Review Problems: ORF 418 - Fall 2025

Midterm - October 22, 2025   version: August 29, 2025

Structure.

- The exam will be in the usual classroom, Aaron Hall 219.
- You may bring one page document (one-side only) prepared by yourself. document.
- You will write your exam on a regular paper.
- You will be asked to scan your handwritten exam at the end of the exam. So bring an electronic devise capable of that.
- However, this electronic devise must be stored away during the exam.

Topics. The midterm will cover the following topics:

1. Armed bandits. (Section 1.3)
   - Problem setting; (1.3.1)
   - Exploration and exploitation; (1.3.2)
   - ϵ-greedy algorithm. (1.3.2)
2. Linear Quadratic Regulator (All of chapter 2)
   - Problem setting: infinite, finite horizon, and the stochastic problems; (2.1, 2.4)
   - Dynamic Programming Equations and optimal control; (2.2.1, 2.2.2)
   - Solution through Riccati equations and the gains matrix; (2.2.3)
   - Solution in finite horizon; (2.2.4)
   - Controllability; (2.4)
   - Deriving the problem from a given a continuous time model (2.3.1).
3. Bayesian Learning (All of chapter 3)
   - Bayes' formulae and its use; (3.1)
   - Bayesian estimation; (3.2.1)
   - Maximum Likelihood Estimation (3.2.2)
4. Kalman Filter (Sections 4.1, 4.2)
   - Problem definition; (4.1, 4.1.3)
   - Solution; (4.2)
   - Long time behavior; (4.2.4)

- Some problems are too long as midterm questions. However, each part can be a midterm question.
- If you have corrections, question or comments, please contact me.
- I will update and make corrections. So if you are not sure about the statement of a problem or certain things are not clear, please check the latest version in Canvas.
- Solutions will be posted on Canvas.

---

1 Multi-armed bandit problems

Exercise 1.1. Armed bandit with two arms and the reward distribution of both arms are uniform on the interval $[a, a+2]$ for $a = 1, 2$. We explore each N times with rewards $\{r_k^{(a)}\}_{k=1,\ldots N}$. We use the estimate

$$\widehat{q}(a) := \frac{1}{N} \sum_{k=1}^{N} r_k^{(a)}, \qquad a = 1, 2.$$

Let

$$a^* := \text{argmax}_a \ \widehat{q}(a),$$

and *we choose* 1 *if both are maximizers.* After the first N trials, we *only* use a^* and receive rewards $\{r_t\}_{t=1,\ldots}$. *i.e.*, there is no exploration after N. The limiting reward is defined as

$$J := \lim_{T\to\infty} \frac{1}{T} \sum_{k=1}^{T} r_k.$$

a. Suppose $a^* = 1$ and compute J.

b. Suppose $a^* = 2$ and compute J.

c. Suppose $N = 1$ and compute the distribution of a^*.

d. Suppose $N = 1$ and compute $\mathbb{E}[\,J\,]$.

e. Suppose we use an ϵ-greedy algorithm with $\epsilon = 10\%$. What would be the resulting J value? (Simply state the result and give some reasoning but do not prove it.)

Exercise 1.2. Same problem as above. Assume that the reward distribution of each arm is Bernoulli with success probability $\frac{1}{2a}$ for $a = 1, 2$.

a. Suppose $a^* = 1$ and compute J.

b. Suppose $a^* = 2$ and compute J.

c. Suppose $N = 1$ and compute the distribution of a^*.

d. Suppose $N = 1$ and compute $\mathbb{E}[\,J\,]$.

---

2 Linear-Quadratic and related problems

Exercise 2.1. Consider the LQR with

$$A = \begin{bmatrix} 0.5 & 0 \\ 0 & a \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad N = 1,$$

with $0 < \rho \leq 1$ and the constant a are arbitrary.

a. Is (A, B) controllable?

b. For which values of $x_0 = x$, ρ and a, is the value function $v(x)$ finite?

c. For $a = \rho = 1$, write down the Riccati equation with infinite horizon.

d. Solve the above Riccati equation for $a = \rho = 1$.

Exercise 2.2. Consider the LQR with

$$A = \begin{bmatrix} 1 & 0 \\ 0 & a \end{bmatrix}, \qquad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad N = 1,$$

with $0 < \rho \leq 1$ and the constant a are arbitrary.

a. For which values of a, the system (A, B) controllable?

b. For which values of x, ρ and a, is the value function $v(x)$ finite?

Exercise 2.3. Consider the control problem (which is not a LQR):

$$x_{k+1} = u_k\, x_k + 1,$$

with a cost function for $\mathbf{u} = (\mathbf{u_0}, \mathbf{u_1}, \ldots)$,

$$J(x, \mathbf{u}) = \sum_{\mathbf{k=0}}^{\infty} \mathbf{2^{-k}}\, [\mathbf{x_k^2} + \mathbf{u_k^2}].$$

Let $v(x) = \min_{\mathbf{u}} J(x, \mathbf{u})$ be the value function.

a. Write down the dynamic programming equation for v (do not solve it).

b. Show that $v(0) < \infty$.

c. Show that $v(x) < \infty$ for every x

Exercise 2.4. Consider the control problem (which is not a LQR):

$$x_{k+1} = x_k + u_k\, [z_k + 1],$$

where $z_0, z_1, \ldots$ is an i.i.d. sequence of random variables taking values ± 1 with equal probability. For a finite horizon n and control $\mathbf{u} = (\mathbf{u_0}, \mathbf{u_1}, \ldots)$,

$$J(n, x, \mathbf{u}) := \mathbb{E}\left[x_n - \frac{1}{2}\sum_{k=0}^{n-1} u_k^2\right].$$

Let $v(n, x) = \max_{\mathbf{u}} J(n, x, \mathbf{u})$ be the value function.

a. Write down the dynamic programming equation for v.

b. Show that $v(n, x) = x + \frac{n}{2}$.

c. Show that the optimal control is $u_k^* = 1$ for every k.

Exercise 2.5. Consider the continuous time system

$$y''(t) = \sin(y(t)) + (1 + y(t))\, u(t),$$

where $u(t)$ is the control.

a. Rewrite the above equations as a system of first order equations.

b. Linearize the resulting system around $y(t) \equiv 0$.

c. Discretize (in time) the resulting linear equation.

d. Is the resulting system controllable?

Exercise 2.6. Consider the continuous time system

$$y'''(t) = (y'(t))^3 + u(t),$$

where $u(t)$ is the control.

a. Rewrite the above equations as a system of first order equations.

b. Linearize the resulting system around $y(t) \equiv 0$.

c. Discretize (in time) the resulting linear equation.

d. Is the resulting system controllable?

---

3 Bayesian estimation and related problems

Exercise 3.1. Consider a medical test which that gives three possible results: 'p=positive', 'n=negative', and 'u=uncertain'. A randomly chosen person is tested. Let

$A =$ the person is infected,
$B_p =$ the test is positive
$B_n =$ the test is negative
$B_u =$ the test is uncertain.

We know that $\mathbb{P}(A) = 0.05$ and

$$\mathbb{P}(B_p \,|\, A) = 0.7, \quad \mathbb{P}(B_n \,|\, A) = 0.1, \quad \mathbb{P}(B_p \,|\, A^c) = 0.1, \quad \mathbb{P}(B_n \,|\, A^c) = 0.6,$$

a. Compute $\mathbb{P}(B_u \,|\, A)$, and $\mathbb{P}(B_u \,|\, A^c)$.

b. Compute $\mathbb{P}(A \,|\, B_p)$.

c. Compute $\mathbb{P}(A \,|\, B_u)$.

d. Compute $\mathbb{P}(A \,|\, B_n)$.

Exercise 3.2. In a game, we are told that only one of the three people Halil (first person), Mete (second), or Soner (third) has a reward and it will be given to us if we guess it correctly. All of them know who has the reward. When asked *"who has the reward?"*, they reply truthfully, if they do not have it. But when they have it, with probability p the answer is equal to one of the other persons with the minimum ranking and $(1-p)$ is the other person.

Consider the events:

$A_h =$ Halil has it, $\qquad A_m =$ Mete has it, $\qquad A_s =$ Soner has it.

Our prior is that all are equally likely, i.e., have probability $1/3$. We ask Mete and the answer is Soner. Let B be this event and let $p = 0.25$.

a. Compute $\mathbb{P}(B \,|\, A_h)$, $\mathbb{P}(B \,|\, A_m)$, and $\mathbb{P}(B \,|\, A_s)$.

b. Compute $\mathbb{P}(B)$.

c. Use Bayes' formula to compute all three posterior probabilities $\mathbb{P}(A_h \,|\, B)$, $\mathbb{P}(A_m \,|\, B)$, and $\mathbb{P}(A_s \,|\, B)$.

Now we change the formula for the answer to the question *"who has it?"*. In all below cases, compute the posterior probabilities.

d. Suppose that the answer is one of the other two with equal probability $p = 0.5$ no matter who actually has it.

e. The answer is always with probability $p = 0.25$ the smaller ranked person other than the one asked, and with probability $(1-p)$ the higher ranked person other than the one asked, regardless of who has the reward.

f. If the person has it, the answer is chosen equally likely from all three names including his. When he does not have it, he answers truthfully.

Exercise 3.3. Suppose we flip a coin, with head-probability of P, 100 times. The outcome is 52 heads. Suppose the prior distribution of p is $f_P(p) = 2\,p$ for $p \in (0,1)$.

a. Compute the posterior distribution of P.

b. Compute the Bayesian estimate for P for this prior.

Hint: The Beta distribution with parameters $a, b > 0$, i.e., the p.d.f. is given by,

$$f_P(p) = \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)}\, p^{a-1}\,(1-p)^{b-1}, \qquad p \in (0,1),$$

where Γ is the *gamma function*:

$$\Gamma(a) := \int_0^\infty u^{a-1}\, e^{-u}\, du.$$

Moreover, its *mean is equal to* $a/(a+b)$.

Exercise 3.4. *Suppose that $Y_1, \dots, Y_n \in \{0,1\}$ are i.i.d. from the Bernoulli distribution with parameter $P \in (0,1)$, i.e., $\mathbb{P}(Y_1 = 1 \,|\, P = p) = p$. The parameter P is not known and hence, is a random variable. Suppose that the prior distribution of P is the beta distribution with parameters $a, b > 0$ (see the hint in the above problem).*

Show that the posterior distribution of P again beta with parameters,

$$a + \sum_i x_i, \qquad b + n - \sum_i x_i.$$

Exercise 3.5. Suppose the proportion θ of detective items in a large manufactured lot is unknown. When eight items were selected at random from the lot, and it is found that three are defective.

a. Suppose that and the prior distribution of θ is uniform on the unit interval $[0,1]$. Compute the posterior distribution and the Bayesian estimate of θ.

b. Suppose that and the prior distribution of θ is

$$f_\theta(\theta) = 2(1-\theta), \qquad \theta \in [0,1].$$

Compute the posterior distribution and the Bayesian estimate of θ.

c. What is the maximum likelihood estimator?

Exercise 3.6. Let θ be the proportion of registered voters in a large city who are in favor of a certain proposition. Suppose that the value of θ is not known and has to be estimated. In a random sample of $1,000$ registered voters from the city, it is found that 710 are in favor of the proposition.

a. Suppose that the prior is $f_\theta(\theta) = 2\,\theta$ for $\theta \in [0,1]$. Compute the posterior distribution and the Bayesian estimate of θ.

b. Suppose that the prior is $f_\theta(\theta) = 4\,\theta^3$ for $\theta \in [0,1]$. Compute the posterior distribution and the Bayesian estimate of θ.

c. What is the maximum likelihood estimator?

Exercise 3.7. Suppose that $\mathbf{Y} := (Y_1, Y_2, \dots, Y_n)$ form a random sample from a distribution for which the probability density function (for any $k = 1, \dots, n$) is

$$f_{Y_k|\theta}(y_k \,|\, \theta) \;=\; \theta\, y_k^{\theta-1}, \qquad y_k \in (0,1).$$

Compute the maximum likelihood estimator of θ given the observation $\mathbf{Y} = y = (y_1, \dots, y_n)$.

---

4 Kalman Filter

Exercise 4.1. Consider a one dimensional Kalman filter problem with $a = 1$, $h = 3/2$ $p_0 = q = r = 1$ and $\widehat{x}_0 = 1$.

a. Compute $\bar{p}_k$ and p_k for each $k = 1, 2$.

b. Suppose that all limits $\bar{p}^* := \lim_{k\to\infty} \bar{p}_k$, $K^* := \lim_{k\to\infty} K_k$ and $p^* := \lim_{k\to\infty} p_k$ exists. Compute $\bar{p}^*$, K^*, p^*.

c. Suppose that p_0 is equal to p^* computed in part a. Show that p_k, $\bar{p}_k$, and K_k are all independent of k and are equal to p^*, $\bar{p}^*$, K^*, respectively.

d. With p_0 equal to p^*, show that the innovation process is i.i.d.

Exercise 4.2. Suppose that $d = 2$, $\ell = 1$, $\hat{x} = 0 \in \mathbb{R}^2$, $A_k = I$ is the 2×2 identity matrix, I_2, $H_k = [1,1]$ and $P_0 = Q_k = I_2$, and $R_k = 1$. Hence, $x_0 \sim \mathcal{N}(0, I_2)$,

$$x_{k+1} = x_k + w_k, \quad z_k = [1,1]x_k + \nu_k, \quad \omega_k \sim \mathcal{N}(0, I_2), \quad \nu_k \sim \mathcal{N}(0,1).$$

a. Compute $\overline{P}_1, K_1, P_1, \bar{P}_2, K_2$.

b. Compute $\hat{x}_{1|0}, \hat{x}_1, \hat{x}_{2|1}, \hat{x}_2$ as functions of z_1 and z_2.

c. Let $x_k = (\, v_k \,,\, s_k \,)$. Argue that for each k, $\widehat{v}_k = \widehat{s}_k$ and

$$P_k = \begin{bmatrix} b_k & c_k \\ c_k & b_k \end{bmatrix}, \quad \bar{P}_k = \begin{bmatrix} \bar{b}_k & \bar{c}_k \\ \bar{c}_k & \bar{b}_k \end{bmatrix}, \quad K_k = \begin{bmatrix} m_k \\ m_k \end{bmatrix},$$

for some b_k, c_k, $\bar{b}_k$, $\bar{c}_k$, and m_k. Is this consistent with part a?

d. Set $y_k := v_k + s_k$. Formulate a Kalman filter using only y_k as the state.

e. Let $p_k = \mathbb{E}[(y_k - \widehat{y}_k)^2]$. Show that $p_k = 2\,[b_k \,+\, c_k]$.
