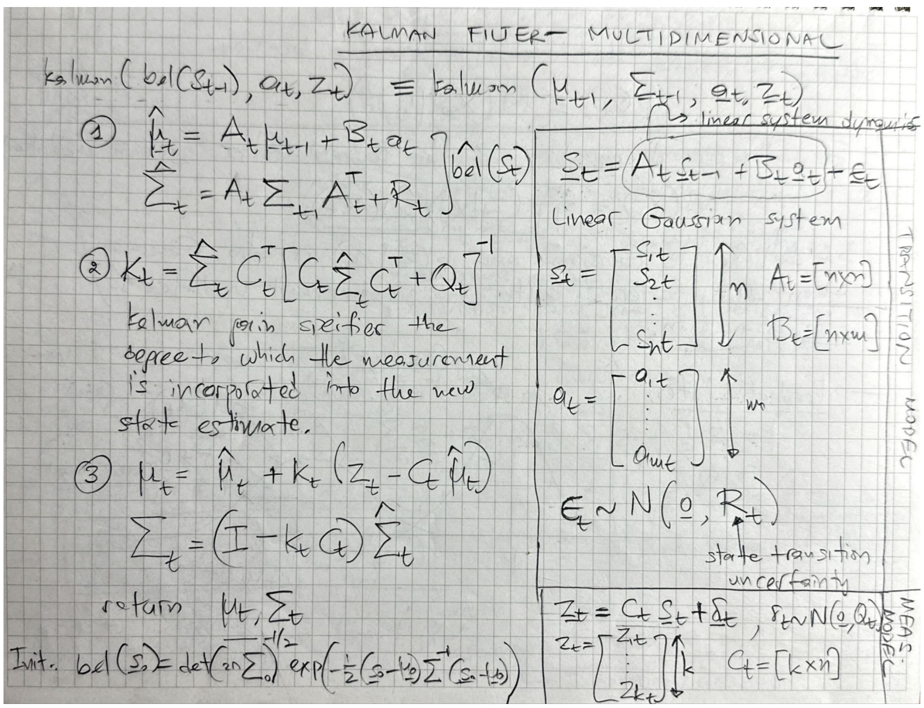
Introduction to AI

NYU Spring 2025 In-Person Final Exam

QUESTION SET 1 (15 points)


KALMAN FILTER — MULTIDIMENSIONAL

$\text{kalman}(bel(s_{t-1}), a_t, z_t) \equiv \text{kalman}(\mu_{t-1}, \Sigma_{t-1}, a_t, z_t)$

① $\hat{\mu}_t = A_t \mu_{t-1} + B_t a_t$

$\hat{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t$ $\Big\} \; \hat{bel}(s_t)$

② $K_t = \hat{\Sigma}_t C_t^T \left[C_t \hat{\Sigma}_t C_t^T + Q_t \right]^{-1}$

Kalman gain specifies the degree to which the measurement is incorporated into the new state estimate.

③ $\mu_t = \hat{\mu}_t + K_t (z_t - C_t \hat{\mu}_t)$

$\Sigma_t = (I - K_t C_t) \hat{\Sigma}_t$

return μ_t, Σ_t

Init. $bel(s_0) = \det(2\pi\Sigma_0)^{-1/2} \exp\left(-\frac{1}{2}(s_0-\mu_0)^T \Sigma^{-1} (s_0-\mu_0)\right)$

linear system dynamics

$s_t = A_t s_{t-1} + B_t a_t + \epsilon_t$

Linear Gaussian system

$s_t = \begin{bmatrix} s_{1t} \\ s_{2t} \\ \vdots \\ s_{nt} \end{bmatrix} n \qquad A_t = [n \times n] \qquad B_t = [n \times m]$

$a_t = \begin{bmatrix} a_{1t} \\ \vdots \\ a_{mt} \end{bmatrix} m$

$\epsilon_t \sim N(0, R_t)$

state transition uncertainty

TRANSITION MODEL

$z_t = C_t s_t + \delta_t, \quad \delta_t \sim N(0, Q_t)$

$z_t = \begin{bmatrix} z_{1t} \\ \vdots \\ z_{kt} \end{bmatrix} k \qquad C_t = [k \times n]$

MEAS. MODEL

Which of the following statements regarding the system described above is / are correct ?

A. If $C_t = \mathbf{0}, \forall t$, then it is not possible to compute an estimate of the state using the Kalman Filter.

B. The Kalman Filter gain K_t depends on the measurements $z_{1:t}$.

C. None of the provided answers is correct.

QUESTION SET 2 (25 points)

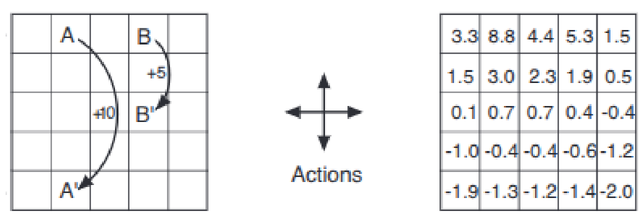
QUESTION SET 2A (10 points)

In the following gridworld environment, the agent can move in four directions: south, west, east, north. The agent receives a reward of +1 for reaching the terminal state (1,4) and -1 for reaching the terminal state (2,4). The walls are represented by "wall" and the empty cells are represented by their coordinates.

	1	2	3	4
1	(1,1)	(1,2)	(1,3)	+1 (terminal)
2	(2,1)	wall	(2,3)	-1 (terminal)
3	(3,1)	(3,2)	(3,3)	(3,4)

What is the number of possible deterministic policies for this grid-world? Explain your answer.

QUESTION SET 2B (15 points)



In the above figure, the gridworld environment is a 5x5 grid that is deterministic.
- The agent can move in 4 directions: North, South, East, West.

- The agent operates under a uniform random policy — the probability of each action at each state is 0.25.
- If a move would take the agent off the grid, it remains in the same state.
- Actions that would take the agent into a wall result in a -1 reward and the agent remains in the same state. Other actions result in a reward of 0 indepedent of the state unless the agent is in a special state as described below.
- Discount factor: $\gamma = 0.9$ - this is just a value used to produce the numbers shown in the figure.
- Two special states:
  - A: when the agent is in state A and chooses any action, it is instantly transported to state A' = A $\rightarrow$ teleport to state A' with reward 10.
  - B: when the agent is in state B and chooses any action, it is transported to B' = B $\rightarrow$ teleport to state B' with reward 5.

The value function $v(s)$ for all states under the random policy, using the Bellman equation for policy evaluation is given by:

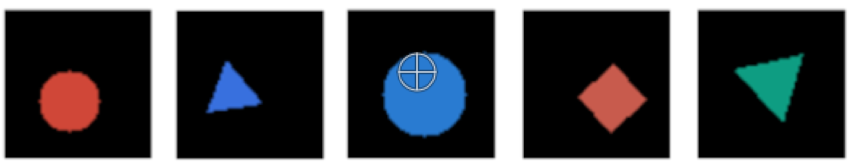
$$v(s) = \sum_a \pi(a|s) \sum_{s'} p(s'|s,a) \left[r(s,a,s') + \gamma v(s') \right]$$

Explain using the equation (this is not a numerical exercise) why the value of state A is less than its immediate reward of 10 while the value of state B is greater than its immediate reward of 5.

QUESTION SET 3 (25 points)

You are given images of geometrical shapes and description of the images in natural language e.g for the shapes below these are:

“red circle”, “blue triangle”, “blue circle”, “red square”, “gren triangle”



You can assume you have a dataset of size $h \times w$ and their aligned descriptions of size s tokens.

QS3A (15 points)

Draw an architecture of a system that can be used to learn the mapping between the images and the descriptions.

QS3B (10 points)

Explain clearly the equation of the loss function and the training process.

QUESTION SET 4 (20 points)

Explain the advantages of llama3-8B over llama2-7B for each of the following features. `B` stands for billion parameters.

Feature	LLaMA 2	LLaMA 3
Model Sizes	7B, 13B, 70B	8B, 70B, 405B
Layers	32 (7B), 40 (13B), 80 (70B)	32 (8B), 80 (70B), 126 (405B)
Latent Size	4096 (7B), 5120 (13B), 8192 (70B)	4096 (8B), 8192 (70B), 16,384 (405B)
Attention Heads	32 (7B), 40 (13B), 64 (70B)	32 (8B), 64 (70B), 128 (405B)
FF Dimension	11,008 (7B), 13,312 (13B), 22,528 (70B)	14,336 (8B), 28,672 (70B), 53,248 (405B)
Vocabulary Size	32,000	128,000
Context Length	4,096 tokens	8,192 tokens
Training Tokens	~2 trillion	~15 trillion

HINT: Ensure that you quote any interdependencies between the features. Any explanation that argues "bigger is better" will not be accepted.

QUESTION SET 5 (15 points)

You are asked to reduce computational complexity of solving the MDP by selecting the γ discount factor. Explain why this is a good idea and how you will select the γ value to effectively reduce the computation associated with estimating the $Q(s, a)$ value function.