LIETUVOS MATEMATIKOS RINKINYS. LMD DARBAI
51 tomas, 2010, 420–425

ISSN 0132-2818
www.mii.lt/LMR/

EM algoritmas bendram Gauso modeliui su paslėptaisiais kintamaisiais

Marijus Radavičius[1,2]

[1] *Matematikos ir informatikos institutas*
Akademijos g. 4, LT-08663 Vilnius
[2] *Klaipėdos universitetas, Gamtos fakultetas*
Herkaus Manto g. 84, LT-92294 Klaipėda
E. paštas: mrad@ktl.mii.lt

Santrauka. Nagrinėjamas Gauso vektoriaus parametru vertinimo uždavinys, kai stebima jo (netiesinė) transformacija. Išvestos EM algoritmo, naudojamo didžiausio tikėtinumo įvertiniui apskaičiuoti, lygtys. Atskiru atveju, kai stebimas dviejų dydžių, aprašomų Gauso regresiniais modeliais, minimumas, gautos išreikštinės EM algoritmo formulės.

Raktiniai žodžiai: EM algoritmas, Gauso regresija, netiesinė transformacija, didžiausio tikėtinumo įvertinys.

Įvadas

Paslėptojo (*latent*) kintamojo modeliai yra viena iš produktyviausių pastaraisiais metais idėjų ne vien statistikos taikymuose, bet ir teorijoje (žr., pavyzdžiui, [4, 2] ir jų nuorodas). Gerai žinomi pavyzdžiai: (ekonometriniai) riboto endogeninio kintamojo modeliai (*Tobit, Heckit*) [2], *asimetrinis normalusis* (*skew normal*) skirstinys [1].

Vietoje to, kad iš karto užrašyti tiesiogiai sebimo proceso modelį, pradžioje sudaromas tiriamo proceso modelis, o tik po to detalizuojama, kaip jis siejasi su stebimu procesu. Praktiškai sudaromi du tarpusavyje susieti modeliai: *struktūrnis* (tiriamojo poceso modelis) ir *matavimo* (*measurement*), t.y. stebimo proceso modelis. Juos patogu traktuoti kaip dvi vieno proceso komponentes, iš kurių viena nėra stebima. Tokiu būdu, paslėptojo kintamojo modeliai papuola į kitą modelių ir su jais susijusių metodų klasę, vadinamą praleistųjų (*missing*) ar nepilnųjų (*incomplete*) duomenų modeliais. EM algritmas yra, matyt, žinomiausias metodas didžiausio tikėtinumo įvertiniams apskaičiuoi, kai duomenys turi praleistų reikšmių [3]. Ši iteratyvi optimizavimo procedūra išsiskiria iš bendrų optimizavimo algoritmų savo santykiniu stabilumu, mažesniu jautrumu pradinių reikšmių parinkimui.

Šiame darbe užrašomos EM algoritmo lygtys bendrai transformuotai (nepilnai stebimai) Gauso sistemai. Įdomu, kad maksimizavimo uždavinio M žingsnyje sprendimas priklauso tik nuo Gauso modelio parametrizacijos, t.y., nuo vidurkio ir kovariacijų matricos parametrinės formos, ir nepriklauso nuo Gauso sistemos transformacijos. Darbe nagrinėjamas pavyzdys, kai yra stebima tik mažesnioji dvimatės atsitiktinės sekos komponentė. Gaunamus tokiu būdu stebinius galima interpretuoti kaip duomenis apie dviejų firmų siūlomų analogiškų prekių pigesnės prekės kainą. Laikoma, kad atsitikinė seka tenkina dvimatės Gauso regresijos modelį. Pateikiamos išreikštinės EM žingsnio formulės nežinomiems modelio parametrams vertinti.

1 EM algoritmas transformuotom Gauso sistemom

Tegu $Y^* \sim N_n(M, V)$, $M \in \mathcal{M}$, $V \in \mathcal{V}$, kur $\mathcal{M}$ ir $\mathcal{V}$ yra duotos n-mačio atsitiktinio vektoriaus (a.v.) Y^* galimų vidurkių M ir galimų kovariacinių matricų V klasės (atitinkamai). Tegu h: $\mathbf{R}^n \to \mathbf{R}^m$, $m \le n$, yra žinoma funkcija, ir stebimas vektorius yra $Y := h(Y^*)$.

Asimetrinis normalusis skirstinys sutampa su $Y^*_{(1)} | \{Y^*_{(2)} > 0_k\}$ skirstiniu, t.y. sąlyginiu atsitiktinio vektoriaus $Y^*_{(1)}$ skirstiniu, kai žinoma, kad $Y^*_{(2)} > 0_k$ [1]. Čia $Y^*_{(1)} \in \mathbb{R}^k$ ir $Y^*_{(2)} \in \mathbb{R}^{n-k}$ yra vektoriaus Y^* komponentės, $k < n$, o nelygybė suprantama pakoordinačiui. Klasikinis *Heckman* modelis gaunamas kaip atskiras šios išraiškos atvejis, kai atsitiktinio vektoriaus Y^* komponenčių poros (y^*_{2i-1}, y^*_{2i}), $i = 1, \ldots, m$, $n = 2m$, yra tarpusavyje nepriklausomos ir $h_i(Y) = y^*_{2i-1} \cdot \mathbf{1}\{y^*_{2i} > 0\}$, $i = 1, \ldots, m$.

Trumpumo dėlei pažymėkime

$$\bar{M}(Y) := \mathbf{E}_{\bar{M},\bar{V}}[Y^* \mid Y] \tag{1}$$

sąlyginį a.v. Y^* vidurkį ir

$$\bar{V}(Y) := \mathbf{E}_{\bar{M},\bar{V}}\big[\big(Y^* - \bar{M}(Y)\big)\big(Y^* - \bar{M}(Y)\big)^\top \mid Y\big] \tag{2}$$

sąlyginę a.v. Y^* kovariacijų matricą, kai žinomas Y, o $Y^* \sim N_n(\bar{M}, \bar{V})$. Atsitiktinio vektoriaus $Y^* \sim N_n(M, V)$ logtikėtinumo funkcija

$$\begin{aligned} l(M, V) &= l\big(M, V \mid Y^*\big) \\ &:= -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln\big(\det(V)\big) - \frac{1}{2}\mathrm{tr}\big(V^{-1}\big(Y^* - M\big)\big(Y^* - M\big)^\top\big). \end{aligned} \tag{3}$$

Parametrų $M \in \mathcal{M}$ ir $V \in \mathcal{V}$ didžiausio tikėtinumo įvertiniui apskaičiuoti taikysime EM algoritmą [3].

E žingsnis. Pilnojo logtikėtinumo funkcijos $l(M, V \mid Y^*)$ sąlyginio vidurkio, kai žinomas Y apskaičiavimas:

$$\begin{aligned} Q(M, V) = Q(M, V \mid \bar{M}, \bar{V}; Y) &:= \mathbf{E}_{\bar{M},\bar{V}}\big[l\big(M, V \mid Y^*\big) \mid Y\big] \\ &= \mathrm{const} - \frac{1}{2}\ln\big(\det(V)\big) - \frac{1}{2}\mathrm{tr}\big(V^{-1}\bar{V}(Y)\big) \\ &\quad - \frac{1}{2}\big(\bar{M}(Y) - M\big)^\top V^{-1}\big(\bar{M}(Y) - M\big). \end{aligned} \tag{4}$$

Paskutinė lygybė gauta pasinaudojus (1)–(3), tapatybe $Y^* - M = Y^* - M(Y) + M(Y) - M$ ir tuo, kad sąlyginė kovariacija tarp $Y^* - \bar{M}(Y)$ ir $\bar{M}(Y) - M$ yra lygi $0_{n\times n}$. Atkreipkime dėmesį, kad $Q(M, V)$ yra atsitiktinis dydis, priklausantis nuo Y.

M žingsnis. Pilnojo tikėtinumo funkcijos sąlyginio vidurkio $Q(M, V \mid \bar{M}, \bar{V})$ maksimizavimas pagal M ir V. Iš (4) išplaukia, kad uždavinio

$$Q(M, V \mid \bar{M}, \bar{V}; Y) \longrightarrow \max_{M \in \mathcal{M}, V \in \mathcal{V}} \tag{5}$$

sprendinys $(M^*, V^*) = (M^*(Y), V^*(Y))$ yra sprendinys ir tokių 2-jų optimizavimo uždavinių:

$$R_{V^*}\big(M \mid \bar{M}(Y)\big) := \big(\bar{M}(Y) - M\big)^\top (V^*)^{-1}\big(\bar{M}(Y) - M\big) \longrightarrow \min_{M \in \mathcal{M}}, \tag{6}$$

$$d\big(V \mid W^*(Y)\big) := \ln\big(\det(V)\big) + \mathrm{tr}\big(V^{-1}W^*(Y)\big) \longrightarrow \min_{V \in \mathcal{V}}, \tag{7}$$

kur

$$\begin{aligned} W^* = W^*(Y) &:= \mathbf{E}_{\bar{M},\bar{V}}\big[\big(Y^* - M^*\big)\big(Y^* - M^*\big)^\top \mid Y\big] \\ &= \big(\bar{M}(Y) - M^*\big)\big(\bar{M}(Y) - M^*\big)^\top + \bar{V}(Y). \end{aligned} \tag{8}$$

Formulės (6) ir (7) nepatogios tuo, kad norint išspręsti (6) reikia jau žinoti V^*, o norint išspręsti (7) reikia jau žinoti M^*. Kadangi svorių matrica svertiniame mažiausių kvadratų metode nėra labai reikšminga (taikant paprastą MKm įvertiniai prie gana bendrų sąlygų gaunasi nors ir neefektyvūs, bet pagrįsti), tai vietoje duotame EM žingsnyje optimalaus kovariacijų matricos įvertinio V^* galima įstatyti ankstesnį („optimalų” praeitame EM žingsnyje) kovariacijų matricos įvertinį $\bar{V}$. Todėl uždavinį (6) siūloma keisti uždaviniu

$$R_{\bar{V}}\big(M \mid \bar{M}(Y)\big) \longrightarrow \min_{M \in \mathcal{M}}. \tag{9}$$

2 Pavyzdys: dviejų egzogeninių kintamųjų minimumas

Nagrinėkime atvejį, kai

$$Y^* = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} (\mathbf{X}_1)\alpha + \mathcal{E}_1 \\ (\mathbf{X}_2)\beta + \mathcal{E}_2 \end{pmatrix},$$

kur $Y_i = (y_i(t),\ t = 1, \ldots, N)^\top$, $\mathbf{X}_i$ yra $N \times k_i$ plano matrica, $i = 1, 2$, o $(\mathcal{E}_1, \mathcal{E}_2)$ sudaro Gauso sistemą, $\mathcal{E}_i = (\varepsilon_i(t),\ t = 1, \ldots, N)^\top$, $\mathcal{E}_i \sim N_n(0_N, \sigma_i^2 I_N)$, $i = 1, 2$, $\mathrm{cov}(\mathcal{E}_1, \mathcal{E}_2) = rI_N$, ir stebimas vektorius yra

$$Y := \min(Y_1, Y_2) =: h(Y^*) \tag{10}$$

(čia operacija „min” atliekama pakoordinačiui). Pastebėkime, kad vektorius $M \in \mathcal{M}$ turi pavidalą,

$$M = M(\alpha, \beta) = \begin{pmatrix} (\mathbf{X}_1)\alpha \\ (\mathbf{X}_2)\beta \end{pmatrix} \tag{11}$$

Pažymėję

$$D = D\big(\sigma_1^2, \sigma_2^2, r\big) = \begin{pmatrix} \sigma_1^2 & r \\ r & \sigma_2^2 \end{pmatrix} \tag{12}$$

matricas $V \in \mathcal{V}$ galime užrašyti tokia forma

$$V = V\big(\sigma_1^2, \sigma_2^2, r\big) = D \otimes I_N. \tag{13}$$

Vektorių M, apskaičiuotą pagal formulę (11) su $(\alpha, \beta) = (\bar{\alpha}, \bar{\beta})$, žymėsime $\bar{M}$. Analogiškai, $\bar{V}$ apskaičiuojamas pagal formulę (13) su $(\sigma_1^2, \sigma_2^2, r) = (\bar{\sigma}_1^2, \bar{\sigma}_2^2, \bar{r})$.

E žingsnis. Kadangi poros $\{(y_1(t), y_2(t)),\ t = 1, \ldots, N\}$ yra tarpusavyje nepriklausomos, pradžioje užtenka nagrinėti tik vieną porą $(y_1, y_2) := (y_1(1), y_2(1))$ (t.y. $N = 1$). Taigi $y = y(1) = \min(y_1, y_2) = Y$, $x_i = x_i(1)$, $i = 1, 2$. Trumpumo dėlei pažymėkime $\bar{\mu}_1 := \bar{\alpha}^\top x_1$, $\bar{\mu}_2 := \bar{\beta}^\top x_2$,

$$g_j\bigl(y \mid \mu, \sigma^2\bigr) := \int_y^\infty u^j \phi\bigl(u \mid \mu, \sigma^2\bigr)\, du; \tag{14}$$

čia $\varphi(u \mid \mu, \sigma^2) := \partial \Phi(u \mid \mu, \sigma^2)/\partial u$, o $\Phi(\cdot \mid \mu, \sigma^2)$ yra atsitiktinio dydžio $\eta \sim N(\mu, \sigma^2)$ pasiskirstymo funkcija.

Tegu $y_i \sim N(\bar{\mu}_i, \bar{\sigma}_i^2)$, $i = 1, 2$, ir $\operatorname{cov}(y_1, y_2) = \bar{r}$. Pagal Gauso sistemų (atsitiktinių vektorių) savybes sąlyginis skirstinys $y_i | \{y_j = y\}$ $(i \neq j)$ yra normalusis,

$$y_i | \{y_j = y\} \sim N\bigl(\bar{\mu}_{ij}(y), \bar{\sigma}_{ij}^2\bigr), \tag{15}$$

$$\bar{\mu}_{ij}(y) := \bar{\mu}_i + \frac{\bar{r}}{\bar{\sigma}_j^2}(y - \bar{\mu}_j), \qquad \bar{\sigma}_{ij}^2(y) := \bar{\sigma}_i^2 - \frac{\bar{r}^2}{\bar{\sigma}_j^2}. \tag{16}$$

Taigi,

$$\mathbf{E}\bigl[y_i^m \mathbf{1}\{y_i > y\} \bigm| y_j = y\bigr] = g_m\bigl(y \bigm| \bar{\mu}_{ij}(y), \bar{\sigma}_{ij}^2\bigr), \quad m = 0, 1, 2,\ i \neq j. \tag{17}$$

Atlikę tiesioginius skaičiavimus, iš (15)–(17) gauname, kad

$$\begin{aligned}
\bar{M}_1(y) &:= \mathbf{E}\bigl[y_1 \bigm| \min(y_1, y_2) = y\bigr] \\
&= y g_0\bigl(y \bigm| \bar{\mu}_{21}(y), \bar{\sigma}_{21}^2\bigr) \cdot \pi_1(y) + g_1\bigl(y \bigm| \bar{\mu}_{12}(y), \bar{\sigma}_{12}^2\bigr) \cdot \pi_2(y), &(18)\\
\bar{M}_2(y) &:= \mathbf{E}\bigl[y_2 \bigm| \min(y_1, y_2) = y\bigr] \\
&= y g_0\bigl(y \bigm| \bar{\mu}_{12}(y), \bar{\sigma}_{12}^2\bigr) \cdot \pi_2(y) + g_1\bigl(y \bigm| \bar{\mu}_{21}(y), \bar{\sigma}_{21}^2\bigr) \cdot \pi_1(y), &(19)\\
\bar{M}_{11}(y) &:= \mathbf{E}\bigl[y_1^2 \bigm| \min(y_1, y_2) = y\bigr] \\
&= y^2 g_0\bigl(y \bigm| \bar{\mu}_{21}(y), \bar{\sigma}_{21}^2\bigr) \cdot \pi_1(y) + g_2\bigl(y \bigm| \bar{\mu}_{12}(y), \bar{\sigma}_{12}^2\bigr) \cdot \pi_2(y), &(20)\\
\bar{M}_{22}(y) &:= \mathbf{E}\bigl[y_2^2 \bigm| \min(y_1, y_2) = y\bigr] \\
&= y^2 g_0\bigl(y \bigm| \bar{\mu}_{12}(y), \bar{\sigma}_{12}^2\bigr) \cdot \pi_2(y) + g_2\bigl(y \bigm| \bar{\mu}_{21}(y), \bar{\sigma}_{21}^2\bigr) \cdot \pi_1(y), &(21)\\
\bar{M}_{12}(y) &:= \mathbf{E}\bigl[y_1 y_2 \bigm| \min(y_1, y_2) = y\bigr] \\
&= y\bigl(g_1\bigl(y \bigm| \bar{\mu}_{21}(y), \bar{\sigma}_{21}^2\bigr) \cdot \pi_1(y) + g_1\bigl(y \bigm| \bar{\mu}_{12}(y), \bar{\sigma}_{12}^2\bigr) \cdot \pi_2(y)\bigr). &(22)
\end{aligned}$$

Kadangi kovariacijų matricos $\bar{V}_1(y) := \bar{V}(y)$ elementai

$$(\bar{V}_1)_{ij}(y) := \mathrm{COV}\bigl[y_1, y_2 \bigm| \min(y_1, y_2) = y\bigr] = \bar{M}_{ij}(y) - \bar{M}_i(y)\bar{M}_j(y), \tag{23}$$

tai ją suskaičiuojame įstatę išraiškas (18), (19) ir (20)–(22).

Nagrinėkime bendrą atvejį, $N \geqslant 1$. Pradėkime nuo pažymėjimų:

$$\bar{M}_i(Y) := \bigl(\bar{M}_i\bigl(y(1)\bigr), \bar{M}_i\bigl(y(2)\bigr), \ldots, \bar{M}_i\bigl(y(N)\bigr)\bigr)^\top, \quad i = 1, 2, \tag{24}$$

$$\bar{V}_{ij}(Y) := \operatorname{diag}\bigl\{(\bar{V}_1)_{ij}\bigl(y(1)\bigr), (\bar{V}_1)_{ij}\bigl(y(2)\bigr), \ldots, (\bar{V}_1)_{ij}\bigl(y(N)\bigr)\bigr\}, \quad i, j = 1, 2, \tag{25}$$

kur $\bar{M}_i(y)$, $i = 1, 2$, yra nusakomas formulėmis (18) ir (19), o $(\bar{V}_1)_{ij}(y)$, $i, j = 1, 2$, yra nusakomas formulėmis (23) ir (20)–(22). Tuomet

$$\bar{M}(Y) = \begin{pmatrix} \bar{M}_1 \\ \bar{M}_2 \end{pmatrix}, \qquad \bar{V}(Y) = \begin{pmatrix} \bar{V}_{11} & \bar{V}_{12} \\ \bar{V}_{21} & \bar{V}_{22} \end{pmatrix}. \tag{26}$$

Pasinaudoję matricine užrašymo forma gausime išreikštinį rizikos funkcijų $R_{\bar{V}}(M \mid \bar{M}(Y))$ ir $d(V \mid \bar{W}(Y))$, nusakytų atitinkamai formulėmis (6) ir (7), pavidalą atžvilgiu nežinomų parametrų $\theta = (\alpha, \beta, \sigma_1^2, \sigma_2^2, r)$. Tuo pačiu bus nusakytas ir funkcijos $\bar{Q}(\theta) := Q(\theta, \bar{\theta})$ išreikštinis pavidalas, žr. formulę (4) bei (6) ir (7). Nuo parametrų $\bar{\theta} = (\bar{\alpha}, \bar{\beta}, \bar{\sigma}_1^2, \bar{\sigma}_2^2, \bar{r})$ priklauso tik vektoriai $\bar{M}, \bar{M}(Y)$ ir matricos $\bar{V}, \bar{V}(Y), \bar{W} = \bar{W}(Y)$, kur $\bar{W}(Y)$ apibrėžta formule (8) su $M^* = \bar{M}$.

Rizikos funkcijos $R_{\bar{V}}(M \mid \bar{M}(Y))$ apibrėžime (6) įstačius vektoriaus M išraišką (11) ir matricos $\bar{V} = V(\bar{\sigma}_1^2, \bar{\sigma}_2^2, \bar{r})$ išraišką, galima užrašyti

$$\begin{aligned} R_{\bar{V}}\big(M \mid \bar{M}(Y)\big) = {} & \frac{1}{\det(\bar{D})}\big(\bar{\sigma}_2^2\big|\bar{M}_1(Y) - \mathbf{X}_1\alpha\big|^2 + \bar{\sigma}_1^2\big|\bar{M}_2(Y) - \mathbf{X}_2\beta\big|^2\big) \\ & - \frac{2\bar{r}}{\det(\bar{D})}\big(\bar{M}_1(Y) - \mathbf{X}_1\alpha\big)^\top\big(\bar{M}_2(Y) - \mathbf{X}_2\beta\big) =: \bar{R}(\alpha, \beta). \end{aligned} \tag{27}$$

Kadangi $V^{-1}W = (D^{-1} \otimes I_N)W$, tai $d(V \mid W^*) = \bar{d}(\sigma_1^2, \sigma_2^2, r)$,

$$\bar{d}\big(\sigma_1^2, \sigma_2^2, r\big) := n\ln(\det D) + \frac{1}{\det D}\big(\sigma_2^2 {w^*}_{11} - 2r{w^*}_{21} + \sigma_1^2 {w^*}_{22}\big), \tag{28}$$

$$\det D = \sigma_1^2\sigma_2^2 - r^2, \tag{29}$$

$${w^*}_{jj} := \operatorname{tr}\big(W_{jj}^*\big) = \big|M_j^* - \bar{M}_j(Y)\big|^2 + \operatorname{tr}(\bar{V}_{jj}), \quad j = 1, 2, \tag{30}$$

$${w^*}_{12} := \operatorname{tr}\big(W_{12}^*\big) = \big(M_1^* - \bar{M}_1(Y)\big)^\top\big(M_2^* - \bar{M}_2(Y)\big) + \operatorname{tr}(\bar{V}_{12}), \tag{31}$$

$$M_j^* := M_j\big(\alpha^*, \beta^*\big), \quad j = 1, 2.$$

M žingsnis. Rizikos funkcijos $\bar{R}(\alpha, \beta)$, apibrėžtos išraiškoje (27), minimizavimo uždavinys faktiškai yra svertinis mažiausių kvadratų metodas, todėl

$$\begin{pmatrix} \alpha^* \\ \beta^* \end{pmatrix} = \begin{pmatrix} \bar{\sigma}_2^2\mathbf{X}_1^\top\mathbf{X}_1 & -\bar{r}\mathbf{X}_1^\top\mathbf{X}_2 \\ -\bar{r}\mathbf{X}_1^\top\mathbf{X}_2 & \bar{\sigma}_1^2\mathbf{X}_2^\top\mathbf{X}_2 \end{pmatrix}^{-1} \begin{pmatrix} \bar{\sigma}_2^2\mathbf{X}_1^\top\bar{M}_1(Y) - \bar{r}\mathbf{X}_1^\top\bar{M}_2(Y) \\ -\bar{r}\mathbf{X}_2^\top\bar{M}_1(Y) + \bar{\sigma}_1^2\mathbf{X}_2^\top\bar{M}_2(Y) \end{pmatrix}. \tag{32}$$

Dabar spręsime funkcijos $\bar{d}(\sigma_1^2, \sigma_2^2, r)$, nusakytos formulėmis (28) ir (29), minimizavimo uždavinį. Prilyginę šios funkcijos išvestines nuliui, randame ieškomus dydžius ${\sigma^*}_j^2 = N^{-1}w_{jj}^*$, $j = 1, 2$, ir $r^* = N^{-1}w_{12}^*$. bei

$$D^* := D({\sigma^*}_1^2, {\sigma^*}_2^2, r^*) = \frac{1}{N}\begin{pmatrix} w_{11}^* & w_{12}^* \\ w_{21}^* & w_{22}^* \end{pmatrix}. \tag{33}$$

Apibendrinkime gautas formules.

EM algoritmas. Tegu duotos pradinės nežinomų vidurkio parametrų (α, β) ir kovariacijų matricos parametrų $(\sigma_1^2, \sigma_2^2, r)$ reikšmės $(\bar{\alpha}, \bar{\beta})$ ir $(\bar{\sigma}_1^2, \bar{\sigma}_2^2, \bar{r})$ (atitinkamai). Naujos (patikslintos) nežinomų parametrų (α, β) ir $(\sigma_1^2, \sigma_2^2, r)$. reikšmės (α^*, β^*) ir $({\sigma^*}_1^2, {\sigma^*}_2^2, r^*)$ apskaičiuojamos taip:

1 Žingsnis. Naudojant formules (18)–(26) ir (14) apskaičiuojame $\bar{M}(Y)$ ir $\bar{V}(Y)$, o taip pat $\bar{W}(Y)$ pagal formulę (8) su M^* pakeistu į $\bar{M}$.

2 Žingsnis. Remiantis formule (32) randame (α^*, β^*).

3 Žingsnis. Naująją matricos D (žr. (12)) reikšmę D^*, o tuo pačiu ir naujas reikšmes $({\sigma^*}_1^2, {\sigma^*}_2^2, r^*)$, apskaičiuojame iš formulių (30), (31) ir (33).

Įdomu būtų išnagrinėti atvejį, kai stebima keleto, sakykim $m, m > 2$, firmų siūlomų prekių pigiausios prekės kaina. Praktiniu ir teoriniu požiūriu patrauklus modelis gaunasi, kai atitinkamų m Gauso regresijos modelių regresijos koeficientai yra tarpusavyje nepriklausomi ir vienodai pasiskirstę atsitiktiniai vektoriai, o $m \to \infty$.

Literatūra

[1] A.B. Arellano-Vale, M.D. Branco and M.G. Genton. A unified view on skewed distributions arising from selections. *Canadian Journal of Statistics*, **34**(4):581–601, 2006.

[2] A.C. Cameron and P.K. Trivedi. *Microeconometrics: Methods and Applications*. Cambridge University Press, Cambridge, 2005.

[3] A.P. Dempster, N.M. Laird and D.B. Rubin. Maximum likelihood from incomplete data via the em algorithm (with discussion). *J. Roy. Statist. Soc., Ser. B*, **39**:1–38, 1977.

[4] M.D. Sammel, L.M. Ryan and J.M. Legler. Latent variable models for mixed discrete and continuous outcomes. *J. Roy. Statist. Soc., Ser. B*, **59**:667–678, 1997.

SUMMARY

The EM algorithm for general Gaussian model with latent variables
M. Radavičius

The problem of estimating parameters of Gaussian vector when only its (nonlinear) transformation is observed is addressed. The EM algorithm equations to calculate maximum likelihood estimator are derived. In particular, closed-form formulas of the EM algorithm are derived in the case when only the minimum of two endogenous variables satisfying Gaussian regression model is observed.

Keywords: EM algorithm, Gaussian regression, nonlinear transform, maximum likelihood estimator.