ROB 501 Exam-II Solutions

17 December 2018

Problem 1:

(a) False. The columns of V are e-vectors of $A^\top A$. Indeed, $A^\top A = V\Sigma^2 V^\top$.

(b) True. This is the main point of the SVD handout.

(c) False. $B = U\widehat{\Sigma}V^\top$ with $\widehat{\Sigma} = \mathrm{diag}(1,1,1,0)$ and thus has rank three.

(d) False. $A^\top A = V\Sigma^2 V^\top$, and hence $\lambda_{\max}(A^\top A) = 20^2 \neq 20$.

Problem 2:

(a) False. $\widehat{x} = S^{-1}C^\top \left(CS^{-1}C^\top\right)^{-1} y$. Also, we need linearly independent *columns* for $(C^\top SC)^{-1}$ to exist.

(b) As written, this problem is False. It should be $\widehat{x} = \left(C^\top SC\right)^{-1} C^\top Sy$. We will accept both answers!

(c) True. This is the BLUE. You can also see this by matrix multiplication:

$$KC = (C^\top Q^{-1}C)^{-1}\left(C^\top Q^{-1}C\right) = I$$

(d) True. Set $P = rI$ and take the limit as $r \to \infty$. We have

$$\widehat{x} = \lim_{r\to\infty} [C^\top Q^{-1}C + (rI)^{-1}]^{-1} C^\top Q^{-1} y = [C^\top Q^{-1}C]^{-1} C^\top Q^{-1} y.$$

Problem 3:

(a) True. Weierstrass Theorem.

(b) False. Completeness is required and we have nothing in the problem that implies that $\mathcal{X}$ is complete.

(c) False. S is not necessarily closed, and hence it does not necessarily contain its limit points.

(d) True. Because S is bounded, $\exists\ r < \infty$ such that $S \subset B_r(0)$. Hence, $S \subset \overline{B}_r(0)$, the closure of the ball, which, being closed and bounded in a finite-dimensional normed space, is compact. Therefore, the sequence (x_n) is contained in a compact set and must have a convergent subsequence. Being convergent, the subsequence is therefore Cauchy.

Problem 4:

(a) False. Take for example the open set $S = (0,1) \subset \mathbb{R}$, then $x = 1 \notin S$ satisfies $d(x, S) = 0$.

(b) True. S is closed if, and only if $\sim S$ is open, and $\sim S$ is open if, and only if $\forall\ x \in \sim S$, $d(x, \sim\sim S) > 0$. But $\sim\sim S = S$, and thus we see that the statement is true. <u>Second solution</u>: S is closed if, and only, if, $d(x,S) = 0 \implies x \in S$, and thus, $x \notin S$ must imply that $d(x,S) > 0$.

(c) False. All open balls are convex and they are not closed.

(d) True. y is a local minimum of f. Because $B_1(0)$ and f are convex, local minima are also global minima. Hence, the statement is true.

Problem 5:

(a) True. From lecture. We can also derive it:

$P_{k|k-1} - P_{k|k} = K_k C_k P_{k|k-1} = P_{k|k-1} C_k^\top \left(C_k P_{k|k-1} C_k^\top + Q_k \right)^{-1} C_k P_{k|k-1}$

Next we use the following matrix facts to show positive semi-definiteness:

$$\forall A \in \mathbb{R}^{n\times m}, \forall B \in \mathbb{R}^{n\times k}, \quad A \succeq 0 \implies B^\top A B \succeq 0 \tag{1}$$
$$\forall A \in \mathbb{R}^{n\times m}, \forall B \in \mathbb{R}^{n\times m}, \quad A \succeq 0, B \succeq 0 \implies A + B \succeq 0 \tag{2}$$
$$\forall A \in \mathbb{R}^{n\times n}, \det(A) \neq 0, \quad A \succeq 0 \implies A^{-1} \succeq 0 \tag{3}$$

The proof proceeds as follows:

- Start with the fact that $P_{k|k-1} \succeq 0$
- (1) $\implies C_k P_{k|k-1} C_k^\top \succeq 0$
- (2) $\implies C_k P_{k|k-1} C_k^\top + Q_k \succeq 0$ (where we also use $Q_k \succeq 0$
- (3) $\implies \left(C_k P_{k|k-1} C_k^\top + Q_k \right)^{-1} \succeq 0$
- (1) $\implies P_{k|k-1} C_k^\top \left(C_k P_{k|k-1} C_k^\top + Q_k \right)^{-1} C_k P_{k|k-1} \succeq 0$

(b) False. This could be larger or smaller depending on the matrices A_k, G_k, R_k. For instance, suppose $A_k = I$, $G_k = I$ and $R_k = I$. Then $P_{k+1|k} - P_{k|k} = I \succeq 0$.

(c) True. If Q_k is very large, then:

- $\left(C_k P_{k|k-1} C_k^\top + Q_k \right)^{-1} \approx Q_k^{-1} \approx 0$ and $K_k \approx 0$
- This gives $\hat{x}_{k|k} \approx \hat{x}_{k|k-1}$ and $P_{k|k} \approx P_{k|k-1}$.

This says that if our measurement is very noisy, it shouldn't affect our state estimate or variance.

(d) True. If Q_k is very small, then:

- $\left(C_k P_{k|k-1} C_k^\top + Q_k \right)^{-1} \approx \left(C_k P_{k|k-1} C_k^\top \right)^{-1}$
- $K_k = P_{k|k-1} C_k^\top \left(C_k P_{k|k-1} C_k^\top \right)^{-1}$
- $C_k \hat{x}_{k|k} = C_k \hat{x}_{k|k-1} + \left(C_k P_{k|k-1} C_k^\top \right) \left(C_k P_{k|k-1} C_k^\top \right)^{-1} \left(C_k \hat{x}_{k|k-1} - y_k \right) = y_k$

This says that if our measurement is very precise, the updated state should be approximately consistent with the value of the measurement.

Problem 6:

(a) Remember to flip the qualifiers and invert the statement. This gives us:

$$\exists x, y \in \mathcal{X}, \lambda \in [0,1], \text{s.t. } f\big(\lambda x + (1-\lambda)y\big) > \lambda f(x) + (1-\lambda) f(y)$$

(b) Since Q is orthogonal, we know $Q^{-1} = Q^\top$, therefore:

$$\begin{aligned}
\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= \begin{bmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}^\top \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} \\
&= \begin{bmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\
\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}
\end{aligned}$$

We then use back-substition to get:

$$\begin{aligned}
x_3 &= 1 \\
x_2 &= -2 - 2x_3 = -4 \\
x_1 &= 2 - 2x_2 - 3x_3 = 7
\end{aligned}$$

i.e. $x = [7, -4, 1]^\top$.

Alternate Solution:
Instead of backsubstitution, many of you directly computed R^{-1} using its adjoint and determinant (this is easy since R is upper triangular):

$$R^{-1} = \frac{1}{\det(R)} \text{adj}(A) = \frac{1}{1} \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

Then the solution can be computed as $x = R^{-1} Q^\top \begin{bmatrix} 0 & 0 & 3 \end{bmatrix}^\top$

(c) From lecture, we can write

$$A = \sigma_1 U_1 V_1^\top + \sigma_2 U_2 V_2^\top,$$

where U_i and V_i are columns of U and V, respectively. From the SVD Handout, we know that the matrix of smallest norm that drops the rank of $A - \Delta A$ is $\Delta A = \sigma_2 U_2 V_2^\top$.

If you want to work out the details, because $\sigma_2 = 0.1$ and $||U_2 V_2^\top|| = 1$, we have $||\sigma_2 U_2 V_2^\top|| = \sigma_2 = 0.1$. Additionally, since $A - \sigma_2 U_2 V_2^\top = \sigma_1 U_1 V_1^\top$ has rank 1, we obtain that $\Delta A = \sigma_2 U_2 V_2^\top$ satisfies the requirements.

Computing, we get:

$$\Delta A = \sigma_2 U_2 V_2^\top = \frac{0.1}{5\sqrt{2}} \begin{bmatrix} -4 & 3 \\ -4 & 3 \end{bmatrix}$$

Alternate Solution
Alternatively, you can get ΔA by selecting only the smallest singular value from Σ to get:

$$\Delta A = U \begin{bmatrix} 0 & 0 \\ 0 & 0.1 \end{bmatrix} V^\top$$

Problem 7: The solutions can be done straight out of the handout on Gaussian Random Vectors

(a) $Y = [2\ 3\ 0] \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$. Therefore, from the handout,

$$\mu_Y = [2\ 3\ 0]\mu_X = [2\ 3\ 0] \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = 2 \text{ and } \Sigma_Y = [2\ 3\ 0]\Sigma_X [2\ 3\ 0]^\top = [2\ 3\ 0] \begin{bmatrix} 6 & 2 & 4 \\ 2 & 3 & 1 \\ 4 & 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix} = 75$$

(b) We apply the formulas from the handout

$$\begin{aligned} \mu_{1|2} &:= \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2) \\ \Sigma_{1|2} &:= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \end{aligned}$$

and thus, in our case

$$\mu_{\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} | \{X_3=4\}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 4 \\ 1 \end{bmatrix} \frac{1}{3}\big(4 - (-1)\big) = \begin{bmatrix} 7\ \frac{2}{3} \\ 1\ \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{23}{3} \\ \frac{5}{3} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 23 \\ 5 \end{bmatrix}$$

and

$$\Sigma_{\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} | \{X_3=4\}} = \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix} - \begin{bmatrix} 4 \\ 1 \end{bmatrix} \frac{1}{3} \begin{bmatrix} 4 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 16 & 4 \\ 4 & 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & 2 \\ 2 & 8 \end{bmatrix}$$

(c) We give two solutions

(i) $Y|\{X_3 = 4\} = [2\ \ 3] \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} |\{X_3 = 4\}$ and thus

$$\mu_{Y|\{X_3=4\}} = [2\ \ 3]\mu_{\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} | \{X_3=4\}} = [2\ \ 3]\frac{1}{3} \begin{bmatrix} 23 \\ 5 \end{bmatrix} = \frac{61}{3}$$

(ii) The above is fine and is super quick. The solution I expected you to give is to first find the density of $Z = \begin{bmatrix} Y \\ X_3 \end{bmatrix}$, and then compute the conditional mean. We do that next.

$$Z = \begin{bmatrix} Y \\ X_3 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

and therefore

$$\begin{aligned} \mu_Z &= \begin{bmatrix} 2 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mu_X \\ &= \begin{bmatrix} 2 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ -1 \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned}
\Sigma_Z &= \begin{bmatrix} 2 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Sigma_X \begin{bmatrix} 2 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}^\top \\
&= \begin{bmatrix} 2 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 & 2 & 4 \\ 2 & 3 & 1 \\ 4 & 1 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 3 & 0 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 75 & 11 \\ 11 & 3 \end{bmatrix}
\end{aligned}$$

Hence, we apply the formula from the handout and arrive at

$$\mu_{Y|\{X_3=4\}} = 2 + 11\frac{1}{3}\big(4 - (-1\big) = \frac{61}{3},$$

the same answer we computed before.

Remark: This part was only worth 3 points so that if you did not find the quick solution from your work in (b), then you would not be penalized very much.

Problem 8:

(a) True. The Projection Theorem tells us that $\widehat{x}$ is unique and is characterized by $x - \widehat{x} \perp M$. Because $< x - x^*, v^i >= 0$ for $1 \le i \le k$, we have that $x^* \perp M$ also, and hence $\widehat{x} = x^*$.

Alternative very good (5): Use the normal equations to solve for $\widehat{x}$. Recognize that we can use the orthonormal basis $\{v\}$, in which the Gram matrix is the identify, $G = I$. Hence, the normal equations give $\widehat{x} = x^*$

Alternative pretty good (5): $\widehat{x}$ is the projection of x onto span$\{y^1, \cdots, y^k\}$ while x^* is the projection of x onto span$\{v^1, \cdots, v^k\}$ (something we worked out in lecture). Because the two spans are equal, we have $\widehat{x} = x^*$

Alternative that is not too bad (4): State that both formulas give the orthogonal projection of x onto M. This answer is not as complete as I would like, but in the stress of an exam, it's not so bad that I want to take off many points. You are using the term "orthogonal projection" as an operator and you are using it with confidence.

Others: There were too many others to list that selected T (true) but really did not give an explanation that hangs together. These received one point for selecting T and one to two points, mostly, for stating useful clues as to why the result is true.

(b) From HW, we know the answer is $f = \alpha_1 t^2 + \alpha_2 \sin(\pi t)$, where $G\alpha = c$. We compute the appropriate terms to have

$$G = \begin{bmatrix} < t^2, t^2 > & < t^2, \sin(\pi, t) > \\ < t^2, \sin(\pi, t) > & < \sin(\pi, t), \sin(\pi, t) > \end{bmatrix} = \begin{bmatrix} \frac{2}{5} & 0 \\ 0 & 1 \end{bmatrix}$$

Hence,

$$\begin{bmatrix} \frac{2}{5} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 4 \\ \pi \end{bmatrix}$$

and therefore $f = 10t^2 + \pi \sin(\pi t)$

(c) First note that:

$$\frac{\partial h}{\partial x} = \begin{bmatrix} e^{x_1} & 1 \\ 1 & 2x_2 \end{bmatrix}$$

We can then compute our Newton-Raphson step:

$$x_{k+1} = x_k - \left(\frac{\partial h}{\partial x}(x_k) \right)^{-1} (h(x_k) - y)$$

and thus

$$
\begin{aligned}
x_1 &= x_0 - \left(\frac{\partial h}{\partial x}(x_0)\right)^{-1} (h(x_0) - y) \\
&= \begin{bmatrix} 0 \\ -1 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}^{-1} \left(\begin{bmatrix} 0 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right) \\
&= \begin{bmatrix} 0 \\ -1 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 \\ -1 \end{bmatrix} + \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix} \\
&\therefore \\
x_1 &= \begin{bmatrix} 2/3 \\ -2/3 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 \\ -2 \end{bmatrix}
\end{aligned}
$$

Problem 9: A+

Inspired by $\mathbb{R}^n$, we define e_i to be an infinite sequence that has all entries equal to zero except a one in the $i - th$ element. That is $e_i := (0, 0, \ldots, 1, 0, 0, \ldots)$, so that

$$\begin{aligned} e_1 &:= (1, 0, 0, 0, \ldots) \\ e_2 &:= (0, 1, 0, 0, \ldots) \\ e_3 &:= (0, 0, 1, 0, \ldots) \\ &\text{etc.} \end{aligned}$$

It is obvious that $\forall\ n \geq 1$, $||e_n||_\infty = 1$ and thus $e_n \in \overline{B}_1(0)$. Also, for $i \neq j \implies ||e_i - e_j||_\infty = 1$, and therefore sequence $(e_n)_{n=1}^\infty$ cannot have a Cauchy subsequence. Indeed, for $n_i \neq n_j$,

$$||e_{n_i} - e_{n_j}||_\infty = 1.$$