Solution of the problem 5.19 (a, b,c)

Victor Mergel

March 11, 2002

part a.

We will use the notations $\vec{x^c}, \vec{y^c}$ instead of $\vec{x_1}, \vec{x_2}$ to present the complete data and $\vec{x}, \vec{y}$ to present observed data.

The density of two dimensional normal distribution is as follow:

$$\varphi(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left\{\frac{x_1^2}{\sigma_1^2} - 2\rho\frac{x_1x_2}{\sigma_1\sigma_2} + \frac{x_2^2}{\sigma_2^2}\right\}\right\};$$

and so the complete-data likelihood:

$$L^c(\sigma_1, \sigma_2, \rho|\vec{x^c}, \vec{y^c}) = \frac{1}{(2\pi\sigma_1\sigma_2)^{12}(1-\rho^2)^6} \exp\left\{-\frac{1}{2(1-\rho^2)} \times \left(\frac{20 + x_9^2 + x_{10}^2 + x_{11}^2 + x_{12}^2}{\sigma_1^2} - 2\rho\frac{2y_5 + 2y_6 - 2y_7 - 2y_8 + 2x_9 + 2x_{10} - 2x_{11} - 2x_{12}}{\sigma_1\sigma_2} + \frac{20 + y_5^2 + y_6^2 + y_7^2 + y_8^2}{\sigma_2^2}\right)\right\} = \frac{1}{(2\pi\sigma_1\sigma_2)^{12}(1-\rho^2)^6} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{20}{\sigma_1^2} + \frac{20}{\sigma_2^2}\right)\right\} \times \prod_{i=9}^{12} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x_i^2}{\sigma_1^2} - 4\rho\frac{x_i(10.5-i)}{\sigma_1\sigma_2|10.5-i|}\right)\right\} \prod_{i=5}^{8} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{y_i^2}{\sigma_2^2} - 4\rho\frac{y_i(6.5-i)}{\sigma_1\sigma_2|6.5-i|}\right)\right\}.$$

Then, incomplete-data likelihood:

$$L(\sigma_1, \sigma_2, \rho|\vec{x}, \vec{y}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} L^c(\sigma_1, \sigma_2, \rho|\vec{x^c}, \vec{y^c})\, dx_9dx_{10}dx_{11}dx_{12} dy_5dy_6dy_7dy_8 = \frac{1}{(2\pi\sigma_1\sigma_2)^{12}(1-\rho^2)^6} \exp\left\{-\frac{10}{1-\rho^2}\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)\right\} \times$$

$$\left(\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2}-4\rho\frac{x}{\sigma_1\sigma_2}\right)\right\} dx\right)^2 \left(\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2}+4\rho\frac{x}{\sigma_1\sigma_2}\right)\right\} dx\right)^2 \times$$

$$\left(\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{y^2}{\sigma_2^2}-4\rho\frac{y}{\sigma_1\sigma_2}\right)\right\} dy\right)^2 \left(\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{y^2}{\sigma_2^2}+4\rho\frac{y}{\sigma_1\sigma_2}\right)\right\} dy\right)^2.$$

To evaluate this we use:

$$\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2\sigma^2}(y^2 \pm 2my)\right\} dy = \sqrt{2\pi}\sigma \exp\left\{\frac{m^2}{2\sigma^2}\right\}.$$

And so,

$$L(\sigma_1,\sigma_2,\rho|\vec{x},\vec{y}) = \frac{1}{(2\pi\sigma_1\sigma_2)^{12}(1-\rho^2)^6}\exp\left\{-\frac{10}{1-\rho^2}\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right)\right\}\left(\sqrt{2\pi}\sqrt{1-\rho^2}\sigma_1 \times \right.$$
$$\left.\exp\left\{\frac{2\rho^2}{(1-\rho^2)\sigma_2^2}\right\}\right)^4 \left(\sqrt{2\pi}\sqrt{1-\rho^2}\sigma_2 \exp\left\{\frac{2\rho^2}{(1-\rho^2)\sigma_1^2}\right\}\right)^4 =$$
$$\frac{1}{(2\pi\sigma_1\sigma_2)^8(1-\rho^2)^2}\exp\left\{-2\left(4+\frac{1}{(1-\rho^2)}\right)(\sigma_1^{-2}+\sigma_2^{-2})\right\}.$$

$$\log L(\sigma_1,\sigma_2,\rho|\vec{x},\vec{y}) = -8\log(2\pi) - 8\log\sigma_1 - 8\log\sigma_2 - 2\left(4+\frac{1}{1-\rho^2}\right)\frac{1}{\sigma_1^2} -$$
$$2\left(4+\frac{1}{1-\rho^2}\right)\frac{1}{\sigma_2^2} - 2\log(1-\rho^2)$$

Therefore

$$\frac{\partial \log L}{\partial \rho} = -\frac{4\rho}{(1-\rho^2)^2}(\sigma_1^{-2}+\sigma_2^{-2}) + \frac{4\rho}{1-\rho^2} = 0;$$
$$\frac{\partial \log L}{\partial \sigma_1} = -\frac{8}{\sigma_1} + \left(4+\frac{1}{1-\rho^2}\right)\frac{4}{\sigma_1^3} = 0;$$
$$\frac{\partial \log L}{\partial \sigma_2} = -\frac{8}{\sigma_2} + \left(4+\frac{1}{1-\rho^2}\right)\frac{4}{\sigma_2^3} = 0.$$

Solutions of this three equations are: $\rho = 0, \sigma_1^2 = \sigma_2^2 = 2.5$ and $\rho = \pm 0.5, \sigma_1^2 = \sigma_2^2 = 2.(6)$. To decide which one correspond to maximum and which one correspond to saddlepoint we need second order derivatives of $\log L$:

$$\frac{\partial^2 \log L}{\partial \rho^2} = -4\frac{1+2\rho^2-3\rho^4}{(1-\rho^2)^4}(\sigma_1^{-2}+\sigma_2^{-2}) + \frac{4(1+\rho^2)}{(1-\rho^2)^2};$$

$$\frac{\partial^2 \log L}{\partial \sigma_1^2} = 8\sigma_1^{-2} - 12\left(4+\frac{1}{1-\rho^2}\right)\sigma_1^{-4};$$

$$\frac{\partial^2 \log L}{\partial \sigma_2^2} = 8\sigma_2^{-2} - 12\left(4 + \frac{1}{1-\rho^2}\right)\sigma_2^{-4};$$

$$\frac{\partial^2 \log L}{\partial \rho \partial \sigma_1} = \frac{8\rho}{(1-\rho^2)^2}\sigma_1^{-3}; \quad \frac{\partial^2 \log L}{\partial \rho \partial \sigma_2} = \frac{8\rho}{(1-\rho^2)^2}\sigma_2^{-3}; \quad \frac{\partial^2 \log L}{\partial \sigma_1 \partial \sigma_2} = 0.$$

To proove that solutions $\rho = \pm 0.5, \sigma_1^2 = \sigma_2^2 = 2.(6)$ are MLE we need the following quadratic form to be negative definite:

$$\begin{pmatrix} -\frac{32}{9} & \alpha\sqrt{\frac{8}{3}} & \beta\sqrt{\frac{8}{3}} \\ \alpha\sqrt{\frac{8}{3}} & -6 & 0 \\ \beta\sqrt{\frac{8}{3}} & 0 & -6 \end{pmatrix}$$

where α, β or +1 or -1.

Since $-\frac{32}{9} < 0$, $\begin{vmatrix} -\frac{32}{9} & \alpha\sqrt{\frac{8}{3}} \\ \alpha\sqrt{\frac{8}{3}} & -6 \end{vmatrix} = 18.(6) > 0$, $\begin{vmatrix} -\frac{32}{9} & \alpha\sqrt{\frac{8}{3}} & \beta\sqrt{\frac{8}{3}} \\ \alpha\sqrt{\frac{8}{3}} & -6 & 0 \\ \beta\sqrt{\frac{8}{3}} & 0 & -6 \end{vmatrix} = -96 < 0$

this quadratic form is negatively definite (Silvester criterion) and thus solutions $\rho = \pm 0.5, \sigma_1^2 = \sigma_2^2 = 2.(6)$ provide local maxima. To prove that points $\rho = 0$, $\sigma_1^2 = \sigma_2^2 = 2.5$ are saddlepoints, just note that $\frac{\partial^2 \log L}{\partial \rho^2}|_{\rho=0,\sigma_1^2=\sigma_2^2=2.5} = 0.8$ and $\frac{\partial^2 \log L}{\partial \sigma_1^2}|_{\rho=0,\sigma_1^2=\sigma_2^2=2.5} = -6.4$ and so we have saddle points. To prove that points $\rho = \pm 0.5, \sigma_1^2 = \sigma_2^2 = 2.(6)$ provide global maximum to likelihood function, note that since we did not use properties of ρ and since two other extrema are saddlepoints we should compare likelihood at this points only with likelihood on infinity (which is zero). So, using nonnegativity of likelihood and its symmetry in σ_1, σ_2 we will have that points $\rho = \pm 0.5, \sigma_1^2 = \sigma_2^2 = 2.(6)$ provide global maxima to likelihood.

part b. Complete likelihood is as follow:

$$L^c(\sigma_1, \sigma_2, \rho | \vec{x^c}, \vec{y^c}) = \frac{1}{(2\pi\sigma_1\sigma_2)^{12}(1-\rho^2)^6} \exp\left\{ -\frac{1}{2(1-\rho^2)} \times \left(20(\sigma_1^{-2} + \sigma_2^{-2}) + \sum_{i=9}^{12}\left(\frac{x_i^2}{\sigma_1^2} - 4\rho\frac{\alpha_i x_i}{\sigma_1\sigma_2}\right) + \sum_{i=5}^{8}\left(\frac{y_i^2}{\sigma_2^2} - 4\rho\frac{\alpha_i y_i}{\sigma_1\sigma_2}\right)\right)\right\},$$

where $\alpha_5 = \alpha_6 = \alpha_9 = \alpha_{10} = 1$; $\alpha_7 = \alpha_8 = \alpha_{11} = \alpha_{12} = -1$. Then

$$\log L^c(\sigma_1, \sigma_2, \rho | \vec{x^c}, \vec{y^c}) = -12\log(2\pi\sigma_1\sigma_2) - 6\log(1-\rho^2) - \frac{1}{2(1-\rho^2)} \times \left(20(\sigma_1^{-2} + \sigma_2^{-2}) + \sum_{i=9}^{12}\left(\frac{x_i^2}{\sigma_1^2} - 4\rho\frac{\alpha_i x_i}{\sigma_1\sigma_2}\right) + \sum_{i=5}^{8}\left(\frac{y_i^2}{\sigma_2^2} - 4\rho\frac{\alpha_i y_i}{\sigma_1\sigma_2}\right)\right)\Bigg\},$$

and

$$Q(\rho|\rho_0) = \int \ldots \int \log L^c \prod_{i=5}^{8} f(y_i) \prod_{i=9}^{12} f(x_i) \prod_{i=5}^{8} dy_i \prod_{i=9}^{12} dx_i,$$

where

$$f(x_i) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{-\frac{x_i^2}{2\sigma_1^2}\right\}, f(y_i) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left\{-\frac{y_i^2}{2\sigma_2^2}\right\}.$$

And so

$$Q(\rho|\rho_0) = -6\log(1-\rho^2) - \frac{1}{2(1-\rho^2)}\left(20(\sigma_1^{-2}+\sigma_2^{-2}) + 4 + 4\right),$$

after differentiating we'll get:

$$\frac{\partial Q}{\partial \rho} = \frac{12\rho}{1-\rho^2} - \frac{2\rho}{(1-\rho^2)^2}\left(10(\sigma_1^{-2}+\sigma_2^{-2}) + 8\right) \;\Rightarrow\; \rho = 0$$

and

$$\frac{\partial^2 Q}{\partial \rho^2} = \frac{12(1+\rho^2)}{(1-\rho^2)^2} - \frac{1+3\rho^2}{(1-\rho^2)^3}\left(20(\sigma_1^{-2}+\sigma_2^{-2}) + 16\right).$$

Then

$\frac{\partial^2 Q}{\partial \rho^2}|_{\rho=0} = -20(\sigma_1^{-2}+\sigma_2^{-2}) - 4 < 0 \;\Rightarrow\; \rho = 0$ - maxima.

part c. If $\rho_0 > 0$ and $\rho_n \to 0 \Rightarrow L(0|\vec{x},\vec{y}) > L(\rho_0 > 0|\vec{x},\vec{y})$.

In our case we have:

$$L(0|\vec{x},\vec{y}) = \frac{1}{(2\pi\sigma_1\sigma_2)^8} \exp\left\{-10(\sigma_1^{-2}+\sigma_2^{-2})\right\};$$

$$L(\rho_0|\vec{x},\vec{y}) = \frac{1}{(2\pi\sigma_1\sigma_2)^8(1-\rho_0^2)^2} \exp\left\{-(8 + \frac{2}{1-\rho^2}(\sigma_1^{-2}+\sigma_2^{-2})\right\}.$$

So

$$g(\rho_0) = \frac{L(0|\vec{x},\vec{y})}{L(\rho_0|\vec{x},\vec{y})} = (1-\rho_0^2)^2 \exp\left\{-(2 - \frac{2}{1-\rho_0^2}(\sigma_1^{-2}+\sigma_2^{-2})\right\}.$$

Easy to check that $\rho_0 = 0$ provide global maxima to $g(*)$ and $g(0) = 1$. So $L(0|\vec{x},\vec{y}) \leq L(\rho_0 > 0|\vec{x},\vec{y})$ and $\rho_n \to \rho^* \neq 0 \;\Rightarrow\; \rho_n$ converge to a maximum.