathbf{B}$ adalah simetris dan dengan Asumsi (2) adalah matrik definit positip. Misalkan $\{v_1,\cdots,v_p\}$ adalah eigenvektor ortonormal dari eigenvalue yang bersesuaian $\{\lambda_1,\cdots,\lambda_p\}$ dari $\Re^p$, maka $\gamma = \sum_{i=1}^{p}\gamma_i^* v_i$ dan

$$\begin{aligned}\|\mathbf{B}\gamma\|^2 &= \left(\sum_{i=1}^{p}\gamma_i^* v_i\right)\mathbf{B}^T\mathbf{B}\left(\sum_{j=1}^{p}\gamma_j^* v_j\right)\\ &= \sum_{i=1}^{p}\sum_{j=1}^{p}\gamma_i^*\gamma_j^* v_i^T\mathbf{B}^T\mathbf{B}v_j = \sum_{i=1}^{p}\sum_{j=1}^{p}\gamma_i^*\gamma_j^* v_i^T\lambda_j v_j\\ &\ge \lambda_*\sum_{i=1}^{p}\sum_{j=1}^{p}\gamma_i^*\gamma_j^* v_i^T\lambda_j v_j\\ &= \lambda_*\sum_{i=1}^{p}\gamma_i^* v_i^T\sum_{j=1}^{p}\gamma_j^* v_j = \lambda_*\gamma^T\gamma = \lambda_*\|\gamma\|^2 = \lambda_*\varsigma^2\end{aligned} \tag{16}$$

dimana $\lambda_*$ adalah eigenvalue terkecil dari matrik $\mathbf{B}^T\mathbf{B}$. Hal ini menunjukkan bahwa $A_2 \ge a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})\lambda_*\varsigma^2$, dan karena Asumsi (2), dan $k > 0$ suatu konstanta positip sedemikian hingga $A_2 \ge a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})kn\varsigma^2$.

Jika dikombinasikan dengan Persamaan (12), untuk setiap $\varepsilon > 0$ diperoleh

$P(A_1 - A_2 \le \varsigma K^*\sqrt{np} - a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})kn\varsigma^2) \ge 1-\varepsilon$

untuk semua $n \ge n_\varepsilon$.

Misalkan $\gamma^T G(\beta) = A_1 - A_2 \le 0$, dan ketidaksamaan,

$$\begin{gathered}\varsigma K^*\sqrt{np} - a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})kn\varsigma^2 \le 0\\ \varsigma K^*\sqrt{np} \le a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})kn\varsigma^2\\ \frac{K^*}{k}\sqrt{\frac{p}{n}}\frac{1}{a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})} \le \varsigma\end{gathered} \tag{17}$$

Sebagai konsekuensi dari Persamaan (10), dan untuk setiap $\varsigma > 0$ dan sejumlah $n_\varsigma$, $n \ge n_\varepsilon$ dengan $\varsigma$ dan $\varepsilon$ positip, maka Persamaan (17) dapat dinyatakan sebagai,

$P(\gamma^T G(\beta) \le 0) \ge 1-\varepsilon$

untuk semua $n \ge \max(n_\varsigma, n_\varepsilon)$ dan $\gamma \in \partial Ber(0,\varsigma)$

Selanjutnya pandang Persamaan (15), yang analog dengan Persamaan (8), yaitu:

$$\begin{aligned}\mathbf{u}^T H_{\ln L}(\beta)\mathbf{u} &= -\sum_{i=1}^{n}(\mathbf{B}_i^T\mathbf{u})^2\sigma_L^{'}(\mathbf{B}_i^T\beta)\\ &\ge a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})\sum_{i=1}^{n}\left(B_i^T\gamma\right)^2 = a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})\|\mathbf{B}\gamma\|^2\\ &\ge a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})kn\|\mathbf{u}\|^2 \ge \inf_{n\in N} a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})kn\|\mathbf{u}\|^2 > 0,\ \mathbf{u} \ne 0\end{aligned}$$

Dari Persamaan (8), jika $na_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})$ konvergen ke-0 dengan $n \to \infty$, diperoleh,

$$\sqrt{\frac{p}{n}}\frac{1}{a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})} = \frac{\sqrt{pn}}{na_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})} \xrightarrow{n\to\infty} \infty$$

hal ini kontradiksi.

Jika $\varsigma$ dipilih sedemikian hingga,

$$\varsigma = \varsigma_n = \frac{K^*}{k}\sqrt{\frac{p}{n}}\frac{1}{a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})}$$

maka menunjukkan konvergen probabilitas ke-1 pada $n \to \infty$,

$$\left\|\hat{\beta} - \beta\right\| = O\left(\frac{p}{n}\frac{1}{(a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta}))^2}\right)$$

dengan $\beta \in \overline{Ber}(\hat{\beta},\varsigma)$. Dengan kata lain, $\frac{n}{p}(a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta}))^2\left\|\hat{\beta} - \beta\right\|$ adalah dalam batas probabilitas.

Pandang matrik Hessian $H_{\varepsilon_{y,\sigma_L}}(\beta)$ yang definit positip dan $\varepsilon_{y,\sigma_L}$ konvex dalam $\overline{Ber}(\hat{\beta},\varsigma)$, dan dibawah Asumsi (1) dan

Asumsi (2), maka minimum absolut $\hat{\beta}$ dari $\varepsilon_{y,\sigma_L}$ (maksimum absolut dari $\ln L$) tidak hanya unik tetapi konvergen dalam probabilitas ke $\hat{\beta}$ pada $n \to \infty$.

Sifat asimtotik kenormalan $\hat{\beta}$ diperlukan asumsi dan teorema sebagai berikut:

Asumsi 3. $B$ adalah variabel random berdistribusi uniform, ada konstanta positif K sedemikian hingga $|B_{ij}| \le K$ untuk semua $i = \{1, \cdots, n\}$, $j = \{1, \cdots, p\}$

Asumsi 4. Jika $\lambda_*$ adalah eigenvalue terkecil dari matrik $\mathbf{B}^T \mathbf{B} \in \Re^{p \times p}$, ada konstanta positif $k > 0$ sedemikian hingga $\lambda_* > kn$ untuk semua n.

Teorema 2. Jika $D(\hat{\beta}) = d_{ij} = \sigma_L^{'}(B_i^T \hat{\beta})$ dan $\Sigma = s_{ij} = \sqrt{\sigma_L^{'}(B_i^T \hat{\beta})}$ untuk $i = j$, maka

$$\frac{\mathbf{e}^T \mathbf{B}^T D(\hat{\beta}) \mathbf{B}(\beta - \hat{\beta})}{\|\Sigma \mathbf{Be}\|} \xrightarrow{n \to \infty} N(0,1)$$

untuk semua $\mathbf{e} \in \Re^p$ dengan $\|\mathbf{e}\| = 1$.

Bukti:

Dari Teorema 1 diketahui bahwa probabilitas mendekati 1 pada $n \to \infty$, estimator $\hat{\beta}$ mempunyai penyelesaian yang unik dari,

$$-\frac{\partial}{\partial \beta} \ln L(\beta) = \mathbf{B}^T (\sigma_L(\mathbf{B}\beta) - \mathbf{y}) = 0$$

yang dapat disederhanakan untuk semua n,

$$\mathbf{B}^T (\sigma_L(\mathbf{B}\beta) - \mathbf{y}) = 0 \tag{18}$$

Persamaan (18) dapat ditulis kembali sebagai,

$$\mathbf{B}^T (\sigma_L(\mathbf{B}(\hat{\beta} + \beta - \hat{\beta})) - \mathbf{y}) = 0$$

$$\Leftrightarrow \quad \mathbf{B}^T (\sigma_L(\mathbf{B}\hat{\beta} + \mathbf{B}(\beta - \hat{\beta})) - \mathbf{y}) = 0$$

$$\Leftrightarrow \sum_{i=1}^{n} B_{ij} \left( \sigma_L(B_i^T \hat{\beta} + B_i^T (\beta - \hat{\beta})) - y_i \right) = 0 \quad , \forall j$$

$$\Leftrightarrow \quad \sum_{i=1}^{n} B_{ij} \left( \sigma_L(\gamma_i^0 + \hat{\varepsilon}_i) - y_i \right) = 0 \quad , \forall j$$

nilai teorema utama, $\sigma_L(\gamma_i^0 + \hat{\varepsilon}_i) = \sigma_L^{'}(\gamma_i^0) + \hat{\varepsilon}_i \sigma_L(\xi_i)$ dimana $\xi_i = \gamma_i^0 + \alpha_i \hat{\varepsilon}_i$ dengan $\alpha_i \in (0,1)$. Kondisi ini ekuivalen dengan,

$$\Leftrightarrow \sum_{i=1}^{n} B_{ij} \left( \hat{\varepsilon}_i \sigma_L^{'}(\xi_i) + \sigma_L(\gamma_i^0) - y_i \right) = 0 \quad , \forall j$$

$$\Leftrightarrow \quad \sum_{i=1}^{n} B_{ij} \left( \hat{\varepsilon}_i \sigma_L^{'}(\xi_i) - r_i \right) = 0 \quad , \forall j$$

Misalkan $\mathbf{e} \in \Re^p$ dengan $\|\mathbf{e}\| = 1$, diperoleh:

$$\begin{aligned} 0 &= \sum_{j=1}^{p} \frac{e_j \sum_{i=1}^{n} B_{ij} (\hat{\varepsilon}_i \sigma_L^{'}(\xi_i) - r_i)}{\|\Sigma \mathbf{Be}\|} \\ &= \sum_{j=1}^{p} \frac{e_j \sum_{i=1}^{n} B_{ij} (\hat{\varepsilon}_i \sigma_L^{'}(\xi_i - \gamma_i^0 + \gamma_i^0) - r_i)}{\|\Sigma \mathbf{Be}\|} \\ &= \underbrace{-\sum_{j=1}^{p} \frac{e_j \sum_{i=1}^{n} B_{ij} r_i}{\|\Sigma \mathbf{Be}\|}}_{A_3} + \underbrace{\sum_{j=1}^{p} \frac{e_j \sum_{i=1}^{n} B_{ij} \hat{\varepsilon}_i (\sigma_L^{'}(\xi_i) - \sigma_L^{'}(\gamma_i^0))}{\|\Sigma \mathbf{Be}\|}}_{A_4} + \underbrace{\sum_{j=1}^{p} \frac{e_j \sum_{i=1}^{n} B_{ij} \hat{\varepsilon}_i \sigma_L^{'}(\gamma_i^0)}{\|\Sigma \mathbf{Be}\|}}_{A_5} \\ &= A_3 + A_4 + A_5 \end{aligned} \tag{19}$$

Untuk $A_3$ dapat ditunjukkan sebagai,

$$-A_3 = \sum_{j=1}^{p} \frac{e_j \sum_{i=1}^{n} B_{ij} r_i}{\|\Sigma \mathbf{Be}\|} = \frac{\mathbf{e}^T \mathbf{B}^T \mathbf{r}}{\|\Sigma \mathbf{Be}\|} \xrightarrow{n \to \infty} N(0,1) \tag{20}$$

Untuk $A_4$ dapat ditunjukkan sebagai,

$$\begin{aligned} \max_{i \in \{1,\cdots,n\}} |\xi_i - \gamma_i^0| &= \max_{i \in \{1,\cdots,n\}} |\gamma_i^0 + \alpha_i \hat{\varepsilon}_i - \gamma_i^0| = \max_{i \in \{1,\cdots,n\}} |\alpha_i B_i^T (\beta - \hat{\beta})| \\ &\le B_{i*}^T (\beta - \hat{\beta}) \quad \text{untuk setiap } i* = \{1, \cdots, n\} \end{aligned}$$

Dengan pertidaksamaan Cauchy-Schwarz dan asumsi (3).,

$$B_{i*}^T (\beta - \hat{\beta}) \le \|B_i\| \|\beta - \hat{\beta}\| = \sqrt{\sum_{j=1}^{p} B_{i*j}^2} \|\beta - \hat{\beta}\| \le \sqrt{p} K \|\beta - \hat{\beta}\|$$

dan, $\beta$ adalah konsisten, $\|\beta - \hat{\beta}\| \xrightarrow[n \to \infty]{P} 0$

dari sini diikuti,

$$\max_{i \in \{1,\cdots,n\}} |\xi_i - \gamma_i^0| \xrightarrow[n \to \infty]{P} 0 \tag{21}$$

Selanjutnya,

$$\begin{aligned} A_4 &= \sum_{j=1}^{p} \frac{e_j \sum_{i=1}^{n} B_{ij} \hat{\varepsilon}_i (\sigma_L^{'}(\xi_i) - \sigma_L^{'}(\gamma_i^0))}{\|\Sigma \mathbf{Be}\|} \\ &= \frac{\|\mathbf{e}\|}{\|\Sigma \mathbf{Be}\|} \sqrt{\sum_{j=1}^{p} \left| \sum_{i=1}^{n} B_{ij} \hat{\varepsilon}_i (\sigma_L^{'}(\xi_i) - \sigma_L^{'}(\gamma_i^0)) \right|^2} \\ &= \frac{1}{\|\Sigma \mathbf{Be}\|} \sqrt{pn^2 K^2 \hat{\varepsilon}^2 \max_{i \in \{1,\cdots,n\}} (\sigma_L^{'}(\xi_i) - \sigma_L^{'}(\gamma_i^0))^2} \\ &= \frac{1}{\|\Sigma \mathbf{Be}\|} \sqrt{p} nK \hat{\varepsilon} \max_{i \in \{1,\cdots,n\}} |\sigma_L^{'}(\xi_i) - \sigma_L^{'}(\gamma_i^0)| \end{aligned}$$

dimana, $\hat{\varepsilon} = \max_{i \in \{1,\cdots,n\}} |\hat{\varepsilon}_i|$. Untuk penyebutnya,

$$\|\Sigma \mathbf{Be}\| = \sqrt{\sum_{i=1}^{n} \sigma_L^{'}(B_i^T \hat{\beta})(\mathbf{Be})_i^2} \ge \|\mathbf{Be}\| \min_{i \in \{1,\cdots,n\}} \underbrace{\sqrt{\sigma_L^{'}(B_i^T \hat{\beta})}}_{s_i}$$

dan ruas kiri dan dengan Asumsi (3), $|B_i^T \hat{\beta}| \le \|B_i\| \|\hat{\beta}\| \le \sqrt{p} K \|\hat{\beta}\|$ sedemikian hingga, $s_i^2 = \sigma_L^{'}(\sqrt{p} K \|\hat{\beta}\|) = k_0$. Sedangkan ruas lainnya, $\|\mathbf{Be}\| \geq \sqrt{kn}$ (yaitu, kombinasi Persamaan (16) dan Asumsi (4), ditunjukkan dalam persamaan,

$$\|\Sigma\mathbf{Be}\| \geq \sqrt{k_o kn} \tag{22}$$

sehingga,

$$|A_4| = \frac{1}{\sqrt{k_o kn}}\sqrt{p}K\sqrt{n}\hat{\varepsilon} \max_{i\in\{1,\cdots,n\}} |\sigma'_L(\xi_i) - \sigma'_L(\gamma_i^0)| \tag{23}$$

Dengan $\sigma'_L$ kontinu dan persamaan (21) diperoleh,

$$\max_{i\in\{1,\cdots,n\}} |\sigma'_L(\xi_i) - \sigma'_L(\gamma_i^0)| \xrightarrow[n\to\infty]{P} 0 \tag{24}$$

Digunakan lagi Cauchy-Schwarz,

$$\hat{\varepsilon} = \max_{i\in\{1,\cdots,n\}} |\hat{\varepsilon}_i| = \max_{i\in\{1,\cdots,n\}} |B_i^T(\beta-\hat{\beta})| \leq \sqrt{p}K\left\|\beta-\hat{\beta}\right\|$$

Karena p tetap dan $\mathbf{B}Ber(\hat{\beta},\varsigma)$ terbatas, K konstanta positif $K > 0$ sedemikian hingga,

$$\left|a_{pn}^{\varsigma}(\mathbf{B},\hat{\beta})\right| = \left|\inf_{\varsigma\in\mathbf{B}Ber(\hat{\beta},\varsigma)} \min_{i\in\{1,\cdots,n\}} \sigma'_L(\zeta_i)\right| \geq K.$$

Selanjutnya $\sqrt{n}\left\|\beta-\hat{\beta}\right\|$ adalah terbatas dalam probabilitas, dan Persamaan (23) dan (24), maka, $A_4 \xrightarrow[n\to\infty]{P} 0$.

Sedangkan untuk $A_5$ diperoleh sebagai,

$$A_5 = \sum_{j=1}^{p} \frac{e_j \sum_{i=1}^{n} B_{ij}\hat{\varepsilon}_i\sigma'_L(\gamma_i^0)}{\left\|\sum\mathbf{Be}\right\|}$$
$$= \sum_{j=1}^{p} \frac{e_j \sum_{i=1}^{n} B_{ij}B_i^T(\beta-\hat{\beta})\sigma'_L(B_i^T\hat{\beta})}{\left\|\sum\mathbf{Be}\right\|}$$
$$= \frac{\mathrm{e}^T\mathbf{B}^T D(\hat{\beta})\mathbf{B}(\beta-\hat{\beta})}{\left\|\sum\mathbf{Be}\right\|}$$

Dari Persamaan (19), merupakan,

$$A_3 + A_5 \xrightarrow[n\to\infty]{P} 0 \tag{25}$$

dimana

$$-A_3 = \sum_{j=1}^{p} \frac{e_j \sum_{i=1}^{n} B_{ij}r_i}{\left\|\sum\mathbf{Be}\right\|} = \frac{\mathbf{e}^T\mathbf{B}^T\mathbf{r}}{\left\|\sum\mathbf{Be}\right\|} \xrightarrow{n\to\infty} N(0,1)$$

dan dari Persamaan (25) menunjukkan bahwa $A_5 \xrightarrow{n\to\infty} N(0,1)$.

Selanjutnya, pandang $\mathbf{e}^T\mathbf{B}^T\mathbf{r} = \rho_i = (\mathbf{Be})_i r_i$ pada $A_3$, maka:

$$\sum_{i=1}^{n} Var(\rho_i) = \sum_{i=1}^{n} (\mathbf{Be})_i^2 Var(r_i) = \sum_{i=1}^{n} (\mathbf{Be})_i^2 s_i^2$$
$$= \sum_{i=1}^{n} (\Sigma\mathbf{Be})_i^2 = \|\Sigma\mathbf{Be}\|^2$$

Dengan menggunakan teorema limit pusat dari Linderberg', untuk semua $t > 0$, maka,

$$\sum_{i=1}^{n} \frac{E(\rho_i^2 \,\|\rho_i| \geq t\|\Sigma\mathbf{Be}\|)}{\|\Sigma\mathbf{Be}\|} \xrightarrow{n\to\infty} 0 \tag{26}$$

Sehingga,

$$\frac{Var(\rho_i)}{\|\Sigma\mathbf{Be}\|^2} = \frac{s_i^2(B_i^T\mathbf{e})^2}{\|\Sigma\mathbf{Be}\|^2} \leq \frac{s_i^2\|B_i\|^2\|\mathbf{e}\|^2}{\|\Sigma\mathbf{Be}\|^2} \leq \xrightarrow{n\to\infty} 0 \tag{27}$$

Persamaan ini adalah konsekuensi dari Asumsi (3), dan menunjukkan bahwa $s_i^2 = \sigma'_L(B_i^T\hat{\beta}) \leq \frac{1}{4}$.

## DAFTAR PUSTAKA

Dudoit S, Fridlyand J & Speed TP. 2002. Comparison of Discriminantion Methods for the Classifications of Tumors Using Gene Expression Data. *Journal of the American Statistical Association*. **97**(457): 77 – 87.

Emond M & Steven G. 1997. An Efficient Estimator for the Generalized Semi linear Model, *Journal of the American Statistical Association.* **92**, 1033-1040.

Eubank R. 1988. *Spline Smoothing and Nonparametric Regression*. New York: Marcel Dekker.

Friedman JH. 1990. *Estimating functions of mixed ordinal and categorical variables using Multivariate Adaptive Regression Splines*. Technical Report LCS 107. Statistics Department. Stanford University.

Friedman JH. 1990. Multivariate Adaptive Regression Splines (with discussion). *Ann. Statist.* **19**: 1 – 141.

Hardle W. 1990. *Smoothing Techniques With Implementation in S.* Springer-Verlag: New York.

Holmes CC & Denison DGT. 2002. *Classification with Bayesian MARS.* Department of Mathematics. Imperial college. London. sw7.2BZ, UK.

Xiang R & Meullenet JF. 2002. Comparasion of Logistic Regression and MARS in modeling the effects of water activity, pH and potassium sorbate on growth – no growth of *Saccharomyces cerevisiae.* Food Science Department. University of Arkansan.