Probabilistic Artificial Intelligence
Questions Pack

January 22, 2025

Time limit: 120 minutes

Instructions. This pack contains all questions for the final exam. It contains the questions only. You must provide your answers to the questions on the accompanying **answer sheet** by blackening out the corresponding squares. Please carefully read the following instructions regarding the answer sheet:

- During the exam, you may use pencil and eraser to mark and edit your answers on the answer sheet. After the time is up, we will collect the questions pack. Make sure that you marked all your answers by then, at least with a pencil. After collecting the questions pack, we provide you with additional **10** minutes of time to blacken out the squares on the answer sheet with a **black pen**.
- As the answer sheet will be graded by a computer, please **make sure to blacken out the whole square. Do not use ticks or crosses**.
- Nothing written on the pages of the question pack will be collected or marked. **Only the separate answer sheet with the filled squares will be marked**.
- Please make sure that your answer sheet is clean and do not write anything on the answer sheet except the squares you blackened out. We reserve the right to classify answers as wrong without further consideration if the sheet is filled out ambiguously.

Collaboration on the exam is strictly forbidden. You are allowed a summary of *two* A4 pages and a simple, non-programmable calculator. The use of any other helping material or collaboration will lead to being excluded from the exam and subjected to disciplinary measures by the ETH Zurich disciplinary committee.

Question Types In this exam, questions award **1, 2, 3, or 4** points if answered correctly, depending on the difficulty of the question. You will encounter the following question types:

- **True or False.**
- **Multiple choice questions with a *single* correct answer.**

  These multiple choice questions have **exactly one** correct choice.
- **Multiple choice questions with *multiple* correct answers.**

  These multiple choice questions are marked with a "♣" and have **one or more** correct choices (at least one choice is correct). You will receive full points if you have chosen **all** the correct choices and **none** of the wrong choices. You will get zero points otherwise.

There are no negative points in this exam, so **0 points** are awarded if a question is answered wrongly or not attempted. The sections and questions are not ordered by difficulty. Thus, if you find a question too difficult, it may make sense to get back to it in the end.

Questions in the same section might be related. In addition, some questions require information from previous parts. Hence, if you intend to skip some questions, make sure that you read all notes and details between the questions.

The notation $\mathcal{N}(\mu, \sigma^2)$ denotes the Gaussian random distribution with mean μ and variance σ^2. For example, $\mathcal{N}(3, 4)$ designates the Gaussian distribution with mean 3 and standard deviation 2.

Good luck!

1 Probability and Regression (11 points)

Question 1 *(1 point)*
For two jointly Gaussian random variables X, Y, the variance of the conditional random variable $X|Y = y$ is always less than or equal to the marginal variance of X.

■ True B False

Solution:
True. *(Lecture 1, slide 52)* The conditional variance is

$$\mathbf{\Sigma}_{X|Y} = \mathbf{\Sigma}_{XX} - \mathbf{\Sigma}_{XY}\mathbf{\Sigma}_{YY}^{-1}\mathbf{\Sigma}_{YX},$$

where the second term is greater or equal to 0.

Question 2 *(1 point)*
Consider random variables X, Y, Z, where X and Y are conditionally independent given Z, i.e. $X \perp Y|Z$. Then for all x, y, z it holds that $P(X = x, Y = y|Z = z) = P(X = x|Z = z) \cdot P(Y = y|Z = z)$

■ True B False

Solution:
True. *(Lecture 1, slide 42)* By definition of conditional independence.

Question 3 *(1 point)*
For n binary random variables $X_1, \ldots, X_n$, we need $\mathcal{O}(n^2)$ parameters in the worst case to specify the joint distribution $P(X_1, \ldots, X_n)$.

A True ■ False

Solution:
False. *(Lecture 1, slide 43)* In the worst case, we need $2^n - 1$ parameters to completely specify the joint distribution.

Question 4 *(1 point)*
Let $(X, Y)^\top \sim \mathcal{N}((\mu_X, \mu_Y)^\top, \Sigma)$ with Σ being a diagonal matrix. It is possible that X and Y are not independent.

A True ■ False

Solution:
False. *(Lecture 1, slide 50)* Since Σ is a diagonal matrix, X, Y are uncorrelated. We know that components of multivariate Gaussians are independent iff they are uncorrelated.

Question 5 *(2 points)*
Let X and Y be i.i.d. Gaussian random variables distributed according to $\mathcal{N}(0, 1)$. Suppose $Z_1 = X + Y$ and $Z_2 = X - Y$. The random variables Z_1 and Z_2 are independent.

CORRECTED

■ True B False

Solution:

True. $(Z_1, Z_2)^\top = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} (X, Y)^\top$, and $(X, Y)^\top$ is multivariate Gaussian, so $(Z_1, Z_2)^\top$ is multivariate Gaussian (see *Lecture 1, slide 54*). Therefore independence of Z_1 and Z_2 is equivalent to them being uncorrelated (see *Lecture 1, slide 50*). Compute

$$\mathrm{Cov}(X + Y, X - Y) = \mathrm{Cov}(X, X) - \mathrm{Cov}(X, Y) + \mathrm{Cov}(Y, X) - \mathrm{Cov}(Y, Y) = 1 - 1 = 0$$

where we used the bi-linearity of the covariance operator.

1.1 Bayesian Linear Regression (5 points)

We want to predict a target variable $y \in \mathbb{R}$ as a linear function of an input feature vector $\boldsymbol{x} \in \mathbb{R}^2$. We consider a Bayesian linear regression model of the form

$$y = \boldsymbol{x}^\top \boldsymbol{\theta} + \epsilon,$$

where $\boldsymbol{\theta}$ represents the model parameters, and ϵ is i.i.d. Gaussian noise, with a mean of zero and a variance of 4. We choose the prior $\boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\mu}_\theta = \mathbf{0}, \boldsymbol{\Sigma}_\theta = \boldsymbol{I})$, where $\boldsymbol{I}$ is the identity matrix and $\mathbf{0} \in \mathbb{R}^2$ is the null vector.

Question 6 *(2 points)* What is the distribution of the predicted value y_* for a new input $\boldsymbol{x}_* = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$? No data has been observed yet.

A $p(y_* \mid \boldsymbol{x}_*) = \mathcal{N}(0, 4)$

B $p(y_* \mid \boldsymbol{x}_*) = \mathcal{N}(0, 5)$

C $p(y_* \mid \boldsymbol{x}_*) = \mathcal{N}(0, 13)$

■ $p(y_* \mid \boldsymbol{x}_*) = \mathcal{N}(0, 17)$

Solution:
D. For the test point $\boldsymbol{x}_*$, we can compute the distribution of y_* using the fact that sums of independent Gaussians also follow a Gaussian distribution, and analogously for multiplications with a vector (see *Lecture 1, slides 54-55*). For

$$p(y_* \mid \boldsymbol{x}_*) = p(\boldsymbol{x}_*^\top \boldsymbol{\theta} + \epsilon \mid \boldsymbol{x}_*)$$

we compute the mean as

$$\bar{y}_* = \boldsymbol{x}_*^\top \boldsymbol{\mu}_\theta + \mu_\epsilon$$

and by plugging in the numbers, we have

$$\bar{y}_* = \begin{pmatrix} 3 \\ 2 \end{pmatrix}^\top \begin{pmatrix} 0 \\ 0 \end{pmatrix} + 0 = 0.$$

Analogously, we compute the variance as

$$\text{Var}(y_*) = \text{Var}(\boldsymbol{x}_*^\top \boldsymbol{\theta}) + \text{Var}(\epsilon) = \boldsymbol{x}_*^\top \boldsymbol{\Sigma}_\theta \boldsymbol{x}_* + \sigma_\epsilon^2$$

and by plugging in the numbers, we have

$$\text{Var}(y_*) = \begin{pmatrix} 3 \\ 2 \end{pmatrix}^\top \boldsymbol{I} \begin{pmatrix} 3 \\ 2 \end{pmatrix} + 4 = 13 + 4 = 17.$$

Hence $p(y_* \mid \boldsymbol{x}_*) = \mathcal{N}(0, 17)$.

Question 7 *(3 points)* We observe a dataset $\mathcal{D} = (\boldsymbol{X}, \boldsymbol{y})$ which consists of three data points

$$\boldsymbol{X} = \begin{pmatrix} 3 & 0 \\ 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad \boldsymbol{y} = \begin{pmatrix} 5 \\ 3 \\ 4 \end{pmatrix}.$$

Assuming the same prior distribution on the model parameters as above, what is the posterior distribution of θ conditioned on this data?

■ $p(\boldsymbol{\theta}|\mathcal{D}) = \mathcal{N}(\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}})$ with $\bar{\boldsymbol{\mu}} = \begin{pmatrix} \frac{9}{7} \\ 1 \end{pmatrix}$ and $\bar{\boldsymbol{\Sigma}} = \begin{pmatrix} \frac{2}{7} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$

B $p(\boldsymbol{\theta}|\mathcal{D}) = \mathcal{N}(\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}})$ with $\bar{\boldsymbol{\mu}} = \begin{pmatrix} \frac{9}{7} \\ 1 \end{pmatrix}$ and $\bar{\boldsymbol{\Sigma}} = \begin{pmatrix} \frac{1}{11} & 0 \\ 0 & \frac{1}{5} \end{pmatrix}$

C $p(\boldsymbol{\theta}|\mathcal{D}) = \mathcal{N}(\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}})$ with $\bar{\boldsymbol{\mu}} = \begin{pmatrix} \frac{18}{11} \\ \frac{8}{5} \end{pmatrix}$ and $\bar{\boldsymbol{\Sigma}} = \begin{pmatrix} \frac{2}{7} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$

D $p(\boldsymbol{\theta}|\mathcal{D}) = \mathcal{N}(\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}})$ with $\bar{\boldsymbol{\mu}} = \begin{pmatrix} \frac{18}{11} \\ \frac{8}{5} \end{pmatrix}$ and $\bar{\boldsymbol{\Sigma}} = \begin{pmatrix} \frac{1}{11} & 0 \\ 0 & \frac{1}{5} \end{pmatrix}$

Solution:
A. *(Lecture 2, slide 18)* The posterior distribution of the model parameters is given by

$$p(\theta|\mathcal{D}) = \mathcal{N}(\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}}) \text{ with } \bar{\boldsymbol{\mu}} = (\boldsymbol{X}^T\boldsymbol{X} + \sigma_\epsilon^2\boldsymbol{I})^{-1}\boldsymbol{X}^T\boldsymbol{y} \text{ and } \bar{\boldsymbol{\Sigma}} = \left(\sigma_\epsilon^{-2}\boldsymbol{X}^T\boldsymbol{X} + \boldsymbol{I}\right)^{-1}.$$

We have

$$\boldsymbol{X}^\top\boldsymbol{X} = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} 3 & 0 \\ 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 10 & 0 \\ 0 & 4 \end{pmatrix}$$

and hence

$$(\boldsymbol{X}^\top\boldsymbol{X} + \sigma_\epsilon^2\boldsymbol{I})^{-1} = \left(\begin{pmatrix} 10 & 0 \\ 0 & 4 \end{pmatrix} + \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}\right)^{-1} = \begin{pmatrix} 14 & 0 \\ 0 & 8 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{14} & 0 \\ 0 & \frac{1}{8} \end{pmatrix}.$$

Further, we have

$$\boldsymbol{X}^\top\boldsymbol{y} = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} 5 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 18 \\ 8 \end{pmatrix}$$

and hence

$$\bar{\boldsymbol{\mu}} = \begin{pmatrix} \frac{1}{14} & 0 \\ 0 & \frac{1}{8} \end{pmatrix}\begin{pmatrix} 18 \\ 8 \end{pmatrix} = \begin{pmatrix} \frac{9}{7} \\ 1 \end{pmatrix}.$$

For the variance, we have

$$\bar{\boldsymbol{\Sigma}} = \left(\sigma_\epsilon^{-2}\boldsymbol{X}^T\boldsymbol{X} + \boldsymbol{I}\right)^{-1} = \left(\frac{1}{4}\begin{pmatrix} 10 & 0 \\ 0 & 4 \end{pmatrix} + \boldsymbol{I}\right)^{-1} = \begin{pmatrix} \frac{7}{2} & 0 \\ 0 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{2}{7} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}.$$

2 Gaussian Processes (12 points)

Question 8 *(1 point)* The polynomial function $k : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ with $k(x, x') = (1 + xx')^m$ is not a valid kernel for odd m because it can take negative values.

A True ■ False

Solution:
False. The polynomial kernel is always well-defined as per *Lecture 3, Slide 16*. A kernel does not have to be nonnegative, it has to be positive semidefinite, the latter does not imply the former.

Question 9 *(1 point)*
The mean function $\mu : \mathbb{R}^n \to \mathbb{R}$ of a Gaussian process has to be non-negative: $\mu(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in \mathbb{R}^n$.

A True ■ False

Solution:
False.
The definition in Lection 3, Slide 13 admits any μ. For example, $\mu(x) = -1$ works.

Question 10 *(1 point)* Every stationary kernel $k : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfies $k(\boldsymbol{x}, \boldsymbol{x}') = \tilde{k}(\|\boldsymbol{x} - \boldsymbol{x}'\|_2)$ for some function $\tilde{k} : \mathbb{R} \to \mathbb{R}$.

A True ■ False

Solution:
False. This is the definition of a narrower class of isotropic kernels. See Lecture 3, Slide 23.
If $\boldsymbol{x} = (x_1, x_2), \boldsymbol{x}' = (x_1', x_2') \in \mathbb{R}^2$, then $k(\boldsymbol{x}, \boldsymbol{x}') = \exp(-|x_1 - x_1'|^2 - 2|x_2 - x_2'|^2)$ is stationary since $k(\boldsymbol{x} + \boldsymbol{u}, \boldsymbol{x}' + \boldsymbol{u}) = k(\boldsymbol{x}, \boldsymbol{x}')$ for all $\boldsymbol{u} \in \mathbb{R}^2$. However, it is not isotropic since it is not a function of $\|\boldsymbol{x} - \boldsymbol{x}'\|$.

Question 11 ♣ *(4 points)*
Consider an undirected square graph, i.e. the graph $G = (V, E)$ with vertices $V = \{1, 2, 3, 4\}$ and edges $E = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 1\}\}$, as depicted below:

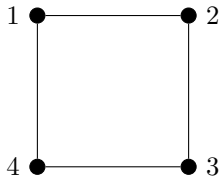


Figure 1: Corresponding to Question 11.

Define the kernel $k_l : V \times V \to \mathbb{R}$ on the vertices of G similarly to the standard squared exponential kernel, by $k_l(x, x') = \exp(-\frac{d(x,x')^2}{2l^2})$, where $d(x, x')$ is the length of the shortest path on G between x and x' and $l \neq 0$. The distance of a vertex from itself, i.e. $d(x, x)$, is zero. **Mark all that apply.**

■ For all permissible values of l, k_l is a symmetric function, i.e. $k_l(x, x') = k_l(x', x)$.

B k_l for $l = 1$ is a valid kernel, i.e. a symmetric positive semidefinite function.

C For all permissible values of l, k_l is a valid kernel, i.e. a symmetric positive semidefinite function.

Hint: $(1, -1, 1, -1)^\top$ is an eigenvector of the matrix $\mathbf{K} = \begin{pmatrix} k_l(1,1) & k_l(1,2) & k_l(1,3) & k_l(1,4) \\ k_l(2,1) & k_l(2,2) & k_l(2,3) & k_l(2,4) \\ k_l(3,1) & k_l(3,2) & k_l(3,3) & k_l(3,4) \\ k_l(4,1) & k_l(4,2) & k_l(4,3) & k_l(4,4) \end{pmatrix}$.

Solution:
A.
A is true simply because the shortest path distance is symmetric: $d(x, x') = d(x', x)$.

The fact that **C** *may* be false is mentioned in Lecture 3, Slide 30. However, it also follows from the fact that **B** is false.

To show that **B** is false, notice that

$$\mathbf{K} = \begin{pmatrix} a & b & c & b \\ b & a & b & c \\ c & b & a & b \\ b & c & b & a \end{pmatrix},$$

where $a = k_l(i, j)$ when $d(i, j) = 0$; $b = k_l(i, j)$ when $d(i, j) = 1$; $c = k_l(i, j)$ when $d(i, j) = 2$. Then

$$\mathbf{K} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = (a - 2b + c) \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix},$$

i.e. $(1, -1, 1, -1)^\top$ is an eigenvector of $\mathbf{K}$ corresponding to the eigenvalue $(a - 2b + c)$. We have

$$a - 2b + c = 1 - 2\exp\left(-\frac{1}{2l^2}\right) + \exp\left(-\frac{4}{2l^2}\right) \overset{l=1}{=} 1 - 2\exp(-1/2) + \exp(-2) \approx -0.08 < 0.$$

This means that $\mathbf{K}$ has a negative eigenvalue, thus $\mathbf{K}$ is not positive semidefinite, and thus k_l is not positive semidefinite.

2.1 Matching Kernels to Samples from the Posterior (5 points)

In Figure 2, we show samples from posterior Gaussian processes with different kernels fitted to the same dataset, with the same fixed value of observation noise. The kernels $k_i : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ considered are:

- $k_1(x, x') = \exp\left(-\frac{(x-x')^2}{2l^2}\right)$ with $l = 0.1$
- $k_2(x, x') = \exp\left(-\frac{(x-x')^2}{2l^2}\right)$ with $l = 0.4$
- $k_3(x, x') = \exp\left(-\frac{|x-x'|}{l}\right)$ with $l = 1$
- $k_4(x, x') = \exp\left(-\frac{|x-x'|}{l}\right)$ with $l = 10$
- $k_5(x, x') = (1 + xx')^m$ with $m = 2$
- $k_6(x, x') = (1 + xx')^m$ with $m = 10$

Match each of the plots in Figure 2 to the corresponding kernel. Each kernel corresponds to a single plot.

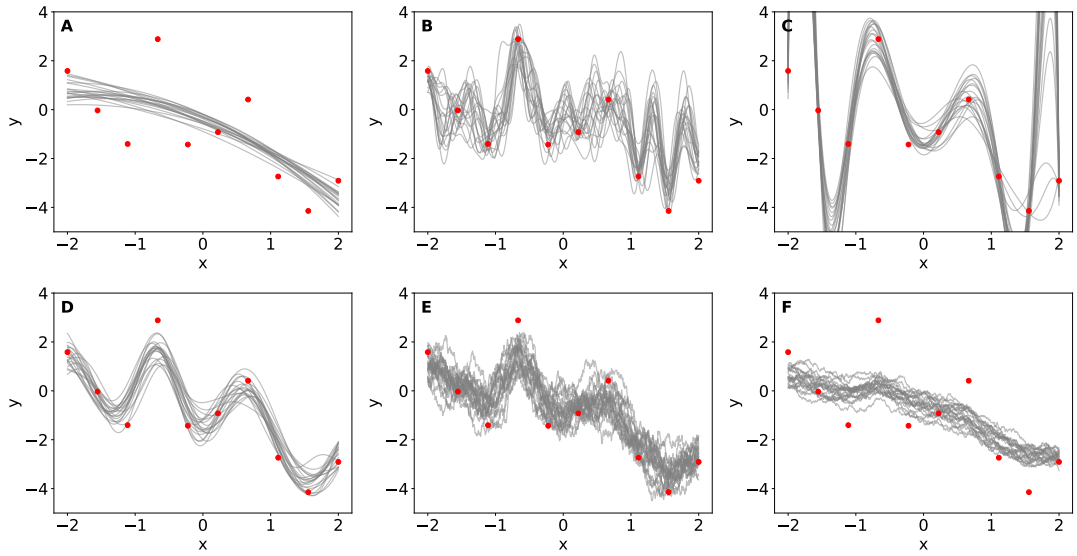


Figure 2: Samples of posterior Gaussian processes for different kernels k. The red dots represent the data.

Question 12 *(1 point)*
Which plot corresponds to k_1, the squared exponential kernel with a smaller length scale?

☐ A ■ B ☐ C ☐ D ☐ E ☐ F

Question 13 *(1 point)*
Which plot corresponds to k_2, the squared exponential kernel with a larger length scale?

☐ A ☐ B ☐ C ■ D ☐ E ☐ F

Question 14 *(1 point)*
Which plot corresponds to k_3, the exponential kernel with a smaller length scale.

☐ A ☐ B ☐ C ☐ D ■ E ☐ F

Question 15 *(1 point)*
Which plot corresponds to k_4, the exponential kernel with a larger length scale.

☐ A ☐ B ☐ C ☐ D ☐ E ■ F

Question 16 *(1 point)*
Which plot corresponds to k_5, the quadratic polynomial kernel.

■ A ☐ B ☐ C ☐ D ☐ E ☐ F

Solution:
Samples corresponding to the exponential kernel should be non-smooth (Lecture 3, Slide 21), only plots E and F fit that. The smaller length scale should correspond to a more complex model, with samples resembling a straight line less, thus k_3 must correspond to **E**. Therefore, k_4 must correspond to **F**.
The samples corresponding to the kernel k_5 should be parabola-like (Lecture 3, Slide 5), only plot **A** fits this criterion.
k_6 is the only kernel that is expected to produce extremely high and extremely low values (it corresponds to polynomial regression, for which it is a common issue), thus it should correspond to plot C, and only plots B and D remain.
The smaller length scale should correspond to a more complex model, with samples that oscillate more frequently. Because of this, k_1 corresponds to **B**. The remaining k_2 corresponds to the plot **D**.

3 Variational Inference and MCMC (12 Points)

Question 17 *(1 point)* The covariance matrix in the Laplace approximation for Bayesian logistic regression depends on the normalization constant of the true posterior.

A True ■ False

Solution:
False. *(Lecture 5, slide 14)* The covariance does not depend on the normalizer.

Question 18 *(1 point)* Let $X_1, X_2, \ldots$ be an ergodic Markov chain over a finite state space. Let T_0 be the burn-in time. Then, for a function f defined on the state space of the chain, the following holds

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} f(X_i) = \lim_{N\to\infty} \frac{1}{N - T_0} \sum_{i=T_0+1}^{N} f(X_i).$$

■ True B False

Solution:
True. *(Lecture 6, slide 41)* By the ergodic theorem, the limit on the left-hand side exists and is equal to $\mathbb{E}_{x\sim\pi}[f(x)]$, where π is the stationary distribution of the chain. Now observe that the sequence $Y_1 = X_{T_0}, Y_2 = X_{T_0+1}, \ldots$ is a Markov chain that has the same stationary distribution π (the transitions are the same, only the initial distribution is different). The limit on the right-hand side, therefore, is equal to $\mathbb{E}_{x\sim\pi}[f(x)]$ by the ergodic theorem. Notice that choosing T_0 to be the burn-in time is irrelevant, and the result holds for any T_0.

Question 19 *(2 points)* In the context of variational inference, which one of the following statements about the Evidence Lower Bound (ELBO) is true?

A Maximizing the ELBO is equivalent to minimizing the Kullback–Leibler divergence of the variational distribution from the likelihood.

B The ELBO provides an upper bound on the log marginal likelihood of the data.

■ Maximizing the ELBO is equivalent to minimizing the Kullback–Leibler divergence of the variational distribution from the posterior distribution.

D The ELBO is always positive.

Solution:
C. *(Lecture 6, slide 9)* We know that maximizing ELBO is equivalent to minimizing $\mathrm{KL}(q\|p(\cdot \mid y))$. Thus, **C** is correct. Clearly **A** is incorrect. **B** is incorrect, as ELBO gives a lower bound on the evidence ($\log p(y)$), not an upper bound. **D** is incorrect. Note that the *maximum* of ELBO is always nonnegative, but the value of ELBO for some variational distribution q might get negative. One example can be the case for Bayesian logistic regression (see example 5.18 in the notes). Letting σ_i to be very large, makes the ELBO negative.

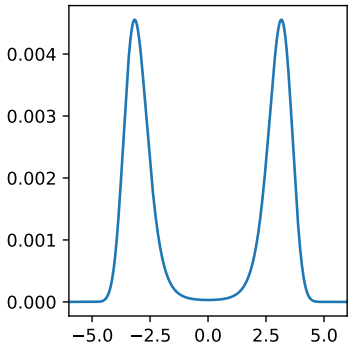


Figure 3: Related to Question 20.

Question 20 *(2 points)* Consider the "double-well" potential $f(x) = 0.05x^4 - x^2$. We use the Metropolis-adjusted Langevin Algorithm (MALA) to sample from the distribution $\pi(x) \propto e^{-f(x)}$ shown in Figure 3. Which of the following is true?

A MALA with the step-size $\tau = 0.01$ does not converge to π.

■ MALA with the step-size $\tau = 5$ converges to the correct distribution π.

Solution:
B. *(Lecture 7, slide 4)* Since MALA (with any step-size) satisfies the detailed balance equation for π (by construction), it would converge to π and does not get stuck indefinitely or overshoot. Note that these effects might happen in finite time, but eventually, the Markov chain mixes and converges to π.

3.1 Sampling from the Cube (6 Points)

Let $n \geq 4$ and $S = \{(x_1, \ldots, x_n) \mid x_i \in \{0, 1\},\ \sum_{i=1}^n x_i \leq 2\} \subset \{0, 1\}^n$ be the set of all n-bit strings with at most two 1-s. Suppose we want to sample from the probability distribution p defined as

$$p(x_1, \ldots, x_n) = \begin{cases} \frac{1}{Z} \exp(\sum_{i=1}^n x_i w_i) & \text{if } (x_1, \ldots, x_n) \in S \\ 0 & \text{otherwise,} \end{cases}$$

where $w_i \in \mathbb{R}$ are fixed weights and Z is a normalization constant that makes p a probability distribution.

Question 21 *(1 point)* Suppose the random vector $\mathbf{X}$ is distributed according to p. Denote the i-th component of the vector $\mathbf{X}$ by X_i. The random variables $X_1, \ldots, X_n$ are independent.

A True ■ False

Solution:
False. If X_i were independent, p should have been decomposed into a product:

$$p(x_1, \ldots, x_n) = p_1(x_1) \cdots p_n(x_n).$$

Observe that $p(x_1, \ldots, x_n)$ is nonzero for any $(x_1, \ldots, x_n) \in S$. Then, $p(1, 1, 0, \ldots) > 0$, implying $p_1(1) > 0$. Similarly, from $p(0, 1, 1, 0, \ldots) > 0$ we get $p_1(0) > 0$. With a similar argument, we infer that $p_i(0), p_i(1) > 0$. Therefore, $p(1, 1, \ldots, 1) = p_1(1) \cdots p_n(1) > 0$, a contradiction.

Consider the following process for sampling from p: Starting from an arbitrary $\boldsymbol{x}^{(0)} \in S$, at each iteration $t = 1, 2, \ldots$ we select an index $i \in \{1, \ldots, n\}$ uniformly at random, and

- if $\boldsymbol{x}_i^{(t-1)} = 1$, set $\boldsymbol{y}$ to be the same as $\boldsymbol{x}^{(t-1)}$ with the ith index flipped to 0 with probability A_i.
- if $\boldsymbol{x}_i^{(t-1)} = 0$, set $\boldsymbol{y}$ to be the same as $\boldsymbol{x}^{(t-1)}$ with the ith index flipped to 1 with probability B_i.
- if $\boldsymbol{y} \in S$, set $\boldsymbol{x}^{(t)} = \boldsymbol{y}$, otherwise, set $\boldsymbol{x}^{(t)} = \boldsymbol{x}^{(t-1)}$.

Here, $0 < A_i \leq 1$ and $0 < B_i \leq 1$ are parameters that will be defined later for all $i \in \{1, \ldots, n\}$.

Question 22 *(2 points)* The process above describes an ergodic Markov chain.

■ True B False

Solution:
True. *(Lecture 6, slide 28)* We have to show that for some t, we can reach from each state to any state with exactly t steps. We claim that $t = 4$ works for this chain. Let us decompose the states into three groups: G_2 are those states with exactly two bits set to 1, G_1 is those with exactly one bit set to 1, and $G_0 = \{(0, 0, \ldots, 0)\}$.
Note that regardless of A, B, there is always the possibility to stay at a state in G_2 (choosing a coordinate other than the two bits that are 1, the chain keeps the state unchanged as otherwise it would lead to a state outside of S). Also note that it is always possible to go from G_2 to G_1 and vice versa. Same for G_1 to G_0 and vice versa.
To go from a state in G_2 to another in G_2, we go to a neighbor in G_1 and then G_0 and then go back up to G_1 and to G_2. To go from a state in G_2 to one in G_1, we first stay where we are, then go to a neighbor in G_1 and then G_0 and then go back up to G_1. Other cases can be verified similarly.

Question 23 *(3 points)* Suppose the Markov chain above satisfies the detailed balance equation for p. Which of the following are values that A_i and B_i could take, such that the chain admits p as its stationary distribution for all possible values of $w_i \in \mathbb{R}$?

A $A_i = B_i = \frac{1}{2}$

■ $A_i = \dfrac{1}{1 + e^{w_i}}$ and $B_i = \dfrac{e^{w_i}}{1 + e^{w_i}}$

C $A_i = e^{-w_i}$ and $B_i = 1$

D $A_i = 1$ and $B_i = e^{w_i}$

E $A_i = \dfrac{1}{1 + w_i}$ and $B_i = \dfrac{w_i}{1 + w_i}$

Solution:
B. *(Lecture 6, slides 29 and 32)* Let $\boldsymbol{x}, \boldsymbol{x'} \in S$ be two elements of S with $\boldsymbol{x}_i = 1$ and $\boldsymbol{x}'_i = 0$ for some i while $\boldsymbol{x}_j = \boldsymbol{x}'_j$ for all $j \neq i$. Then,

$$p(\boldsymbol{x}) \cdot p(\boldsymbol{x} \to \boldsymbol{x'}) = p(\boldsymbol{x'}) \cdot p(\boldsymbol{x'} \to \boldsymbol{x})$$

which implies

$$e^{w_i} \cdot \frac{1}{n} \cdot A = \frac{1}{n} \cdot B$$

Among the choices, only option B satisfies this property and ensures that $0 < A_i, B_i \leq 1$. (Recall that $w_i \in \mathbb{R}$, thus e^{w_i} or e^{-w_i} might be more than 1). Note also that satisfying the detailed balance equation for p is a sufficient condition for having p as a stationary distribution.

4 Bayesian Deep Learning and Active Learning (10 points)

Question 24 *(2 points)*
Consider a Bayesian optimization problem of maximizing an unknown function on a discrete domain $\{x_1, x_2, x_3, x_4\}$. At each round we select a point x from this domain and observe the value $f(x) + \epsilon$ where ϵ is Gaussian noise and f is an unknown function. Suppose at some round t, the posterior distribution of the vector $\big(f(x_1), f(x_2), f(x_3), f(x_4)\big)^T$ is a multivariate normal distribution $\mathcal{N}(\mu_t, \Sigma_t)$ where

$$\mu_t = \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \end{pmatrix}, \Sigma_t = \begin{pmatrix} 9 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 1 & 16 \end{pmatrix}.$$

Which input would we choose at this round if we use the upper confidence bound algorithm with $\beta_t = 1$?
Hint: Recall that the upper confidence bound algorithm selects the input x with maximum upper confidence bound: $\mu_t(x) + \beta_t \sigma_t(x)$ where $\mu_t(x)$ and $\sigma_t(x)$ are the posterior mean and posterior marginal standard deviation of $f(x)$ respectively.

■ x_1 B x_2 C x_3 D x_4

Solution:
A. *(Lecture 8, Slide 41)*To apply the upper confidence bound algorithm, for each point we compute the UCB: $\mu_t(x) + \beta_t \sigma_t(x)$ and then select the point with largest UCB. Note that $\sigma_t(x_i) = \sqrt{[\Sigma_t]_{i,i}}$, or the square root of the ith element of the diagonal of Σ_t. Plugging in we get UCBs of 4, 3, 2, 3 for x_1, x_2, x_3, x_4 respectively and so the point selected is x_1.

4.1 Properties of Information Gain (2 points)

Recall that for random variables X, Y, the information gain between X and Y is given by

$$I(X; Y) = H(X) - H(X \mid Y)$$

where $H(X)$ is the entropy of X and $H(X \mid Y)$ is the conditional entropy of X given the observation of Y. Now given a new random variable Z, we similarly define the information gain between X and (Y, Z) as

$$I(X; Y, Z) = H(X) - H(X \mid Y, Z).$$

Question 25 *(1 point)*
It holds that $I(X; Y) \geq 0$.

■ True B False

Solution:
True. This can be seen since the information gain between two random variables x and y with joint density $p(x, y)$ can be expressed as $I(x; y) = \mathrm{KL}(p(x, y) \| p(x)p(y))$. You have shown in homework 3, exercise 1b that a KL-divergence is non-negative.

Question 26 *(1 point)*
It holds that $I(X; Y, Z) - I(X; Y) \geq 0$.

■ True B False

Solution:
True. Lecture 8, Slide 13. We have

$$I(X; Y, Z) - I(X \mid Y) = H(X \mid Y) - H(X \mid Y, Z).$$

This is non-negative iff $H(X \mid Y) \geq H(X \mid Y, Z)$ which is the "information never hurts" property of conditional entropy.

4.2 Bayesian Deep Learning (6 points)

We are modeling a binary classification problem with a 2-layer Bayesian neural network where the probability that the binary label $y \in \{0, 1\}$ of input $\boldsymbol{x}$ is equal to one is

$$p(y = 1 \mid \boldsymbol{w}, \boldsymbol{W}, \boldsymbol{x}) = \sigma(\boldsymbol{w}^\top \sigma(\boldsymbol{W}\boldsymbol{x}))$$

where σ is the element-wise applied sigmoid activation function defined as $\sigma(z) = (1 + \exp(-z))^{-1}$. The network depends on two *unknown* parameters $\boldsymbol{w} \in \mathbb{R}^2$ and $\boldsymbol{W} \in \mathbb{R}^{2\times 3}$. We use the following priors:

$$\boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, 0.5 \cdot \boldsymbol{I}), \qquad \boldsymbol{W}_{ij} \sim \mathcal{N}(0, 0.5)$$

where all entries $\boldsymbol{W}_{ij}$ and $\boldsymbol{w}$ are independent of each other and $\boldsymbol{I}$ is the 2-dimensional identity matrix.

Question 27 *(3 points)*
Given the i.i.d. dataset $\{\boldsymbol{x}_i, y_i\}_{i=1}^{100}$, which of the following optimization problems finds the maximum a posteriori (MAP) estimate of the model parameters $\boldsymbol{w}$ and $\boldsymbol{W}$?
Note: Recall that the Frobenius matrix norm is defined via $\|\boldsymbol{A}\|_F^2 = \sum_{i,j} \boldsymbol{A}_{ij}^2$, while $\|\cdot\|_2$ denotes the spectral norm (the largest singular value of a matrix and for vectors the Euclidean norm).

A $\arg\min_{\boldsymbol{w},\boldsymbol{W}} \|\boldsymbol{w}\|_2^2 + \|\boldsymbol{W}\|_F^2 - \frac{1}{100}\sum_{i=1}^{100} y_i \log(\sigma(\boldsymbol{w}^\top\sigma(\boldsymbol{W}\boldsymbol{x}_i))) + (1-y_i)\log(1-\sigma(\boldsymbol{w}^\top\sigma(\boldsymbol{W}\boldsymbol{x}_i)))$

■ $\arg\min_{\boldsymbol{w},\boldsymbol{W}} \|\boldsymbol{w}\|_2^2 + \|\boldsymbol{W}\|_F^2 - \sum_{i=1}^{100} y_i \log(\sigma(\boldsymbol{w}^\top\sigma(\boldsymbol{W}\boldsymbol{x}_i))) + (1-y_i)\log(1-\sigma(\boldsymbol{w}^\top\sigma(\boldsymbol{W}\boldsymbol{x}_i)))$

C $\arg\min_{\boldsymbol{w},\boldsymbol{W}} \|\boldsymbol{w}\|_2^2 + \|\boldsymbol{W}\|_F^2 - \frac{1}{100}\sum_{i=1}^{100} y_i \log(\sigma(\boldsymbol{w}^\top\sigma(\boldsymbol{W}\boldsymbol{x}_i)))$

D $\arg\min_{\boldsymbol{w},\boldsymbol{W}} \|\boldsymbol{w}\|_2^2 + \|\boldsymbol{W}\|_F^2 - \sum_{i=1}^{100} y_i \log(\sigma(\boldsymbol{w}^\top\sigma(\boldsymbol{W}\boldsymbol{x}_i)))$

E $\arg\min_{\boldsymbol{w},\boldsymbol{W}} \|\boldsymbol{w}\|_2^2 + \|\boldsymbol{W}\|_2^2 - \sum_{i=1}^{100} y_i \log(\sigma(\boldsymbol{w}^\top\sigma(\boldsymbol{W}\boldsymbol{x}_i))) + (1-y_i)\log(1-\sigma(\boldsymbol{w}^\top\sigma(\boldsymbol{W}\boldsymbol{x}_i)))$

Solution:
B. The solution follows by calculating the likelihood times the prior probability and taking the negative log (lecture 7, slide 25):

$$\arg\min_{\boldsymbol{w},\boldsymbol{W}} -\log p(\boldsymbol{w}, \boldsymbol{W}) - \sum_{i=1}^{100} \log p(y_i \mid \boldsymbol{x}_i, \boldsymbol{w}, \boldsymbol{W}).$$

The negative log-prior is proportional to the Euclidean norm of the parameter vector, which in the case of matrices is the Frobenius norm. That is,

$$-\log p(\boldsymbol{w}, \boldsymbol{W}) = \|\boldsymbol{w}\|_2^2 + \|\boldsymbol{W}\|_F^2.$$

The log-likelihood of data point $(\boldsymbol{x}_i, y_i)$ is

$$\log p(y_i \mid \boldsymbol{x}_i, \boldsymbol{w}, \boldsymbol{W}) = y_i \log(\sigma(\boldsymbol{w}^\top\sigma(\boldsymbol{W}\boldsymbol{x}_i))) + (1-y_i)\log(1-\sigma(\boldsymbol{w}^\top\sigma(\boldsymbol{W}\boldsymbol{x}_i))).$$

Question 28 *(3 points)*
We choose to approximate the posterior of $\boldsymbol{w}$ by a Laplace approximation $\boldsymbol{w} \sim \mathcal{N}(\hat{\boldsymbol{w}}, \boldsymbol{\Lambda}^{-1})$ with MAP estimate $\hat{\boldsymbol{w}}$ and precision matrix $\boldsymbol{\Lambda}$. Which of the following represents the correct precision matrix?
Note: $\ell(y_i \mid \boldsymbol{w}, \boldsymbol{W}, \boldsymbol{x}_i)$ denotes the negative log-likelihood of data point $(\boldsymbol{x}_i, y_i)$ and $\boldsymbol{H}_{\boldsymbol{w}} f(\boldsymbol{w})|_{\boldsymbol{w}=\hat{\boldsymbol{w}}}$ denotes the Hessian of a function f with respect to $\boldsymbol{w}$ evaluated at $\hat{\boldsymbol{w}}$.

A $\boldsymbol{\Lambda} = 2\boldsymbol{I} + \frac{1}{100}\sum_{i=1}^{100} \boldsymbol{H}_{\boldsymbol{w}}\, \ell(y_i \mid \boldsymbol{w}, \boldsymbol{W}, \boldsymbol{x}_i)|_{\boldsymbol{w}=\hat{\boldsymbol{w}}}$

B $\boldsymbol{\Lambda} = \frac{1}{100}\sum_{i=1}^{100} \boldsymbol{H}_{\boldsymbol{w}}\, \ell(y_i \mid \boldsymbol{w}, \boldsymbol{W}, \boldsymbol{x}_i)|_{\boldsymbol{w}=\hat{\boldsymbol{w}}}$

■ $\boldsymbol{\Lambda} = 2\boldsymbol{I} + \sum_{i=1}^{100} \boldsymbol{H}_{\boldsymbol{w}}\, \ell(y_i \mid \boldsymbol{w}, \boldsymbol{W}, \boldsymbol{x}_i)|_{\boldsymbol{w}=\hat{\boldsymbol{w}}}$

D $\boldsymbol{\Lambda} = \sum_{i=1}^{100} \boldsymbol{H}_{\boldsymbol{w}}\, \ell(y_i \mid \boldsymbol{w}, \boldsymbol{W}, \boldsymbol{x}_i)|_{\boldsymbol{w}=\hat{\boldsymbol{w}}}$

E $\boldsymbol{\Lambda} = -2\boldsymbol{I} - \sum_{i=1}^{100} \boldsymbol{H}_{\boldsymbol{w}}\, \ell(y_i \mid \boldsymbol{w}, \boldsymbol{W}, \boldsymbol{x}_i)|_{\boldsymbol{w}=\hat{\boldsymbol{w}}}$

F $\boldsymbol{\Lambda} = -\sum_{i=1}^{100} \boldsymbol{H}_{\boldsymbol{w}}\, \ell(y_i \mid \boldsymbol{w}, \boldsymbol{W}, \boldsymbol{x}_i)|_{\boldsymbol{w}=\hat{\boldsymbol{w}}}$

Solution:
C. Lecture 5, slide 14. The precision matrix of the Laplace approximation is

$$\begin{aligned}\boldsymbol{\Lambda} &= -\boldsymbol{H}_{\boldsymbol{w}} \log p(\boldsymbol{w} \mid \boldsymbol{x}_{1:n}, y_{1:n}, \boldsymbol{W})|_{\boldsymbol{w}=\hat{\boldsymbol{w}}} \\ &= -\boldsymbol{H}_{\boldsymbol{w}} \log p(y_{1:n} \mid \boldsymbol{x}_{1:n}, \boldsymbol{w}, \boldsymbol{W})|_{\boldsymbol{w}=\hat{\boldsymbol{w}}} - \boldsymbol{H}_{\boldsymbol{w}} \log p(\boldsymbol{w})|_{\boldsymbol{w}=\hat{\boldsymbol{w}}}.\end{aligned}$$

The solution follows by observing $-\log p(y_{1:n} \mid \boldsymbol{x}_{1:n}, \boldsymbol{w}, \boldsymbol{W}) = \ell(y_i \mid \boldsymbol{w}, \boldsymbol{W}, \boldsymbol{x}_i)$, using independence of data, and plugging in the prior $-\boldsymbol{H}_{\boldsymbol{w}} \log p(\boldsymbol{w})|_{\boldsymbol{w}=\hat{\boldsymbol{w}}} = 2\boldsymbol{I}$.

5 MDPs and Reinforcement Learning (15 points)

Question 29 ♣ *(2 points)*
Which of the following algorithms, as taught in the course, are off-policy RL algorithms? **Mark all that apply.**

- ■ Deep Q-Networks (DQN)
- ■ Soft Actor Critic (SAC)
- C PPO (Proximal Policy Optimization)
- ■ Deep Deterministic Policy Gradients (DDPG)
- E REINFORCE
- F Advantage Actor Critic (A2C)

Solution:
A, B, D. Lecture 12, Slides 15, 16, 22, 32, 49. On-policy algorithms (as opposed) to off-policy algorithms require sampling from the true underlying MDP when they are searching for optimal policy.
E: This algorithm requires Monte Carlo samples from the trajectory. C: The same as above, just with added regularization. F: Use monte Carlo samples to estimate advantage function.

Question 30 ♣ *(2 points)* Which of the following algorithms are based on the "optimism in the face of uncertainty" principle? **Mark all that apply.**

■ R-Max B PETS C DDPG ■ H-UCRL

Solution:
A, D. Optimistic algorithm design an optimal policy for the environment (transitions, rewards, Q-function) that is highest possible given the past observations. An optimistic algorithm can be both model-free (optimistic w.r.t to Q function) or model-based (transition and rewards).

Using the logic we can infer:
R-Max: Initializes all reward to be very high i.e. **optimistic** to ensure they are visited. (Lecture 10, Slide 17).
PETS: The resulting sequence of actions optimizes **average performance** by averaging over the sampled trajectories (Lecture 13, Slide 39-40.).
DDPG: The policy of this actor-critic algorithm is chosen **greedily** according to the current estimate of the Q-function. The final executed action is perturbed with noise (Lecture 12, Slide 32.).
H-UCRL: This is an algorithm that you saw in model-based RL. The algorithm searches over the model space and chooses an action that **maximizes reward over the possible model space** (Lecture 13, Slide 46.).

Question 31 ♣ *(3 points)* Consider an RL problem with reward $r(x, a)$, where $x \in \mathbb{R}, a \in \mathbb{R}$, are the state and action, respectively. We roll out a parametric policy $\pi_\theta(a|x)$ for T steps and denote with $\tau \sim \pi_\theta$ the resulting trajectory detailed as $\tau = \{x_0, a_0, x_1, a_1, \ldots, x_T, a_T\}$. The expected sum of rewards under π_θ is defined as $J(\theta) = \mathbb{E}_{\tau\sim\pi_\theta}[\sum_{t=0}^{T} r(x_t, a_t)]$. Which of the following is equal to $\nabla J(\theta)$, the gradient of this objective with respect to θ? Suppose $b \leq 0$ is a constant and **mark all that apply.**

- A $\mathbb{E}_{\tau\sim\pi_\theta}[\sum_{t=0}^{T}(\sum_{t'=0}^{T} r(x_{t'}, a_{t'}))\nabla\pi_\theta(a_t|x_t)]$
- ■ $\mathbb{E}_{\tau\sim\pi_\theta}[\sum_{t=0}^{T}(\sum_{t'=0}^{T} r(x_{t'}, a_{t'}))\nabla \log\pi_\theta(a_t|x_t)]$
- C $\nabla\mathbb{E}_{\tau\sim\pi_\theta}[\sum_{t=0}^{T}(\sum_{t'=0}^{T} r(x_{t'}, a_{t'}))\log\pi_\theta(a_t|x_t)]$
- D $\mathbb{E}_{\tau\sim\pi_\theta}[\sum_{t=0}^{T}(\sum_{t'=0}^{T} r(x_{t'}, a_{t'}) - b)\nabla\pi_\theta(a_t|x_t)]$

■ $\mathbb{E}_{\tau\sim\pi_\theta}[\sum_{t=0}^{T}(\sum_{t'=0}^{T} r(x_{t'},a_{t'})-b)\nabla\log\pi_\theta(a_t|x_t)]$

F $\nabla\mathbb{E}_{\tau\sim\pi_\theta}[\sum_{t=0}^{T}(\sum_{t'=0}^{T} r(x_{t'},a_{t'})-b)\log\pi_\theta(a_t|x_t)]$

■ $\mathbb{E}_{\tau\sim\pi_\theta}[\sum_{t=0}^{T}(\sum_{t'=t}^{T} r(x_{t'},a_{t'}))\nabla\log\pi_\theta(a_t|x_t)]$

H $\mathbb{E}_{\tau\sim\pi_\theta}[\sum_{t=0}^{T}(\sum_{t'=0}^{t-1} r(x_{t'},a_{t'}))\nabla\log\pi_\theta(a_t|x_t)]$

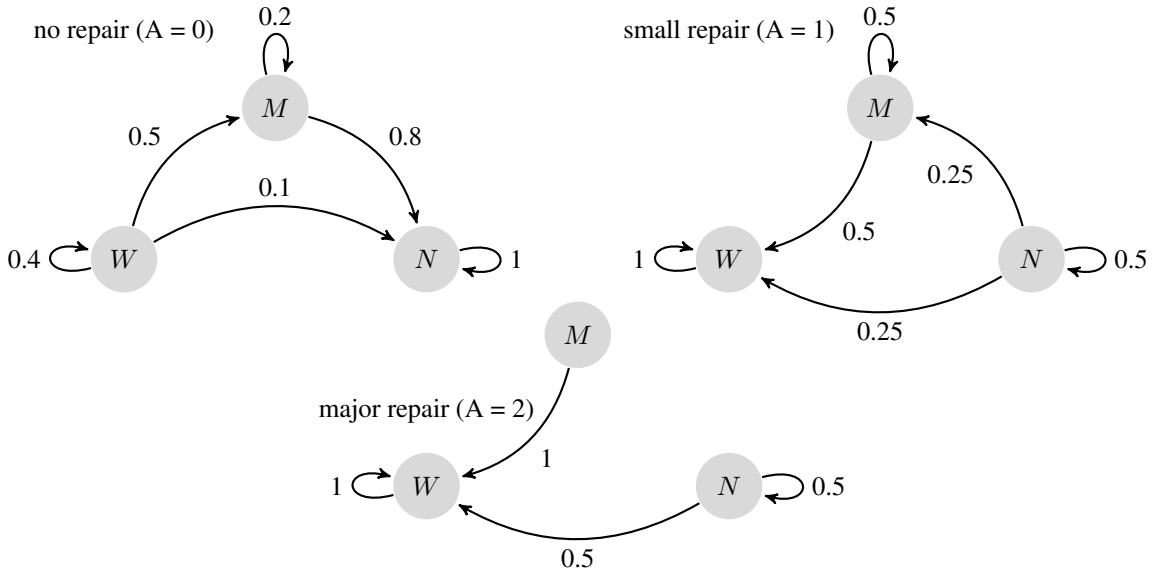
Solution:
B, E, G. (See Lecture 11, Slides: 32, 35, 37).
The solution is directly inferred from the slides by setting $\gamma = 1$. Choice (B) is the policy gradient identity.
In slides 35-37 we prove that any b, independent of (x'_t, a'_t) for $t' \geq t$ may be subtracted from the reward, leading to the correctness of choice (E).
Setting $b = \sum_{t'=0}^{t-1} r(x_{t'}, a_{t'})$ yields the choice (G).

Other choices (A), (D) feature a wrong-parameterization of policy. Options (C) and (F) take a gradient of a function that is not a value function. Option (H) is not correct, since the possible b baseline that could result in this update depends on the future, which is not allowed.

5.1 Markov Decision Processes (8 points)

Consider a machine that can exist in one of three states: *working* (W), *maintenance needed* (M), and *not working* (N). Our goal is to minimize the time that the machine spends in the *not working* state while avoiding unnecessary repair costs.

We develop a policy for repairing the machine at times $t = 1, 2, 3, \ldots$, given the state of the machine at this time. Three options are available: no repairs ($A = 0$), minor repairs ($A = 1$), or a major repair ($A = 2$). The three diagrams below show the transition probabilities between states when we play each of the respective actions $A = 0, 1$ or 2. The action corresponding to each diagram is written next to it.



Question 32 *(1 points)* Suppose we design a policy that performs a major repair at $t = 2$, but does no repairs at $t = 1, 3, 4, \ldots$, regardless of the state at that time. Is this policy a stationary policy?

A True ■ False

Solution:
False. The policy is **not** stationary as it has different behaviors at different times. For instance, there is a non-zero chance for the agent to remain in the same state (W) for $t = 1, 2$: in this case, it would first take action $a_1 = 0$, and then action $a_2 = 2$, and the policy would need to depend on t. You saw the definition of non-stationary Markov chains in Lecture 6, slide 27 and Tutorial 9. The policy results in non-stationary Markov chain hence its non-stationary.

Question 33 *(3 points)* Assume that at $t = 1$ the machine is in working condition. What is the probability that the machine is in the *maintenance needed* (M) state once the above policy is applied for 3 rounds?

A 0.145 B 0.38 C 0.5 ■ 0.475

Solution:
D. The transition matrix without repair is $P_0 = \begin{pmatrix} 0.4 & 0 & 0 \\ 0.5 & 0.2 & 0 \\ 0.1 & 0.8 & 1 \end{pmatrix}$, and $P_2 = \begin{pmatrix} 1 & 1 & 0.5 \\ 0 & 0 & 0 \\ 0 & 0 & 0.5 \end{pmatrix}$ with major overhaul. The solution is obtained by applying $P_0 P_2 P_0$ to $b_0 = (1, 0, 0)$, and reading out the middle coefficient equal to $0.95/2$

We model our objective of having a working machine at low repairs cost via the reward function

$$R(s,a) = R_1(s) + R_2(a)$$

where s is the state, a is the action and

$$R_1(s) = \begin{cases} -6 & s = N \\ 0 & s \in \{W, M\} \end{cases}, \qquad R_2(a) = \begin{cases} -5 & a = 2 \\ -2.5 & a = 1 \\ 0 & a = 0 \end{cases}$$

Question 34 *(4 points)* Assume a discount factor $\gamma = 0.5$. Assume that the optimal value function that we obtained by value iteration is $V^*(W) = -2, V^*(M) = -4, V^*(N) = -12$. Given this, which of the following is a corresponding optimal policy? Note that the function A^* : state $\rightarrow$ action, assigns optimal actions to each state.

■ $A^*(W) = 0, A^*(M) = 1, A^*(N) = 0$

B $A^*(W) = 0, A^*(M) = 1, A^*(N) = 1$

C $A^*(W) = 0, A^*(M) = 0, A^*(N) = 2$

D $A^*(W) = 0, A^*(M) = 1, A^*(N) = 2$

E $A^*(W) = 0, A^*(M) = 0, A^*(N) = 1$

Solution:
A. *(Lecture 11, slide 3.)*The optimal policy is the one which leads to the optimal value function as a function of actions. In other words, $a^* = \arg\max_a Q^*(s, a)$. First, we calculate, the optimal Q function as,

$$Q^*(s, a) = R(s, a) + \gamma \sum_{s'} P(s'|s, a)V^*(s').$$

Now, for each $s \in \{W, M, N\}$, we can calculate the largest a: $[-2., -3.5, -6.], [-5.2, -4., -6.], [-12., -12.25, -14.5]$ for states W, M, N, respectively. This is maximized, respectively, by actions $0, 1, 0$.

CORRECTED

Answer Sheet of the Probabilistic Artificial Intelligence 2024/25 Exam

0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9

⟵ Please encode your student number on the left, and write your first and last names below.

Firstname and Lastname:

. . . . . .

Question 1: ■ B
Question 2: ■ B
Question 3: A ■
Question 4: A ■
Question 5: ■ B
Question 6: A B C ■
Question 7: ■ B C D
Question 8: A ■
Question 9: A ■
Question 10: A ■
Question 11: ■ B C
Question 12: A ■ C D E F
Question 13: A B C ■ E F
Question 14: A B C D ■ F
Question 15: A B C D E ■
Question 16: ■ B C D E F
Question 17: A ■
Question 18: ■ B
Question 19: A B ■ D
Question 20: A ■
Question 21: A ■
Question 22: ■ B
Question 23: A ■ C D E
Question 24: ■ B C D
Question 25: ■ B
Question 26: ■ B
Question 27: A ■ C D E
Question 28: A B ■ D E F
Question 29: ■ ■ C ■ E F
Question 30: ■ B C ■
Question 31: A ■ C D ■ F ■ H
Question 32: A ■
Question 33: A B C ■
Question 34: ■ B C D E