ROB 501 Exam-II Solutions

Fall 2020 (Prof. Gregg)

Problem 1:

(a) True, because $\{a, b\} = \{a\} \bigcup \{b\}$, which is a finite union of two closed sets and thus closed.

(b) False. Not containing all limit points means not closed, but that doesn't mean the set is open. For example, the interval [0,1) is neither open nor closed.

(c) True, e.g., $(0, 2] \bigcap [1, 3) = [1, 2]$ or $[0, 2) \bigcap (1, 3] = (1, 2)$.

(d) False. The existence of a convergent subsequence does not guarantee that the sequence itself converges, even in a compact set.

Problem 2:

(a) True, $a_n = 1 + \frac{1}{n}^n$ converges to e. In fact, as covered in recitation, all irrational numbers are limit points of $\mathbb{Q} \subset (\mathbb{R}, |\cdot|)$

(b) True. Compact $\implies$ the set is closed and bounded, so A is a closed subset of a finite dimensional subspace $\implies$ A must be complete

(c) False - this question is asking if log(n) is Cauchy, and it is not.

(d) False. Consider the following counter example: Let $A = [-2, 2]$ and let $x^* = 0$. Define $f(x)$ such that $f(x) = x^2$ on $[-2, 0)$ and $(x-1)^2 + 1$ on $[0, 2]$. The $f(x)$ has an infimum at 0, but a minimum value is not actually achieved.

Problem 3:

(a) True. This form is sometimes given in the covariance definition. We can also derive it:

$$\begin{aligned}
E\{(X_1 - E\{X_1\})(X_2 - E\{X_2\})^\top\} &= E\{X_1X_2^\top - X_1E\{X_2\}^\top - E\{X_1\}X_2^\top + E\{X_1\}E\{X_2\}^\top\} \\
&= E\{X_1X_2^\top\} - E\{X_1E\{X_2\}^\top\} - E\{E\{X_1\}X_2^\top\} + E\{E\{X_1\}E\{X_2\}^\top\}\} \\
&= E\{X_1X_2^\top\} - E\{X_1\}E\{X_2\}^\top - E\{X_1\}E\{X_2\}^\top + E\{X_1\}E\{X_2\}^\top \\
&= E\{X_1X_2^\top\} - E\{X_1\}E\{X_2\}^\top
\end{aligned}$$

(b) False. Uncorrelated does not imply independent unless RVs are jointly Gaussian distributed.

(c) False. Relate X and Y by defining their joint Gaussian random vector through a linear combination:

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} I \\ A \end{bmatrix} X + \begin{bmatrix} 0 \\ b \end{bmatrix}, \quad \text{where} \quad \text{cov}\left(\begin{bmatrix} X \\ Y \end{bmatrix}\right) = \begin{bmatrix} I \\ A \end{bmatrix} \Sigma_X \begin{bmatrix} I & A^\top \end{bmatrix} = \begin{bmatrix} \Sigma_X & \Sigma_X A^\top \\ A\Sigma_X & A\Sigma_X A^\top \end{bmatrix}$$

Then the conditional covariance of Y given X is

$$\Sigma_{Y|X} = A\Sigma_X A^\top - A\Sigma_X(\Sigma_X)^{-1}\Sigma_X A^\top = A\Sigma_X A^\top - A\Sigma_X A^\top = 0$$

The covariance given in the question is $\Sigma_{X|Y}$.

(d) True. Relate X_1 and Y by defining their joint Gaussian random vector through a linear combination:

$$\begin{bmatrix} X_1 \\ Y \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 + X_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

$$
\operatorname{cov}\left(\begin{bmatrix} X_1 \\ Y \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \Sigma_{11} & 0 & 0 \\ 0 & \Sigma_{22} & \Sigma_{23} \\ 0 & \Sigma_{32} & \Sigma_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}^\top
$$
$$
= \begin{bmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} + 2\Sigma_{23} + \Sigma_{33} \end{bmatrix}
$$

Since the off-diagonal terms are zero, these variables are uncorrelated. Since they are jointly normal, independent $\iff$ uncorrelated.

Problem 4:

(a) False. We need the constraint $KC = I$ for the estimate to be unbiased.

(b) This problem was poorly constructed, as the dimensions of Q do not match the dimensions of x in the inner product definition. This was supposed to be underdetermined least squares, but we messed up and gave credit for either answer.

(c) False. The conditions for MVE are met (note the constraint $CPC^\top + Q > 0$ is satisfied because $Q > 0$, even though $CPC^\top \geq 0$). When we have a non-zero constant mean, we must offset both the estimate *and the observation* by the mean. The correct estimate would be $\hat{x}_t = \bar{x} + PC^\top (CPC^\top + Q)^{-1}(y_t - C\bar{x})$.

(d) True. We showed this in the MVE derivation. See the notes for details.

Problem 5:

(a) True. Let $N = V\Sigma^T U$, then $M = N^T N = V\Sigma^T U^T U \Sigma V^T = V\Sigma^T \Sigma V^T$, where $\Sigma^T \Sigma$ contains the singular values of M.

(b) True if A is full rank, which we assumed in lecture. The Q matrix columns form a basis for the range of A, so the span of the columns of Q is the same as the range of A. However, we also gave credit for False because you can find cases online where A is singular, and thus the span of Q is not equal to the span of A.

(c) True. This problem is simply a variation of HW 8, problem 6. Let $r = \begin{bmatrix} -p \\ 1 \end{bmatrix}$. We can rewrite L as $V = \{x \in \mathbb{R}^2 | < x, r >= q\}$. By underdetermined least squares, we know that there exists a minimum norm vector in V.

(d) False, only drops rank if $\beta = \sigma_r$.

Problem 6:

(a) Turn above information into the familiar matrix equation:

$$\underbrace{\begin{bmatrix} d_t \\ n_t \end{bmatrix}}_{y_t} = \underbrace{\begin{bmatrix} \frac{1}{c_1} & 0 \\ 0 & \frac{\Delta t N}{2\pi r} \end{bmatrix}}_{C} \underbrace{\begin{bmatrix} x_t \\ v_t \end{bmatrix}}_{z_t} + \underbrace{\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix}}_{\epsilon_t}$$
$$= \begin{bmatrix} 500 & 0 \\ 0 & 47.7465 \end{bmatrix} \begin{bmatrix} x_t \\ v_t \end{bmatrix} + \epsilon_t,$$

where $\epsilon_t \sim \mathcal{N}(0, Q)$ and $Q = \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix} = \begin{bmatrix} 3 \times 10^4 & 0 \\ 0 & 1 \times 10^3 \end{bmatrix}$. Incorporating three measurements:

$$\underbrace{\begin{bmatrix} d_0^{(1)} \\ n_0^{(1)} \\ d_0^{(2)} \\ n_0^{(2)} \\ d_0^{(3)} \\ n_0^{(3)} \end{bmatrix}}_{y_0} = \underbrace{\begin{bmatrix} C \\ C \\ C \end{bmatrix}}_{C_0} \underbrace{\begin{bmatrix} x_0 \\ v_0 \end{bmatrix}}_{z_0} + \underbrace{\begin{bmatrix} \epsilon_0^{(1)} \\ \epsilon_0^{(2)} \\ \epsilon_0^{(3)} \end{bmatrix}}_{\epsilon_0}$$

with $\epsilon_0 \sim \mathcal{N}(0, Q_0)$, $Q_0 = \text{diag}(Q, Q, Q)$. Then the estimate is:

$$\begin{bmatrix} \hat{x}_0 \\ \hat{v}_0 \end{bmatrix} = (C_0^\top Q_0^{-1} C_0)^{-1} C_0^\top Q_0^{-1} y_0 = \begin{bmatrix} 0.5867 \\ 0.0070 \end{bmatrix}$$

The covariance is:

$$P_0 = (C_0^\top Q_0^{-1} C_0)^{-1} = \begin{bmatrix} 0.0400 & 0 \\ 0 & 0.1462 \end{bmatrix}$$

Alternate solution: We will also accept solutions which assumed that the terms ϵ_1 and ϵ_2 were noise applied to the state. In this case:

$$\underbrace{\begin{bmatrix} d_t \\ n_t \end{bmatrix}}_{y_t} = \underbrace{\begin{bmatrix} \frac{1}{c_1} & 0 \\ 0 & \frac{\Delta t N}{2\pi r} \end{bmatrix}}_{C} \underbrace{\begin{bmatrix} x_t \\ v_t \end{bmatrix}}_{z_t} + \underbrace{\begin{bmatrix} \frac{1}{c_1}\epsilon_1 \\ \frac{\Delta t N}{2\pi r}\epsilon_2 \end{bmatrix}}_{\epsilon_t}$$

Note that it doesn't affect our calculations if we multiply ϵ_t by -1, so we make it positive for simplicity. The noise terms will still have mean zero, but their covariances must be updated:

$$\text{cov}\left(\frac{1}{c_1}\epsilon_1\right) = \left(\frac{1}{c_1}\right)^2 q_1 = 7.5 \times 10^9$$
$$\text{cov}\left(\frac{\Delta t N}{2\pi r}\epsilon_2\right) = \left(\frac{\Delta t N}{2\pi r}\right)^2 q_2 = 2.2797 \times 10^7$$
$$Q = \begin{bmatrix} 7.5 \times 10^9 & 0 \\ 0 & 2.2797 \times 10^6 \end{bmatrix}$$

These covariances are huge! Using these values, the solution is:

$$\begin{bmatrix} \hat{x}_0 \\ \hat{v}_0 \end{bmatrix} = \begin{bmatrix} 0.5867 \\ 0.0070 \end{bmatrix}, \quad P_0 = \begin{bmatrix} 1 \times 10^4 & 0 \\ 0 & 333.3333 \end{bmatrix}$$

(b) Again, we can turn the equations of motion into matrix form:

$$\underbrace{\begin{bmatrix} x_{t+1} \\ v_{t+1} \end{bmatrix}}_{z_{t+1}} = \underbrace{\begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_t \\ v_t \end{bmatrix}}_{z_t} + \underbrace{\begin{bmatrix} \frac{\Delta t^2}{mr} \\ \frac{\Delta t}{mr} \end{bmatrix}}_{B} \underbrace{\tau_t}_{u_t} + \underbrace{\begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix}}_{\delta_t}$$
$$= \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_t \\ v_t \end{bmatrix} + \begin{bmatrix} 0.25 \\ 0.5 \end{bmatrix} \tau_t + \delta_t$$

We use the Kalman filter to estimate the distribution at time $t_1 = 1 \cdot \Delta t$. The way the problem is worded, the control signal is applied before we take our measurement, so we apply the prediction step first, using the values from (a) as the initial estimate.

$$\hat{z}_{1|0} = A\hat{z}_0 + B\tau_0 = \begin{bmatrix} 0.7902 \\ 0.4070 \end{bmatrix}$$
$$P_{1|0} = AP_0A^\top + R = \begin{bmatrix} 0.1266 & 0.0831 \\ 0.0831 & 0.2262 \end{bmatrix}$$

Next, we incorporate the measurement, $y_1 = \begin{bmatrix} 410 & 18 \end{bmatrix}^\top$. The measurement update step gives:

$$K_1 = P_{1|0}C^\top(CP_{1|0}C^\top + Q)^{-1}$$
$$\hat{z}_{1|1} = \hat{z}_{1|0} + K_1(y_1 - C\hat{z}_{1|0})$$
$$P_{1|1} = P_{1|0} - K_1CP_{1|0}$$

Then the final distribution is $z_{1|1} \sim \mathcal{N}(\hat{z}_{1|1}, P_{1|1})$, where:

$$\hat{z}_{1|1} = \begin{bmatrix} 0.8029 \\ 0.4046 \end{bmatrix}, \quad P_{1|1} = \begin{bmatrix} 0.0590 & 0.0279 \\ 0.0279 & 0.1365 \end{bmatrix}$$

Alternate solution: If you used the alternate Q matrix from (a), then the solution is:

$$\hat{z}_{1|0} = \begin{bmatrix} 0.7902 \\ 0.4070 \end{bmatrix}, \quad P_{1|0} = \begin{bmatrix} 1.0083 \times 10^4 & 166.6767 \\ 166.6767 & 333.4133 \end{bmatrix}, \quad \hat{z}_{1|1} = \begin{bmatrix} 0.7948 \\ 0.3996 \end{bmatrix}, \quad P_{1|1} = \begin{bmatrix} 7.5351 \times 10^3 & 93.6036 \\ 93.6036 & 249.6550 \end{bmatrix}$$

Problem 7:

(a) The columns of U are the eigenvectors of AA^T, giving

$$U = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The columns of V are the eigenvectors of A^TA, giving

$$V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

The rectangular diagonal matrix Σ has square roots of the eigenvalues of both AA^T and A^TA, giving

$$\Sigma = \begin{bmatrix} b\sqrt{2} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix}.$$

Note that $b\sqrt{2} > \sqrt{2}$ for $b > 1$, which was assumed.

(b) First note that $M = A^T A = V\Sigma^T U^T U \Sigma V^T = V\Sigma^T \Sigma V^T = V \begin{bmatrix} 2b^2 & 0 \\ 0 & 2 \end{bmatrix} V^T$, where $U^T U = I$ because U is an orthogonal matrix by construction. This is the singular value decomposition of M, where $\sigma_1 = 2b^2$, $\sigma_2 = 2$. We can now obtain the expansion $M = \sigma_1 v_1 v_1^T + \sigma_2 v_2 v_2^T$, where $v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Hence,

$$\Delta M = -\sigma_2 v_2 v_2^T = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

is the minimum-norm perturbation that drops the rank of M. Specifically,

$$||\Delta M||_2 = \sigma_2 = 2.$$

Problem 8:

(a) TRUE. Note that $\mathbb{P}^n$ is finite-dimensional, and $\mathbb{P}^m$ is a subspace of $\mathbb{P}^n$ for $m < n$. Any finite-dimensional subspace is complete, which by definition means any Cauchy sequence (x_k) in $\mathbb{P}^m$ has a limit in $\mathbb{P}^m$. You could also just say that $\mathbb{P}^m$ is a finite-dimensional vector space and thus complete. Alternatively, recognize that (x_k) can be represented as a sequence of m real coefficients, and $\mathbb{R}^m$ is complete.

(b) FALSE. This set is non-convex, and balls must be convex in any normed space. This follows from the definition of a convex set and the triangle inequality of a norm. See `https://proofwiki.org/wiki/Open_Ball_is_Convex_Set` for the proof. Sadly this means there is no "frog norm" (unless you can draw a convex frog), but there is hope in more general metric spaces: `https://math.stackexchange.com/questions/2900904/are-there-ball-that-are-not-convex`

(c) FALSE. First note that $\bigcap\limits_{n=1}^{\infty} (-\frac{n+1}{n}, \frac{n+1}{n}) = [-1, 1]$ and $\bigcup\limits_{n=1}^{\infty} (-\frac{1}{n}, \frac{1}{n}) = (-1, 1)$. Then $[-1, 1] \bigcap (-1, 1) = (-1, 1)$, which is clearly an open set. This cannot be compact in a finite-dimensional normed space, by Bolzano-Weierstrass.