COMP/ENGN8535 Homework Assignment #4

General Instructions

- Due Date: Friday 26 May 2023, 23:59pm (AEST).
- Full marks for this homework = 100 points.
- Please complete as many problems as you can, and answer the questions clearly and concisely. In all your solutions, you need to show the key steps for how you reach your answers. The clarity of your answers may affect the final marks.
- Submit your answers in a single "pdf" file, with file name U4056789-HW-4.pdf (Note: replace the U4056789 with your own university ID) to Wattle before the deadline. It is a good idea to also include a cover page with your full name and university ID so that we can identify your submission easily when marking.
- In preparing your PDF submission, both typed and handwritten-and-scanned answers are acceptable, though the former is preferred due to its clarity.
- Late submissions will not be accepted and incur a Mark of 0.

Academic Integrity: Your homework must be completed individually and independently. You must comply with the University Policy on Academic Integrity. Plagiarism will lead to severe consequences, including academic misconduct. You are allowed to discuss the problems and solution ideas with other students, but your final solutions must be your own work and written in your own language.

Problems

1. (20 points) (Kurtosis of a sum) Let X and Y be two independent random variables with zero mean and unit variance, i.e.

$$\mathbb{E}[X] = 0 \quad \mathbb{E}\left[X^2\right] = 1$$
$$\mathbb{E}[Y] = 0 \quad \mathbb{E}\left[Y^2\right] = 1$$

Show that

$$\text{kurt}(X + Y) = \text{kurt}(X) + \text{kurt}(Y)$$

Hint: Recall that the kurtosis of a random variable X with zero mean is

$$\text{kurt}(X) = \mathbb{E}\left[X^4\right] - 3\left(\mathbb{E}\left[X^2\right]\right)^2, \tag{1}$$

and use the binomial formula and the linearity of expectation.

2. (20 points) (Negentropy and Mutual Information) Let Y_1 and Y_2 be two (potentially dependent) continuous random variables both with the same covariance (matrix) Σ. The mutual information between Y_1 and Y_2 is defined as

$$I(Y_1; Y_2) = \int\int p(y_1, y_2) \log \frac{p(y_1, y_2)}{p(y_1)p(y_2)} dy_1 dy_2.$$

The negentropy of Y_1 is

$$J(Y_1) = H(Y_g) - H(Y_1),$$

and the negentropy of Y_2 is

$$J(Y_2) = H(Y_g) - H(Y_2)$$

where

$$H(Y_i) = -\int p(y_i) \log p(y_i) dy_i$$

is the entropy of Y_i for $i = 1, 2$, and

$$H(Y_g) = \frac{1}{2} \log\left(\det(2\pi e \Sigma)\right)$$

is the entropy of a Gaussian random variable with the same covariance Σ.

(a) (10 points) Show that the mutual information satisfies

$$I(Y_1; Y_2) = H(Y_1) + H(Y_2) - H(Y_1, Y_2)$$

where $H(Y_1, Y_2)$ is the joint entropy of Y_1 and Y_2 given by

$$H(Y_1, Y_2) = -\int\int p(y_1, y_2) \log p(y_1, y_2) dy_1 dy_2.$$

(Note that $H(Y_1, Y_2)$ is equivalently the entropy $H(Y)$ of the collection random variable $Y \triangleq \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$.)

(b) (10 points) Show then that

$$I(Y_1; Y_2) = C - \sum_{i=1}^{2} J(Y_i).$$

where C is a number that depends only on $H(Y_1, Y_2)$ and the entropy $H(Y_g)$.

Hint: Use properties of the logarithm and the product/sum rules of probability.

3. (20 points) (Under-constrained equations) Solve the following minimization problems by hands. Please list your key steps. Please also verify that your solutions are correct by drawing the graphs for the problems.
   Let $\Phi = \begin{bmatrix} 4 & 1 \end{bmatrix}$ and $y = 1$.
   (a) (10 points) Solve $\min\limits_{x \in \mathbb{R}^2} \|x\|_2 \quad \text{subject to} \quad \Phi x = y.$
   (b) (10 points) Solve $\min\limits_{x \in \mathbb{R}^2} \|x\|_1 \quad \text{subject to} \quad \Phi x = y.$

4. (20 points) (Sparsity) Assume we have:

$$\Phi = \begin{bmatrix} 1 & 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1 & 1/\sqrt{2} \end{bmatrix}$$

   (a) (10 points) Suppose that for some fixed $\mathbf{b} \in \mathbb{R}^2$, the equation $\Phi\mathbf{x} = \mathbf{b}$ has a 1-sparse solution. Show this solution is unique, and so we can recover any 1-sparse solution $\mathbf{x} \in \mathbb{R}^4$.
   (b) (10 points) Let $\mathbf{b}$ a vector in $\mathbb{R}^2$. Show that the equation $\Phi\mathbf{x} = \mathbf{b}$ can have as many as six distinct 2-sparse solutions.

5. (20 points) (Sparsity) Consider the underdetermined linear equation $\Phi\mathbf{x} = \mathbf{b}$, where Φ is the matrix in Question 4, $\mathbf{x} \in \mathbb{R}^4$, and $\mathbf{b} = [0, 3]^t$ (here the superscript t denotes the transpose).
   (a) (5 points) Verify that the vector $\mathbf{x} = [0, 0, 3, 0]^t$ is a 1-sparse solution.
   (b) (5 points) Find the minimum norm solution to $\Phi\mathbf{x} = \mathbf{b}$ using the ℓ^2 norm. (Suggestion: Solve $\Phi\mathbf{x} = \mathbf{b}$ for x_1 and x_3 in terms of x_2 and x_4, then express $\|\mathbf{x}\|_2^2$ in terms of just x_2 and x_4 and minimize in these two variables.) What's the sparsity of this solution?
   (c) (10 points) Find the minimum norm solution to $\Phi\mathbf{x} = \mathbf{b}$ using the ℓ^1 norm $\|\mathbf{x}\|_1$ and the same approach as part (b). Although $\|\mathrm{x}\|_1$ isn't differentiable, it's easy to find the minimum graphically after you've expressed $\|\mathbf{x}\|_1$ as a function of two variables, by plotting $\|\mathbf{x}\|_1$ as a function of x_2 and x_4.

END.