MATH 668 Homework 3 Solutions

1. Using Theorem 5.2.2 with $\boldsymbol{x} = \boldsymbol{y}$, $\mathbf{A} = \boldsymbol{a}$, and $\boldsymbol{y} = y_n$ we see that

$E(\boldsymbol{a}^\top \boldsymbol{y} y_n) = E(\boldsymbol{y}^\top \boldsymbol{a} y_n) = E(\boldsymbol{x}^\top \mathbf{A}\boldsymbol{y}) = \text{tr}(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}^\top \mathbf{A}\boldsymbol{\mu} = \text{tr}(\mathbf{A}\boldsymbol{\Sigma}_{yx}) + \boldsymbol{\mu}_x^\top \mathbf{A}\boldsymbol{\mu}_y =$

$\text{tr}(\boldsymbol{a}\text{cov}(y_n, \boldsymbol{y})) + \boldsymbol{\mu}^\top \boldsymbol{a} E(y_n) = \text{tr}(\text{cov}(y_n, \boldsymbol{y})\boldsymbol{a}) + \boldsymbol{a}^\top \boldsymbol{\mu} E(y_n) = \text{cov}(y_n, \boldsymbol{y})\boldsymbol{a} + \boldsymbol{a}^\top \boldsymbol{\mu} E(y_n)$.

2. (a) $\mathbf{B}^2 = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 10 & 5 \\ 5 & 5 \end{pmatrix} = 5\boldsymbol{\Sigma}$; Note that $\boldsymbol{\Sigma}^{1/2} = \frac{1}{\sqrt{5}}\mathbf{B}$.

   (b) If $\boldsymbol{y} \sim N_2(\mathbf{0}, \boldsymbol{\Sigma})$, then Theorem 4.4.1 implies that $\boldsymbol{z} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{y} \sim N_2(\mathbf{0}, \mathbf{I})$ since $\mathbf{E}(\boldsymbol{\Sigma}^{-1/2}\boldsymbol{y}) = \boldsymbol{\Sigma}^{-1/2}E(\boldsymbol{y}) = \boldsymbol{\Sigma}^{-1/2}\mathbf{0} = \mathbf{0}$ and $\text{cov}(\boldsymbol{\Sigma}^{-1/2}\boldsymbol{y}) = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{-1/2} = (\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}^{1/2})(\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{-1/2}) = \mathbf{I}$.

Note that $\mathbf{A}$ is idempotent since $\mathbf{AA} = \mathbf{A}$ and it has rank 1 since $\text{tr}(\mathbf{A}) = .5 + .5 = 1$. Then Theorem 5.5.1 implies that $\boldsymbol{z}^\top \mathbf{A}\boldsymbol{z} \sim \chi^2(1)$. So, since $\boldsymbol{\Sigma}^{-1/2} = (\frac{1}{\sqrt{5}}\mathbf{B})^{-1} = \sqrt{5}\mathbf{B}^{-1}$ we see that

$\boldsymbol{z}^\top \mathbf{A}\boldsymbol{z} = \boldsymbol{y}^\top \boldsymbol{\Sigma}^{-1/2}\mathbf{A}\boldsymbol{\Sigma}^{-1/2}\boldsymbol{y} = \boldsymbol{y}^\top(\sqrt{5}\mathbf{B}^{-1})\mathbf{A}(\sqrt{5}\mathbf{B}^{-1})\boldsymbol{y} = 5\boldsymbol{y}^\top \mathbf{B}^{-1}\mathbf{A}\mathbf{B}^{-1}\boldsymbol{y} \sim \chi^2(1)$.

   (c) By Theorem 5.6.1, we see that $\boldsymbol{j}^\top \mathbf{B}^{-1}\boldsymbol{y}$ and $\boldsymbol{y}^\top \mathbf{B}^{-1}\mathbf{A}\mathbf{B}^{-1}\boldsymbol{y}$ are independent since

$\boldsymbol{j}^\top \mathbf{B}^{-1}\boldsymbol{\Sigma}\mathbf{B}^{-1}\mathbf{A}\mathbf{B}^{-1} = \boldsymbol{j}^\top \mathbf{B}^{-1}(\frac{1}{5}\mathbf{B}^2)\mathbf{B}^{-1}\mathbf{A}\mathbf{B}^{-1} = \frac{1}{5}\boldsymbol{j}^\top(\mathbf{B}^{-1}\mathbf{B}\mathbf{B}\mathbf{B}^{-1})\mathbf{A}\mathbf{B}^{-1}$

$= \frac{1}{5}\boldsymbol{j}^\top(\mathbf{I})\mathbf{A}\mathbf{B}^{-1} = \frac{1}{5}(\boldsymbol{j}^\top \mathbf{A})\mathbf{B}^{-1} = \frac{1}{5}\mathbf{0}^\top \mathbf{B}^{-1} = \mathbf{O}_{1\times 2}$.

   (d) By Theorem 4.4.1, $\boldsymbol{j}^\top \mathbf{B}^{-1}\boldsymbol{y} \sim N(0, \frac{2}{5})$ since $E(\boldsymbol{j}^\top \mathbf{B}^{-1}\boldsymbol{y}) = \boldsymbol{j}^\top \mathbf{B}^{-1}\mathbf{0} = 0$ and $\text{var}(\boldsymbol{j}^\top \mathbf{B}^{-1}\boldsymbol{y}) = \boldsymbol{j}^\top \mathbf{B}^{-1}\boldsymbol{\Sigma}\mathbf{B}^{-1}\boldsymbol{j} = \boldsymbol{j}^\top \mathbf{B}^{-1}(\frac{1}{5}\mathbf{B}^2)\mathbf{B}^{-1}\boldsymbol{j} = \frac{1}{5}\boldsymbol{j}^\top(\mathbf{B}^{-1}\mathbf{B}^2\mathbf{B}^{-1})\boldsymbol{j} = \frac{1}{5}\boldsymbol{j}^\top \boldsymbol{j} = \frac{1}{5}(2) = \frac{2}{5}$. So $\sqrt{\frac{5}{2}}\boldsymbol{j}^\top \mathbf{B}^{-1}\boldsymbol{y} \sim N(0, 1)$.

The random variables $\sqrt{\frac{5}{2}}\boldsymbol{j}^\top \mathbf{B}^{-1}\boldsymbol{y} \sim N(0, 1)$ and $5\boldsymbol{y}^\top \mathbf{B}^{-1}\mathbf{A}\mathbf{B}^{-1}\boldsymbol{y} \sim \chi^2(1)$ are independent so $\dfrac{\sqrt{\frac{5}{2}}\boldsymbol{j}^\top \mathbf{B}^{-1}\boldsymbol{y}}{\sqrt{5\boldsymbol{y}^\top \mathbf{B}^{-1}\mathbf{A}\mathbf{B}^{-1}\boldsymbol{y}/1}} = \left(\dfrac{1}{\sqrt{2}}\right)\dfrac{\boldsymbol{j}^\top \mathbf{B}^{-1}\boldsymbol{y}}{\sqrt{\boldsymbol{y}^\top \mathbf{B}^{-1}\mathbf{A}\mathbf{B}^{-1}\boldsymbol{y}}} \sim t(1)$ by Definition 5.4.2.

3. (a) Differentiating $\tilde{Q}$ with respect to b_0 and b_1, we obtain

$\dfrac{\partial \tilde{Q}}{\partial b_0} = -2\displaystyle\sum_{i=1}^n (y_i - b_0 - b_1 x_i)$ and

$\dfrac{\partial \tilde{Q}}{\partial b_1} = -2\displaystyle\sum_{i=1}^n x_i(y_i - b_0 - b_1 x_i) + 2\lambda b_1$.

Setting both equations to 0, we denote the solutions as $\hat{\beta}_{0,\lambda}$ and $\hat{\beta}_{1,\lambda}$. From the equation for the partial with respect to b_0, we have

$-2\sum_{i=1}^n (y_i - \hat{\beta}_{0,\lambda} - \hat{\beta}_{1,\lambda}x_i) = 0 \implies n\hat{\beta}_{0,\lambda} + \hat{\beta}_{1,\lambda}\sum_{i=1}^n x_i = \sum_{i=1}^n y_i \implies \hat{\beta}_{0,\lambda} = \bar{y} - \hat{\beta}_{1,\lambda}\bar{x}$. Substituting this into the equation for the partial with respect to b_1, we obtain

$-2\sum_{i=1}^n x_i(y_i - \hat{\beta}_{0,\lambda} - \hat{\beta}_{1,\lambda}x_i) + 2\lambda\hat{\beta}_{1,\lambda} = 0 \implies \sum_{i=1}^n x_i y_i = \sum_{i=1}^n x_i\hat{\beta}_{0,\lambda} + \hat{\beta}_{1,\lambda}(\sum_{i=1}^n x_i^2 + \lambda)$

$\implies \sum_{i=1}^n x_i y_i = \sum_{i=1}^n x_i(\bar{y} - \hat{\beta}_{1,\lambda}\bar{x}) + \hat{\beta}_{1,\lambda}(\sum_{i=1}^n x_i^2 + \lambda) \implies \sum_{i=1}^n x_i y_i = n\bar{x}\bar{y} - \hat{\beta}_{1,\lambda}\bar{x}^2 + \hat{\beta}_{1,\lambda}(\sum_{i=1}^n x_i^2 + \lambda)$

$\implies \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y} = \hat{\beta}_{1,\lambda}(\sum_{i=1}^n x_i^2 - \bar{x}^2 + \lambda) \implies (n-1)s_{x,y} = \hat{\beta}_{1,\lambda}((n-1)s_x^2 + \lambda) \implies$ $\hat{\beta}_{1,\lambda} = \dfrac{(n-1)s_{x,y}}{(n-1)s_x^2 + \lambda}$ if $s_x > 0$ or $\lambda > 0$.

$\tilde{Q}$ is a convex function since its matrix of second partial derivatives is nonnegative definite; this is seen since $\dfrac{\partial^2 \tilde{Q}}{\partial b_0^2} = 2n > 0$ and

$\left| \frac{\partial}{\partial \boldsymbol{b}} \frac{\partial \tilde{Q}}{\partial \boldsymbol{b}^\top} \right| = \begin{vmatrix} 2n & 2n\bar{x} \\ 2n\bar{x} & 2\sum_{i=1}^n x_i^2 + 2\lambda \end{vmatrix} = 4n(\sum_{i=1}^n x_i^2 + \lambda) - 4n^2\bar{x}_2 = 4n(\sum_{i=1}^n x_i^2 + \lambda - n\bar{x}^2) = 4n(\sum_{i=1}^n x_i^2 - n\bar{x}^2 + \lambda) = 4n(n-1)s_x^2 + 4n\lambda \geq 0$. (If $s_x > 0$ or $\lambda > 0$, then the matrix of second partial derivatives is positive definite in which case $\tilde{Q}$ is strictly convex and the minimizer is unique.)

Thus, it is minimized at its critical value $(\hat{\beta}_{0,\lambda}, \hat{\beta}_{1,\lambda})$.

(b) $\lim_{\lambda \to \infty} \hat{\beta}_{1,\lambda} = \lim_{\lambda \to \infty} \dfrac{(n-1)s_{x,y}}{(n-1)s_x^2 + \lambda} = 0$ and

$\lim_{\lambda \to \infty} \hat{\beta}_{0,\lambda} = \lim_{\lambda \to \infty} \bar{y} - \hat{\beta}_{1,\lambda}\bar{x} = \bar{y} - \lim_{\lambda \to \infty} \hat{\beta}_{1,\lambda}\bar{x} = \bar{y} - 0 = \bar{y}$