ORIE 4740 Spring 2022

Homework 0

Due Friday Feb 4th by 11:59pm **on Gradescope** (`https://www.gradescope.com`). Regrade requests should be sent via Gradescope.

Name:

NetID:

1. Normal distributions.
   (a) Suppose $X_1 \sim N(1,4)$, $X_2 \sim N(3,7)$ and X_1, X_2 are independent. What is the distribution of $X_1 + X_2$?
   (b) Suppose that $X \sim N_k(\mu, \Sigma)$ and that A is a $\ell \times k$ matrix with full rank. What is the distribution of the vector AX? (Here $N_k(\mu, \Sigma)$ is the k-dimensional Normal distribution with mean vector $\mu \in \mathbb{R}^k$ and covariance matrix $\Sigma \sim \mathbb{R}^{k\times k}$.)
   (c) Suppose that $y = X\beta + \epsilon$, where $\epsilon \sim N_n(\mathbf{0}, I)$ ($\mathbf{0}$ is the all-zero vector and I denotes the $n \times n$ identity matrix), $\beta \in \mathbb{R}^k$ is a *fixed* vector, $X \in \mathbb{R}^{n\times k}$ is a matrix with i.i.d. entries distributed as $N(0,1)$, and ϵ and X are independent of each other. What is the distribution of y?

2. Mean and variance.
   (a) Express $\mathrm{Var}(X_1 + 2X_2)$ using the variances and covariance of X_1, X_2 (X_1 and X_2 are *not* necessarily independent).
   (b) Suppose that $X_1, \ldots X_n$ are i.i.d. real-valued random variables with finite variances. Show that
   $$\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\mathrm{Var}(X_1).$$
   (c) Assume that X, Y are independent random variables with $\mathbb{E}[X] = 0$, $\mathbb{E}[Y] = 1$, $\mathrm{Var}(X) = 1$, $\mathrm{Var}(Y) = 3$. Compute $\mathbb{E}\left[(3X + Y)(6Y + 2X - 1)\right]$.

3. Suppose that a is a fixed number and X is a random variable.
   (a) Prove that
   $$\mathbb{E}\left[(X-a)^2\right] = \left(\mathbb{E}[X] - a\right)^2 + \mathrm{Var}(X).$$
   (b) Find the (deterministic) value of a that minimizes the quantity $\mathbb{E}\left[(X-a)^2\right]$.

4. Suppose that $Y, X_1, \ldots, X_k$ are random variables, where Y takes values $1, 2, 3, \ldots, 10$. If $X_1, \ldots X_k$ are independent *conditioned on* Y, express the conditional probability $\mathbb{P}\left(Y = 2 | X_1 = x_1, \ldots, X_k = x_k\right)$ in terms of $\mathbb{P}(Y = y)$ and $\mathbb{P}(X_k = x_k | Y = y)$, $y = 1, 2 \ldots, 10$.

5. For arbitrary numbers $x_1, \ldots, x_n, t_1, \ldots, t_n \in \mathbb{R}$, find the values of a and b that minimize the expression $\sum_{i=1}^{n}(x_i - a - bt_i)^2$.

6. For each number $1 \le q < \infty$, the so-called ℓ_q norm of a vector $u \in \mathbb{R}^k$ is defined as $\|u\|_q \triangleq \left(\sum_{i=1}^k |u_i|^q\right)^{1/q}$. Prove that $\|u\|_2 \le \|u\|_1 \le \sqrt{k}\,\|u\|_2$.

7. True or false (capital letters denote matrices, which are assumed to be invertible; lower case letters denote column vectors):

   (a) $AA^{-1} = A^{-1}A = I$.
   (b) $(AB)^{-1} = A^{-1}B^{-1}$.
   (c) $(AB)^\top = A^\top B^\top$.
   (d) $(A^{-1})^\top = (A^\top)^{-1}$.
   (e) $(A+B)^{-1} = A^{-1} + B^{-1}$.
   (f) $\|u\|^2 = u^\top u = \text{trace}(uu^\top)$.

8. Suppose that $A \in \mathbb{R}^{p\times n}$ is a p-by-n matrix. Prove that the matrix $AA^\top$ is positive semidefinite.

9. Which of the following matrices has/have full rank? (An n-by-m matrix Z is called full rank if $\text{rank}(Z) = \min(n, m)$.) $A = \begin{pmatrix} 2 & 0 & 5 \\ 1 & 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 0 \\ 5 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}$.

10. Suppose that $A \in \mathbb{R}^{p\times p}$ is a real symmetric and positive semidefinite matrix. Find a matrix $Z \in \mathbb{R}^{p\times p}$ such that $Z^2 = A$. (Hint: Use the eigendecomposition of A.)

11. This is not a question, but it is very important that you are comfortable with the concepts and notations used in the following statement, and understand why the statement is true.

    Suppose that X is a given $n \times p$ matrix of the form

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix}$$

    and y is a given n-dimensional column vector of the form

$$y = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}^\top = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

    If $p > n$, then:

    - The matrix $X^\top X$ is singular (that is, not invertible).
    - In general there are infinitely many vectors $\beta = \begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_p \end{bmatrix}^\top$ that satisfy the linear equation system

$$X\beta = y.$$

      To be rigorous, this is true when the matrix X has linearly independent rows, or equivalently, when X has full row rank.

    These facts are useful when we discuss overfitting and model flexibility in linear regression.