HW4 : ECE/CS 434 : Real-World Algorithms for IoT and Data Science : Homework 4 : Due 11 :59pm, Wed, Apr 15, 2026

Problem 1 : State True/False with 1 line justifications [3x6 = 18 points]

a. Consider using the steepest gradient descent algorithm to find the minimum of function $F(x)$ starting from x_0. (T/F) If a learning rate α will end in the global minimum, a smaller learning rate $\frac{\alpha}{2}$ will also end in the global minimum.

b. A signal that is below the noise floor (i.e., amplitude of the signal is less than the noise amplitude on average) cannot be detected through cross-correlation.

c. In the early stage of UnLoc, landmarks closer to the origin (e.g. the entrance) are more reliable and of lower localization error compared to landmarks far away from the origin.

d. Echoes of the same signal source, when arriving at the receiver, are correlated.

e. Array consists of $N = 3$ antennas will always suffer from front-back ambiguity in AoA estimation.

f. Order of rotations does not matter. That is : $R_x(R_y(R_z)) \equiv R_y(R_z(R_x))$.

Problem 2 : Beamforming, Triangulation, ToF, TWR [4x11 = 44 points]

(a) Write one similarity and one difference between the steering matrix and the Fourier matrix.

(b) The received signal at a microphone array is written as $y = As + n$, where y is the received signal vector (of dimension $M \times 1$ where M is the number of microphones), A is the steering matrix, s are the signal sources, and n is the noise vector. Write out this equation by showing all the vectors and matrices (assume $K < M$ different sound sources, and assume only line-of-sight path for each of them). *For example, an equation showing all vectors and matrices in $A \times x = b$ would look like :*

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} k \\ l \end{bmatrix}$$

(c) For the same problem above, assume there is only a single source but $K < M$ echoes. Now write the same equation showing all the vectors and matrices.

(d) Explain the Delay-Sum algorithm using the equation in part (c) above.

(e) True/False : The accuracy of the Delay-Sum Algorithm is expected to improve with more microphones as receivers. Answer with a mathematical justification.

(f) You are performing WiFi triangulation and you have computed the 3 AoA spectrums at 3 WiFi access points (AP). The AoA spectrum is a vector of probabilities for each angle θ, and often a WiFi AP announces its AoA estimate as the angle with maximum probability. What is the correct way to estimate the location of the user from these AoA spectrums. Explain with equations.

(g) Complete the sentence : MUSIC does not work well when signals are correlated because

(h) Consider MUSIC applied on to a 3-microphone array receiving a voice signal over a single steering vector a_1 (and no reverberations). Draw the steering vector and the noise sub-space in a complex 3D space. Label your diagram, including the 3 axes of the 3D space as well as the Eigenvectors.

(i) For MVDR, assume we have four receivers with received signals y_1, y_2, y_3, y_4, and two incoming signals at angles θ_1, θ_2. Derive the optimization function that gives the optimal weight w if we want to amplify the signal arriving from θ_1 while canceling out the signal arriving from θ_2. (You do not need to solve the optimization problem, and you can also denote steering vectors for AoA=θ_i as a_i).

(j) Assume we have a pair of unsynchronized transmitter and receiver, but they can communication through both acoustic and wireless radio. Propose how ToF estimation can be performed through single-sided ranging.

(k) In the Symmetric Double-Sided TWR method, assume that $D_a = D_b$. Also, assume that both e_a, e_b are independent Gaussian random variables, distributed as $N(0, \sigma_1^2)$ and $N(0, \sigma_2^2)$. What is the mean and variance of the error that you can expect from this ranging method.

Problem 3 : Probability, HMM, and Applications [4x7 = 28 points]

(a) Suppose random variables A and B are conditionally independent, given C. This means $P(A, B|C) = P(A|C)P(B|C)$. One real life example of such conditional independence is the following : people's height and vocabulary are not independent, but they are conditionally independent given that people are from the same class. Can you give one more example.

(b) Three universities (A, B, C) have 100, 200, 300 students respectively, with 50, 75, and 100 of those students interested in mobile computing. Google randomly picks a university with equal probabilities (there is a 1/3 chance of selecting either A, B or C), and then draws a student's name at random from that university. The student proves to be interested in mobile computing. What is the probability the student comes from university C ?

(c) In a HMM, prove that $P(s_k|m_{1:n}) = P(s_k|m_{1:k})P(m_{k+1:n}|s_k)$.

(d) Complete the sentence :
When performing HMM for IMU based dead-reckoning, $P(m_k|s_{1:k}) \neq P(m_k|s_k)$ because ...

(e) Imagine running an HMM for handwriting recognition. The state variable s takes on values "A, B, C, ... Z" and we have measurements from people's handwriting, say m. Describe an example cost function, $P(m|s)$ in plain language, and then express the same as a mathematical equation.
Hint : *Treat each known state, say "A", as an image composed of black or white pixels.*

(f) Kalman filters rely on the fact that linear combinations of independent Gaussian random variables are also Gaussian. Explain with equations why this is the case.

(g) Show that the Kalman gain K completely uses the measurement when the measurement error covariance $R_m = 0$, and only uses the model when the process error covariance $R_p = 0$.

Problem 4 : Beamforming [4x5=20 points]

1. Derive the equation for the radiation pattern for N sensors with D separation, for a frequency F. Assume the wavelength is λ.
2. Visualize the radiation patterns for $N = 2, 4, 8, 16, 32$ for a fixed $D = \lambda/2$.
3. Visualize the radiation patterns when F increases while D remains unchanged.
4. Explain why $D = \lambda/2$ is necessary ; what happens when D grows larger.
5. What happens when D is made smaller ?