X1-109 (2015/07/30)

برازش، تخمین، نایقینی I

محمد خرمی mamwad@mailaps.org

برازش ِ یک دسته داده بررسی میشود. مشخصتر، بهینه-سازی یِ پارامترها یِ برازش و نایقینی یِ این پارامترها، بر اساس ِ نتایج ِ سنجش، بررسی میشود.

0 درامد

بین ِ یک دسته متغیر ِ $\boldsymbol{X}$ روابط ی برقرار است. این روابط با یک دسته پارامتر ِ $\boldsymbol{A}$ مشخص میشوند:

$$\boldsymbol{Q}(\boldsymbol{A}, \boldsymbol{X}) = 0. \tag{1}$$

اگر $\boldsymbol{X}$ به دقت معلوم باشد، و تعداد ِ مئلفِها یِ $\boldsymbol{Q}$ با تعداد ِ مئلفِها یِ $\boldsymbol{A}$ باشد، به شرط ی که معادلِها یِ (1) مستقل باشند (که از این پس فرض میکنم چنین است)، میشود $\boldsymbol{A}$ را حساب کرد. عملَن $\boldsymbol{X}$ با سنجش به دست میئاید، پس خطا دارد، و تعداد ِ مئلفِها یِ $\boldsymbol{Q}$ معمولَن بیش از تعداد ِ مئلفِها یِ $\boldsymbol{A}$ است. پس اگر مقدارها یِ سنجیده برا یِ $\boldsymbol{X}$ به جا یِ خُد ِ $\boldsymbol{X}$ به کار رود، ممکن است معادلات ِ (1) برا یِ $\boldsymbol{A}$ جواب نداشته باشند (معمولَن چنین است). جواب اگر وجود داشته باشد هم با مقدار ِ واقعی فرق دارد، چون به جا یِ $\boldsymbol{X}$ تخمین ی از آن به کار رفته. هدف این است که با داشتن ِ مقدارها یِ سنجیده،

تخمین ی برا یِ $\boldsymbol{X}$ و $\boldsymbol{A}$ به دست آید، و انحراف ِ این تخمینها از مقدارها یِ واقعی تخمین زده شود.

مقدار ی که با سنجش برا یِ $\boldsymbol{X}$ به دست میئاید را با $\boldsymbol{Y}$ نشان میدهم. چگالی-یِ-احتمال را هم با pro نشان میدهم. از

$$\mathsf{pro}(\boldsymbol{Y}=\boldsymbol{y}|\boldsymbol{A}=\boldsymbol{a},\boldsymbol{X}=\boldsymbol{x})\,\mathsf{pro}(\boldsymbol{A}=\boldsymbol{a},\boldsymbol{X}=\boldsymbol{x})$$
$$=\mathsf{pro}(\boldsymbol{A}=\boldsymbol{a},\boldsymbol{X}=\boldsymbol{x}|\boldsymbol{Y}=\boldsymbol{y})\,\mathsf{pro}(\boldsymbol{Y}=\boldsymbol{y}) \tag{2}$$

یک شکل ِ قضیه یِ بیز [1] نتیجه میشود (مثلَن [2]):

$$\mathsf{pro}(\boldsymbol{A}=\boldsymbol{a},\boldsymbol{X}=\boldsymbol{x}|\boldsymbol{Y}=\boldsymbol{y})=\frac{\mathsf{pro}(\boldsymbol{A}=\boldsymbol{a},\boldsymbol{X}=\boldsymbol{x})}{\mathsf{pro}(\boldsymbol{Y}=\boldsymbol{y})}$$
$$\mathsf{pro}(\boldsymbol{Y}=\boldsymbol{y}|\boldsymbol{A}=\boldsymbol{a},\boldsymbol{X}=\boldsymbol{x}). \tag{3}$$

مقدار ِ واقعی یِ $\mathfrak{Q}$ را با $\hat{\mathfrak{Q}}$ نشان میدهم.

1 تُزیع ِ احتمال

برا یِ $\mathsf{pro}(\boldsymbol{A}=\boldsymbol{a},\boldsymbol{X}=\boldsymbol{x})$ این معلوم است که $\boldsymbol{a}$ و $\boldsymbol{x}$ باید قید ی شبیه ِ (1) را براورَند. پس

$$\mathsf{pro}(\boldsymbol{A}=\boldsymbol{a},\boldsymbol{X}=\boldsymbol{x})=g(\boldsymbol{a},\boldsymbol{x})\,\delta[\boldsymbol{Q}(\boldsymbol{a},\boldsymbol{x})], \tag{4}$$

که g هموار فرض میشود. متغیرها یِ $(\boldsymbol{A},\boldsymbol{X})$ را به متغیرها یِ $(\boldsymbol{B},\boldsymbol{B}')$ تغییر میدهم، چنان که $\boldsymbol{B}'=0$ با $\boldsymbol{Q}=0$ همئَرز است. از جمله میشود $\boldsymbol{B}'$ را خُد ِ $\boldsymbol{Q}$ گرفت. $(\boldsymbol{A},\boldsymbol{X})$ ی که (1) را براورَد چنین میشود

$$\boldsymbol{A}=\boldsymbol{f}_0(\boldsymbol{B}). \tag{5}$$

$$\boldsymbol{X}=\boldsymbol{f}_1(\boldsymbol{B}). \tag{6}$$

عنصر ِ حجم برا یِ $(\boldsymbol{A},\boldsymbol{X})$ را هم میشود بر حسب ِ $\boldsymbol{B}$ و $\boldsymbol{B}'$ نوشت. به این ترتیب، با انتگرالگیری از (3) بر $\boldsymbol{B}'$ چگالی یِ احتمال برا یِ $\boldsymbol{B}$ به شرط ِ $\boldsymbol{Y}=\boldsymbol{y}$ به دست میئاید:

$$\mathsf{pro}(\boldsymbol{B}=\boldsymbol{b}|\boldsymbol{Y}=\boldsymbol{y})=\frac{g[\boldsymbol{f}_0(\boldsymbol{b}),\boldsymbol{f}_1(\boldsymbol{b})]}{\mathsf{pro}(\boldsymbol{Y}=\boldsymbol{y})}\,h(\boldsymbol{b})$$
$$\mathsf{pro}[\boldsymbol{Y}=\boldsymbol{y}|\boldsymbol{A}=\boldsymbol{f}_0(\boldsymbol{b}),\boldsymbol{X}=\boldsymbol{f}_1(\boldsymbol{b})], \tag{7}$$

که h ناشی از عنصرِ انتگرالگیری و رابطه یِ $\boldsymbol{B}'$ با $\boldsymbol{Q}$ است.

تعدادِ متغیرها (تعدادِ سنجشها) را با N، تعدادِ پارامترها را با m، و تعدادِ رابطِها یِ بینِ اینها (تعدادِ قیدها) را با η نشان میدهم. تعدادِ متغیرها یِ مستقل (تعدادِ مئلفِها یِ $\boldsymbol{B}$) را هم با ϕ نشان میدهم. دیده میشود

$$\phi = N + m - \eta. \tag{8}$$

اگر خطا یِ سنجش کم باشد، چگالی یِ احتمالِ $\boldsymbol{Y}$ به شرطِ $\boldsymbol{X}$ و $\boldsymbol{A}$ حُلِ $\boldsymbol{X}$ جایگزیده و تیز است. با فرضِ این که بقیه یِ عبارتها یِ طرفِ راستِ (7) نسبت به $\boldsymbol{b}$ کندتغییر َند، میشود در آنها به جا یِ $\boldsymbol{b}$ مقدار ی را گذاشت که $\mathsf{pro}[\boldsymbol{Y} = \boldsymbol{y}|\boldsymbol{A} = \boldsymbol{f}_0(\boldsymbol{b}), \boldsymbol{X} = \boldsymbol{f}_1(\boldsymbol{b})]$ را بیشینه میکند. به این ترتیب،

$$\mathsf{pro}(\boldsymbol{B} = \boldsymbol{b}|\boldsymbol{Y} = \boldsymbol{y}) = \mathcal{N}_1\, \mathsf{pro}[\boldsymbol{Y} = \boldsymbol{y}|\boldsymbol{A} = \boldsymbol{f}_0(\boldsymbol{b}), \boldsymbol{X} = \boldsymbol{f}_1(\boldsymbol{b})], \tag{9}$$

که $\mathcal{N}_1$ یک ضریبِ بهنجارش است.

وقت ی چگالی یِ احتمالِ $\boldsymbol{Y}$ به شرطِ $\boldsymbol{X}$ و $\boldsymbol{A}$ حُلِ $\boldsymbol{X}$ جایگزیده و تیز باشد، میشود آن را با یک گاؤسی تقریب کرد:

$$\begin{aligned}\mathsf{pro}(\boldsymbol{Y} = \boldsymbol{y}|\boldsymbol{A} = \boldsymbol{a}, \boldsymbol{X} = \boldsymbol{x}) &= \mathcal{N}_2\, \exp\left[-\frac{1}{2}\, C^{-1}_{i\,j}\, (y^i - x^i)\, (y^j - x^j)\right],\\ &= \mathcal{N}_2\, \exp\left[-\frac{1}{2}\, \boldsymbol{C}^{-1}\, (\boldsymbol{y} - \boldsymbol{x}, \boldsymbol{y} - \boldsymbol{x})\right]\end{aligned} \tag{10}$$

که $\mathcal{N}_2$ یک ضریبِ بهنجارش است و $\boldsymbol{C}$ همبستِگی یِ $\boldsymbol{Y}$ است:

$$\mathsf{E}\{[Y^i - \mathsf{E}(Y^i)]\, [Y^j - \mathsf{E}(Y^j)]\} = C^{i\,j}. \tag{11}$$

E مقدارِ چشمداشتی ست. به این ترتیب،

$$\begin{aligned}\mathsf{pro}(\boldsymbol{B} = \boldsymbol{b}|\boldsymbol{Y} = \boldsymbol{y}) &= \mathcal{N}_3\, \exp\left\{-\frac{1}{2}\, C^{-1}_{i\,j}\, [y^i - f_1^i(\boldsymbol{b})]\, [y^j - f_1^j(\boldsymbol{b})]\right\},\\ &= \mathcal{N}_3\, \exp\left\{-\frac{1}{2}\, \boldsymbol{C}^{-1}\, [\boldsymbol{y} - \boldsymbol{f}_1(\boldsymbol{b}), \boldsymbol{y} - \boldsymbol{f}_1(\boldsymbol{b})]\right\},\end{aligned} \tag{12}$$

که $\mathcal{N}_3$ یک ضریبِ بهنجارش است.

$\mathsf{pro}(\boldsymbol{B} = \boldsymbol{b}|\boldsymbol{Y} = \boldsymbol{y})$ نسبت به $\boldsymbol{b}$ تیز است. پس میشود آن را با یک گاؤسی (نسبت به $\boldsymbol{b}$) تقریب

کرد:

$$\mathsf{pro}(\boldsymbol{B}=\boldsymbol{b}|\boldsymbol{Y}=\boldsymbol{y}) = \mathcal{N}_4\,\exp\left\{-\frac{1}{2}\,L_{\alpha\,\beta}\,[b^\alpha-\bar{B}^\alpha(\boldsymbol{y})]\,[b^\beta-\bar{B}^\beta(\boldsymbol{y})]\right\},$$

$$= \mathcal{N}_4\,\exp\left\{-\frac{1}{2}\,\boldsymbol{L}\,[\boldsymbol{b}-\bar{\boldsymbol{B}}(\boldsymbol{y}),\boldsymbol{b}-\bar{\boldsymbol{B}}(\boldsymbol{y})]\right\} \qquad (13)$$

که $\mathcal{N}_4$ یک ثابتِ بهنجارش است، $\bar{\boldsymbol{B}}(\boldsymbol{y})$ و $\boldsymbol{L}$ مستقل از $\boldsymbol{b}$ اند. دیده میشود $\bar{\boldsymbol{B}}(\boldsymbol{y})$ مقدار ی برا یِ $\boldsymbol{b}$ است که $\mathsf{pro}(\boldsymbol{B}=\boldsymbol{b}|\boldsymbol{Y}=\boldsymbol{y})$ را بیشینه میکند. پس میشود آن را با بیشینه-کردنِ (12) نسبت به $\boldsymbol{b}$ حساب کرد:

$$C^{-1}_{i\,j}\,\{f_1^i[\bar{\boldsymbol{B}}(\boldsymbol{y})]-y^i\}\,\{(\mathsf{D}_\beta\,f_1^j)[\bar{\boldsymbol{B}}(\boldsymbol{y})]\}=0, \qquad (14)$$

که D مشتقگیری ست. بیشینه-کردنِ نما یِ گاؤسی در (12) نسبت به $\boldsymbol{b}$، همٔرز با بیشینه-کردنِ نما یِ گاؤسی در (10) نسبت به $\boldsymbol{x}$ (و $\boldsymbol{a}$) با قیدِ (1) است. این بیشینه-کردن را میشود بدونِ حلِ قید هم انجام داد. به روشِ ضریبها یِ لَگرانژ [3]، تعریف میکنم

$$F(\boldsymbol{a},\boldsymbol{x},\boldsymbol{\lambda}) = -\frac{1}{2}\,C^{-1}_{i\,j}\,(y^i-x^i)\,(y^j-x^j)+\lambda_\nu\,Q^\nu(\boldsymbol{a},\boldsymbol{x}). \qquad (15)$$

شرطِ فرینه-شدن با قید میشود (مثلَن [4])

$$(\mathsf{D}_{0,p}\,F)[\bar{\boldsymbol{A}}(\boldsymbol{y}),\bar{\boldsymbol{X}}(\boldsymbol{y}),\bar{\boldsymbol{\Lambda}}(\boldsymbol{y})]=0, \qquad (16)$$

$$(\mathsf{D}_{1,k}\,F)[\bar{\boldsymbol{A}}(\boldsymbol{y}),\bar{\boldsymbol{X}}(\boldsymbol{y}),\bar{\boldsymbol{\Lambda}}(\boldsymbol{y})]=0, \qquad (17)$$

$$(\mathsf{D}_{2,\rho}\,F)[\bar{\boldsymbol{A}}(\boldsymbol{y}),\bar{\boldsymbol{X}}(\boldsymbol{y}),\bar{\boldsymbol{\Lambda}}(\boldsymbol{y})]=0, \qquad (18)$$

که D_0 و D_1 و D_2 مشتقگیری نسبت به، به ترتیب، $\boldsymbol{a}$ و $\boldsymbol{x}$ و $\boldsymbol{\lambda}$ اند. این رابطِها میشوند

$$[\bar{\Lambda}_\nu(\boldsymbol{y})]\,\{(\mathsf{D}_{0,p}\,Q^\nu)[\bar{\boldsymbol{A}}(\boldsymbol{y}),\bar{\boldsymbol{X}}(\boldsymbol{y})]\}=0, \qquad (19)$$

$$C^{-1}_{k\,j}\,[y^j-\bar{X}^j(\boldsymbol{y})]+[\bar{\Lambda}_\nu(\boldsymbol{y})]\,\{(\mathsf{D}_{1,k}\,Q^\nu)[\bar{\boldsymbol{A}}(\boldsymbol{y}),\bar{\boldsymbol{X}}(\boldsymbol{y})]\}=0, \qquad (20)$$

$$Q^\rho[\bar{\boldsymbol{A}}(\boldsymbol{y}),\bar{\boldsymbol{X}}(\boldsymbol{y})]=0. \qquad (21)$$

دیده میشود (16) و (17) و (18)، یا همٔرز با آنها (19) و (20) و (21)، به ترتیب m و N و η معادله

اند. رُشن است که

$$\bar{\mathfrak{Q}}(\hat{\boldsymbol{Y}}) = \hat{\mathfrak{Q}}. \tag{22}$$

$$\boldsymbol{f}_1(\hat{\boldsymbol{B}}) = \hat{\boldsymbol{Y}}. \tag{23}$$

از اینجا دیده میشود

$$\boldsymbol{f}_1(\boldsymbol{b}) - \boldsymbol{y} = [\boldsymbol{f}_1(\boldsymbol{b}) - \boldsymbol{f}_1(\hat{\boldsymbol{B}})] - (\boldsymbol{y} - \hat{\boldsymbol{Y}}). \tag{24}$$

اگر تُزیعها تیز باشند، فقط برا یِ انحرافها یِ کوچک ِ $\boldsymbol{y}$ از $\hat{\boldsymbol{Y}}$ و $\boldsymbol{b}$ از $\hat{\boldsymbol{B}}$ است که چگالی-یِ-احتمالها چشمگیر َند. پس،

$$f_1^i(\boldsymbol{b}) - \boldsymbol{f}_1^i(\hat{\boldsymbol{B}}) = \Gamma^i{}_\alpha\,(b^\alpha - \hat{B}^\alpha), \tag{25}$$

که $\Gamma^i{}_\alpha$ مشتق ِ f_1^i نسبت به b^α است، که جا یی نزدیک ِ $\hat{\boldsymbol{B}}$، از جمله $\bar{\boldsymbol{B}}(\boldsymbol{y})$، حساب شده.

با جاگذاری یِ (25) و (24) در (12)، نتیجه میشود

$$L_{\alpha\,\beta} = C^{-1}_{i\,j}\,\Gamma^i{}_\alpha\,\Gamma^j{}_\beta. \tag{26}$$

شرط ِ این که (13) یک چگالی-یِ-احتمال (ِ بهنجارش-پذیر) برا یِ $\boldsymbol{B}$ باشد این است که $\boldsymbol{L}$ مثبت- ِ-معین باشد. و شرط ِ این-یکی هم آن است که ϕ (تعداد ِ متغیرها یِ مستقل) از N (تعداد ِ متغیرها) بیشتر نباشد. اگر ϕ بزرگتر از N باشد، معادلات ِ

$$\boldsymbol{f}_1(\boldsymbol{b}) = \boldsymbol{y} \tag{27}$$

برا یِ $\boldsymbol{b}$ جواب دارند و بُعد ِ مجموعه یِ این جوابها $(\phi - N)$ است. در این صورت طرف ِ راست ِ (12) بر این مجموعه ثابت است، که نشان میدهد طرف ِ راست ِ (13) بر این مجموعه ثابت است. پس یک زیرفضا یِ $(\phi - N)$ بُعدی هست که اثر ِ $\boldsymbol{L}$ بر آن صفر میشود. از این پس فرض میکنم ϕ کوچکتر از N است (همرَز با آن، m کوچکتر از η است) و $\boldsymbol{L}$ مثبت- ِ-معین (و در نتیجه وارون-پذیر) است. وارون ِ $\boldsymbol{L}$ را با $\boldsymbol{K}$ نشان میدهم.

برا یِ تُزیعها یِ تیز، (14) به این شکل در میئاید.

$$C^{-1}_{i\,j}\,\{\Gamma^i{}_\alpha\,[\bar{B}^\alpha(\boldsymbol{y}) - \hat{B}^\alpha] - (y^i - \hat{Y}^i)\}\,\Gamma^j{}_\beta = 0, \tag{28}$$

که نتیجه میدهد

$$\bar{B}^\alpha(\boldsymbol{y}) - \hat{B}^\alpha = K^{\alpha\,\beta}\,[C^{-1}_{i\,j}\,\Gamma^j{}_\beta\,(y^i - \hat{Y}^i)]. \tag{29}$$

از جمله دیده میشود

$$\frac{\partial \bar{B}^{\alpha}}{\partial y^{i}} = K^{\alpha\,\beta}\, C^{-1}_{i\,j} \Gamma^{j}{}_{\beta}. \tag{30}$$

شاخصها را با $\boldsymbol{C}$ و $\boldsymbol{K}$ بالا-وُ-پایین میبرم. به این ترتیب،

$$\frac{\partial [\bar{B}^{\alpha}(\boldsymbol{y})]}{\partial y^{i}} = \Gamma_{i}{}^{\alpha}. \tag{31}$$

همچنین،

$$\Gamma_{i}{}^{\alpha}\, \Gamma^{i\,\beta} = K^{\alpha\,\beta}. \tag{32}$$

$$\Gamma_{i}{}^{\alpha}\, \Gamma^{i}{}_{\beta} = \delta^{\alpha}_{\beta}. \tag{33}$$

2 گشتاورها

رابطه یِ (11)، وقت ی $\mathsf{E}(\boldsymbol{Y})$ با $\hat{\boldsymbol{Y}}$ جایگزین شود میشود

$$\mathsf{E}[(Y^{i} - \hat{Y}^{i})\,(Y^{j} - \hat{Y}^{j})] = C^{i\,j}. \tag{34}$$

به این ترتیب، با استفاده از (29) یا (31)،

$$\begin{aligned}\mathsf{E}[(\bar{B}^{\alpha} - \hat{B}^{\alpha})\,(\bar{B}^{\beta} - \hat{B}^{\beta})] &= \Gamma_{i}{}^{\alpha}\, \Gamma_{j}{}^{\beta}\, C^{i\,j},\\ &= K^{\alpha\,\beta}.\end{aligned} \tag{35}$$

رابطِها یِ (5) و (6) هم برا یِ انحرافها یِ کم از مقدارها یِ واقعی میشوند

$$\bar{A}^{p} - \hat{A}^{p} = \Theta^{p}{}_{\alpha}\,(\bar{B}^{\alpha} - \hat{B}^{\alpha}), \tag{36}$$

$$\bar{X}^{i} - \hat{X}^{i} = \Gamma^{i}{}_{\alpha}\,(\bar{B}^{\alpha} - \hat{B}^{\alpha}), \tag{37}$$

که $\Theta^{p}{}_{\alpha}$ مشتق f^{i}_{0} نسبت به b^{α} است، که جا یی نزدیک ِ $\hat{\boldsymbol{B}}$ (از جمله $\bar{\boldsymbol{B}}$) حساب شده. به این ترتیب،

$$\bar{A}^{p}(\boldsymbol{y}) - \hat{A}^{p} = \underset{\sim}{\Theta}^{p}{}_{j}\,(y^{j} - \hat{Y}^{j}), \tag{38}$$

$$\bar{X}^{i}(\boldsymbol{y}) - \hat{X}^{i} = \underset{\sim}{\Gamma}^{i}{}_{j}\,(y^{j} - \hat{Y}^{j}), \tag{39}$$

که

$$\utilde{\boldsymbol{\mathfrak{Q}}} = \boldsymbol{\mathfrak{Q}}\,\frac{\partial[\bar{\boldsymbol{B}}(\boldsymbol{y})]}{\partial \boldsymbol{y}}. \tag{40}$$

این ضمنَن یعنی $\utilde{\boldsymbol{\Theta}}$ و $\utilde{\boldsymbol{\Gamma}}$ ماتریسها یِ مشتق ِ، به ترتیب، $\bar{\boldsymbol{A}}(\boldsymbol{y})$ و $\bar{\boldsymbol{X}}(\boldsymbol{y})$ نسبت به $\boldsymbol{y}$ اند. از (38) و (39) نتیجه میشود

$$\mathsf{E}[(\bar{A}^p - \hat{A}^p)\,(\bar{A}^q - \hat{A}^q)] = \utilde{\Theta}^p{}_k\,\utilde{\Theta}^{q\,k}. \tag{41}$$

$$\mathsf{E}[(\bar{X}^i - \hat{X}^i)\,(\bar{X}^j - \hat{X}^j)] = \utilde{\Gamma}^i{}_k\,\utilde{\Gamma}^{j\,k}. \tag{42}$$

یا صریحتر،

$$\mathsf{E}[(\bar{A}^p - \hat{A}^p)\,(\bar{A}^q - \hat{A}^q)] = C^{k\,l}\,\utilde{\Theta}^p{}_k\,\utilde{\Theta}^q{}_l. \tag{43}$$

$$\mathsf{E}[(\bar{X}^i - \hat{X}^i)\,(\bar{X}^j - \hat{X}^j)] = C^{k\,l}\,\utilde{\Gamma}^i{}_k\,\utilde{\Gamma}^j{}_l. \tag{44}$$

و به شکل ی دیگر،

$$\mathsf{E}[(\bar{A}^p - \hat{A}^p)\,(\bar{A}^q - \hat{A}^q)] = \Theta^p{}_\alpha\,\Theta^{q\,\alpha}. \tag{45}$$

$$\mathsf{E}[(\bar{X}^i - \hat{X}^i)\,(\bar{X}^j - \hat{X}^j)] = \Gamma^i{}_\alpha\,\Gamma^{j\,\alpha}. \tag{46}$$

یا صریحتر،

$$\mathsf{E}[(\bar{A}^p - \hat{A}^p)\,(\bar{A}^q - \hat{A}^q)] = K^{\alpha\,\beta}\,\Theta^p{}_\alpha\,\Theta^q{}_\beta. \tag{47}$$

$$\mathsf{E}[(\bar{X}^i - \hat{X}^i)\,(\bar{X}^j - \hat{X}^j)] = K^{\alpha\,\beta}\,\Gamma^i{}_\alpha\,\Gamma^j{}_\beta. \tag{48}$$

3 تخمین ِ گشتاورها

از (43) و (44) دیده میشود همبستگی یِ $\boldsymbol{A}$ و نیز همبستِگی یِ $\boldsymbol{X}$ به $\boldsymbol{C}$ مربوط است. $\boldsymbol{C}$ هم نماینده یِ خطا یِ سنجش ِ $\boldsymbol{Y}$ است. این خطا را میشود از رو یِ کمیتها یِ سنجیده تخمین زد. از (24) و (25) و (31) دیده میشود برا یِ انحرافها یِ کم از مقدارها یِ واقعی

$$\begin{aligned} y^i - f_1^i(\bar{\boldsymbol{B}}) &= y^i - \hat{Y}^i - \Gamma^i{}_\alpha\,[\bar{B}^\alpha(\boldsymbol{y}) - \hat{B}^\alpha], \\ &= y^i - \hat{Y}^i - \Gamma^i{}_\alpha\,\Gamma_k{}^\alpha\,(y^k - \hat{Y}^k). \end{aligned} \tag{49}$$

به این ترتیب،

$$\begin{aligned}
&\mathsf{E}\{\boldsymbol{C}^{-1}[\boldsymbol{Y}-\boldsymbol{f}_1(\bar{\boldsymbol{B}}),\boldsymbol{Y}-\boldsymbol{f}_1(\bar{\boldsymbol{B}})]\}\\
&=\mathsf{E}\{C^{-1}_{i\,j}\,[Y^i-f_1^i(\bar{\boldsymbol{B}})]\,[Y^j-f_1^j(\bar{\boldsymbol{B}})]\}\\
&=C^{-1}_{i\,j}\,\{\mathsf{E}[(Y^i-\hat{Y}^i)\,(Y^j-\hat{Y}^j)]-2\,\Gamma^i{}_\alpha\,\Gamma_k{}^\alpha\,\mathsf{E}[(Y^k-\hat{Y}^k)\,(Y^j-\hat{Y}^j)]\\
&\quad+\Gamma^i{}_\alpha\,\Gamma_k{}^\alpha\,\Gamma^j{}_\beta\,\Gamma_l{}^\beta\,\mathsf{E}[(Y^k-\hat{Y}^k)\,(Y^l-\hat{Y}^l)]\},\\
&=C^{-1}_{i\,j}\,[C^{i\,j}-2\,\Gamma^i{}_\alpha\,\Gamma_k{}^\alpha\,C^{k\,j}+\Gamma^i{}_\alpha\,\Gamma_k{}^\alpha\,\Gamma^j{}_\beta\,\Gamma_l{}^\beta\,C^{k\,l}],
\end{aligned}\tag{50}$$

که نتیجه میدهد

$$\begin{aligned}
\mathsf{E}\{\boldsymbol{C}^{-1}[\boldsymbol{Y}-\boldsymbol{f}_1(\bar{\boldsymbol{B}}),\boldsymbol{Y}-\boldsymbol{f}_1(\bar{\boldsymbol{B}})]\}&=C^i_i-K^\alpha_\alpha,\\
&=N-\phi,
\end{aligned}\tag{51}$$

یا، با استفاده از (8)،

$$\mathsf{E}\{\boldsymbol{C}^{-1}[\boldsymbol{Y}-\boldsymbol{f}_1(\bar{\boldsymbol{B}}),\boldsymbol{Y}-\boldsymbol{f}_1(\bar{\boldsymbol{B}})]\}=\eta-m.\tag{52}$$

از اینجا این تخمین به دست میئاید.

$$\mathsf{es}\{\boldsymbol{C}^{-1}\,[\boldsymbol{y}-\bar{\boldsymbol{X}}(\boldsymbol{y}),\boldsymbol{y}-\bar{\boldsymbol{X}}(\boldsymbol{y})]\}=\eta-m.\tag{53}$$

این در واقع یک رابطه یِ تخمینی بین ِ مثلفِها یِ $\boldsymbol{C}$ است: اگر

$$\boldsymbol{C}=\sigma^2\,\underset{\circ}{\boldsymbol{C}},\tag{54}$$

که σ یک عدد ِ مثبت ِ نامعلوم است و $\underset{\circ}{\boldsymbol{C}}$ معلوم است،

$$\mathsf{es}(\sigma^2)=\frac{\underset{\circ}{\boldsymbol{C}}^{-1}\,[\boldsymbol{y}-\bar{\boldsymbol{X}}(\boldsymbol{y}),\boldsymbol{y}-\bar{\boldsymbol{X}}(\boldsymbol{y})]}{\eta-m}.\tag{55}$$

این را در (43) و (44) مینشانم:

$$\mathsf{es}\{\mathsf{E}[(\bar{A}^p-\hat{A}^p)\,(\bar{A}^q-\hat{A}^q)]\}=\frac{\underset{\circ}{\boldsymbol{C}}^{-1}\,[\boldsymbol{y}-\bar{\boldsymbol{X}}(\boldsymbol{y}),\boldsymbol{y}-\bar{\boldsymbol{X}}(\boldsymbol{y})]}{\eta-m}\,\underset{\circ}{C}^{k\,l}\,\underset{\sim}{\Theta}^p{}_k\,\underset{\sim}{\Theta}^q{}_l.\tag{56}$$

$$\mathsf{es}\{\mathsf{E}[(\bar{X}^i-\hat{X}^i)\,(\bar{X}^j-\hat{X}^j)]\}=\frac{\underset{\circ}{\boldsymbol{C}}^{-1}\,[\boldsymbol{y}-\bar{\boldsymbol{X}}(\boldsymbol{y}),\boldsymbol{y}-\bar{\boldsymbol{X}}(\boldsymbol{y})]}{\eta-m}\,\underset{\circ}{C}^{k\,l}\,\underset{\sim}{\Gamma}^i{}_k\,\underset{\sim}{\Gamma}^j{}_l.\tag{57}$$

دیده میشود طرفها یِ راست ِ رابطِها یِ بالا با ضرب-کردن ِ $\boldsymbol{C}$ در یک عدد تغییر نمیکنند. پس برا یِ تخمین ِ خطا یِ پارامترها و متغیرها مقدار ِ مطلق ِ خطا یِ سنجش لازم نیست.

4 پانوشتها

[1] Bayes

[2] Morris H. deGroot & Mark J. Schervish; “probability and statistics” 4th edition (Addison-Wesley, 2002)

[3] Lagrange

[4] Tom M. Apostol; “mathematical analysis” 2nd edition (Addison-Wesley, 1975) section 13.7