Fall 2016
Pattern Recognition Theory 18794

Homework 1

Due: Friday, Sep. 30, 2016, at the beginning of class, in hard copy.

Do all problems by hand, and do not use Matlab, unless stated otherwise.

To receive full credit, show the steps to your solution. Please be clear, neat, and eligible when you write your solutions. If the TAs cannot follow your work, you will not receive credit for that problem. Homework will be done individually: each student must hand in their own answers. It is acceptable for students to collaborate in figuring out answers and helping each other solve the problems. We will be assuming that, as participants in a graduate course, you will be taking the responsibility to make sure you personally understand the solution to any work arising from such collaboration. You must also indicate on each homework with whom you collaborated. **This homework has 110 points.**

You're encouraged to type up your solutions. Feel free to use Latex. Many free online latex environments exist such as ShareLatex or Overleaf. Check them out.

Problem 1 - because we all like linear algebra (22 pts) [Dipan]

1. A vector can be broken up into its direction and magnitude. The equation $\mathbf{Av} = \mathbf{w}$ can seen as applying a linear transformation to vector $\mathbf{v}$ to get a different vector $\mathbf{w}$. If $\mathbf{v}$ happens to be an eigenvector of $\mathbf{A}$, describe what happens to the direction of $\mathbf{v}$ after transformation?

2. Given

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 3 & 3 & 3 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} -1 & 2 \\ 2 & 2 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 2 & 8 & 1 & 0 \\ 3 & 12 & 2 & 0 \\ 2 & 8 & 3 & -1 \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} 2 & 8 & 1 & 1 \\ 3 & 12 & 2 & 1 \\ 2 & 8 & 3 & -1 \end{bmatrix}$$

(a) What is the rank of each matrix ?

(b) What is the trace of each matrix ?

(c) What is the determinant of matrices $\mathbf{A}, \mathbf{B}$?

(d) Find the eigenvalues and eigenvectors of $\mathbf{B}$.

(e) What is the dimension of the null space of each of the matrices ?

Problem 2 - because we all really like linear algebra (30 points) [Dipan]

Practice with Linear Algebra:

1. (2 pts) Prove that if $\mathbf{P}$ is an orthogonal matrix, then $\mathbf{P}\mathbf{P}^T$ is a projection matrix.

2. (7 pts) Under what condition(s), for arbitrary non-zero matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$, do we have $rank(\mathbf{A} + \mathbf{B}) = rank(\mathbf{A}) + rank(\mathbf{B})$? (Hint: Think singular values/vectors)

3. (3 pts) In general, does a relation exist between the rank of a sum of matrices and the sum of the ranks of the individual matrices ?

4. (3 pts) Under the premise of problem 1.2, you arrived at a condition. Is the condition sufficient to ensure that $\sigma_{max}(A+B) = \max(\sigma_{max}(A), \sigma_{max}(B))$? If not, what else do you need to constrain ? Here, $\sigma_{max}(A)$ is the maximum singular value of $\mathbb{A}$

5. (10 pts) Let $\mathbf{A} \in \mathbb{R}^{N \times N}$ be a full-rank matrix, with columns $\mathbf{a}_1, ..., \mathbf{a}_N$. Prove that for every $\beta \in \mathbb{R}^N$ with $\beta_i \geq 0$, $\sum_i^N \beta_i = 1$, we have $rank(\mathbf{A} - \mathbf{Q}) = N - 1$. Where $\mathbf{Q} = [\mathbf{A}\beta, \mathbf{A}\beta, ..., \mathbf{A}\beta]$ is formed by repeating the column vector $(\mathbf{A}\beta)$, N times, i.e. $\mathbf{Q} \in \mathbb{R}^{N \times N}$. (Hint: Try to scale the problem down and approach it from a geometric paradigm.)

Problem 3 - Lagrange's favorite multipliers (10 points) [Nancy]

Lagrange Multipliers – Overview (more information will be posted as supplementary)

Lagrange multipliers allows us to maximize or minimize functions with the constraint that we only consider points on a certain surface. To find critical points of a function $f(x, y, z)$ on a level surface $g(x, y, z) = C$ (or subject to the constraint $g(x, y, z) = C$), we must solve the following system of simultaneous equations:

$$\bigtriangledown f(x, y, z) = \lambda \bigtriangledown g(x, y, z)$$
$$g(x, y, z) = C$$

, where $\bigtriangledown f$ and $\bigtriangledown g$ are vectors, we can write this as a collection of four equations in the four unknowns x, y, z and λ

$$f_x(x, y, z) = \lambda g_x(x, y, z)$$
$$f_y(x, y, z) = \lambda g_y(x, y, z)$$
$$f_z(x, y, z) = \lambda g_z(x, y, z)$$
$$g(x, y, z) = C$$

, where λ is "Lagrange multiplier"

Lagrange Multipliers – Problem

1. Find the maximum and minimum values of $f(x, y) = 81x^2 + y^2$ subject to the constraint $4x^2 + y^2 = 9$
2. Find the maximum and minimum values of $f(x, y, z) = x + y + 2z$ on the surface $x^2 + y^2 + z^2 = 3$

Problem 6 - decisions decisions decisions (10 points) [Nancy]

The two random variables $\mathbf{X}$ and $\mathbf{Y}$ are jointly Gaussian. Let

$$\begin{aligned} \mathbf{X} &\sim \mathcal{N}(\mu_x, \sigma_x^2), \\ \mathbf{Y} &\sim \mathcal{N}(\mu_y, \sigma_y^2). \end{aligned}$$

The correlation coefficient ρ between $\mathbf{X}$ and $\mathbf{Y}$ is 0.5. Two new PDFs are generated as follows,

$$\begin{aligned} \mathbf{Z} &= \mathbf{X} + \mathbf{Y} \\ \mathbf{W} &= 3\mathbf{X} + 2\mathbf{Y}. \end{aligned}$$

Determine the decision boundaries of the two PDFs which give the minimum errors using $\mu_x = 0$, $\mu_y = 4$, $\sigma_x = 1$, $\sigma_y = 2$ for the cases below.

1. (8 pts) $P(\omega_1) = P(\omega_2)$
2. (8 pts) $P(\omega_1) = 3P(\omega_2)$

Problem 7 - no lab like MATLAB (10 pts) [Nancy]

MATLAB. For each distribution pair below: Generate 400 samples from each of the distributions. (Hint: Look at the *randn* documentation). Print a plot with your samples. Use “x” for points from one distribution, and “o” for points from the other distribution. Make the axis limits *axis([-6 10 -8 8])*. On your plot printouts, draw the general shape (it does not need to be exact) of the decision boundary between the distributions (assume equal priors). Do not compute the decision boundary mathematically. Attach the plots and your .m file. Describe how the decision boundary changes when the distributions do not have the same covariance.

1. $$\mathbf{x}_1 \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & 1\\1 & 8\end{bmatrix}\right),\ \mathbf{x}_2 \sim \mathcal{N}\left(\begin{bmatrix}4\\0\end{bmatrix}, \begin{bmatrix}1 & 1\\1 & 8\end{bmatrix}\right)$$
2. $$\mathbf{x}_1 \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & 1\\1 & 8\end{bmatrix}\right),\ \mathbf{x}_2 \sim \mathcal{N}\left(\begin{bmatrix}4\\0\end{bmatrix}, \begin{bmatrix}2 & 0\\0 & 2\end{bmatrix}\right)$$

Problem 8 - because there’s nothing like no error (20 points) [Dipan & Nancy]

A two-dimensional feature vector is use to select one of two classes shown in Figure 1. The joint PDFs are uniform over a square for class ω_1 and a triangle for class ω_2. For each case, determine the probability of assigning $\mathbf{x}$ to the wrong class ω (compute ε_1 and ε_2), and the total probability of error (compute P_e).

1. (8 pts) Case 1: $P(\omega_1) = P(\omega_2)$
2. (8 pts) Case 2: $P(\omega_1) = 2P(\omega_2)$

Bonus (8 points) - just for you !

Pattern recognition involves gathering data information with a sensor, extracting features, building a classifier, and then making a decision. For your project you will need to: (1) gather data or choose from an existing database, (2) select features such as raw pixels, SIFT, SURF, HOG, wavelets, Gabor, or even Deep

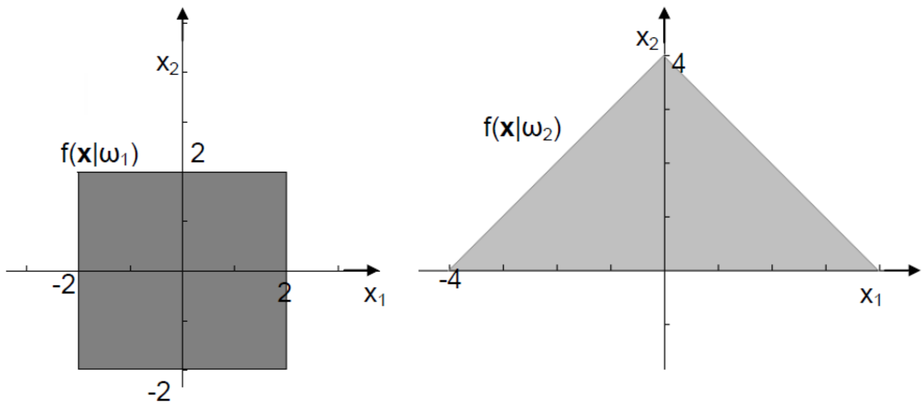

Figure 1: Corresponds to Problem 8

learning features and (3) design a classifier such as Bayesian, Deep Neural Network, SVM, correlation filter, LDA, Kernels, etc. Some of these topics will be cover later in the class but you will need to be read up on these topics to get ideas on how you want to implement them in your project. Feel free to talk to the professor or TAs if you're having trouble deciding on a topic.

1. (2 extra-credit pts) Write the name(s) of your project partner(s). Teams of 2 or 3 people.

2. (6 extra-credit pts) Describe in a few sentences, two ideas of possible projects that you want to pursue.