Homework 11 Solutions

Math 198: Math for Machine Learning

Due Date:
Name:
Student ID:

Instructions for Submission

Please include your name and student ID at the top of your homework submission. You may submit handwritten solutions or typed ones (LaTeX preferred). If you at any point write code to help you solve a problem, please include your code at the end of the homework assignment, and mark which code goes with which problem. Homework is due by start of lecture on the due date; it may be submitted in-person at lecture or by emailing a PDF to both facilitators.

1 More Probability Proofs

1. Show that for any random variables X and Y, $\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$.

$$\begin{aligned}\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] &= \mathbb{E}[XY - X\mathbb{E}[Y] - \mathbb{E}[Y]X + \mathbb{E}[X]\mathbb{E}[Y]] \\ &= \mathbb{E}[XY] - 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]\end{aligned}$$

2. Show that for any random variables X, Y, Z and constants α, β,

$$\text{Cov}(\alpha X + \beta Y, Z) = \alpha \text{Cov}(X, Z) + \beta \text{Cov}(Y, Z)$$

$$\begin{aligned}\text{Cov}(\alpha X + \beta Y, Z) &= \mathbb{E}[\alpha XZ + \beta YZ] - \mathbb{E}[\alpha X + \beta Y]\mathbb{E}[Z] \\ &= \mathbb{E}[\alpha XZ] + \mathbb{E}[\beta YZ] - \mathbb{E}[\alpha X]\mathbb{E}[Z] - \mathbb{E}[\beta Y]\mathbb{E}[Z] \\ &= \alpha\mathbb{E}[XZ] - \alpha\mathbb{E}[X]\mathbb{E}[Z] + \beta\mathbb{E}[YZ] - \beta\mathbb{E}[Y]\mathbb{E}[Z] \\ &= \alpha\text{Cov}(X, Z) + \beta\text{Cov}(Y, Z)\end{aligned}$$

3. Show that for independent random variables X and Y, $\text{Cov}(X, Y) = 0$.
For discrete X, Y

$$\begin{aligned}\text{Cov}(X, Y) &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\ &= \sum_{x \in \Omega_x} \sum_{y \in \Omega_y} xyp(xy) - \mathbb{E}[X]\mathbb{E}[Y] \\ &= \sum_{x \in \Omega_x} \sum_{y \in \Omega_y} xp(x)yp(y) - \mathbb{E}[X]\mathbb{E}[Y] \\ &= \mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[X]\mathbb{E}[Y] \\ &= 0\end{aligned}$$

The continuous case has an analogous construction.

4. Show that for uncorrelated random variables $X_1, \ldots, X_n$, $\mathrm{Var}(X_1 + \ldots + X_n) = \sum_{i=1}^n \mathrm{Var}(X_i)$.

$$\begin{aligned}
\mathrm{Var}(\sum_{i=1}^n X_i) &= \mathbb{E}[(\sum_{i=1}^n X_i)^2] - \mathbb{E}[\sum_{i=1}^n X_i]^2 \\
&= \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}[X_i X_j] - \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}[X_i]\mathbb{E}[X_j] \\
&= \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j] \\
&= \sum_{i=1}^n \sum_{j=1}^n \mathrm{Cov}(X_i, X_j)
\end{aligned}$$

Since the X_i are uncorrelated with one another, all the $\mathrm{Cov}(X_i, X_j)$ terms will be 0 when $i \neq j$, so this further simplifies to $\sum_{i=1}^n \mathrm{Var}(X_i)$.

5. Show that for any random vector $\mathbf{X}$, its covariance matrix $\mathbf{\Sigma}$ is positive semi-definite.
For any $\mathbf{x}$,

$$\begin{aligned}
\mathbf{x}^\top \mathbf{\Sigma} \mathbf{x} &= \mathbf{x}^\top \mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X})^\top]\mathbf{x} \\
&= \mathbb{E}[\mathbf{x}^\top (\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^\top \mathbf{x}] \\
&= \mathbb{E}[((\mathbf{X} - \mathbb{E}[\mathbf{X}])^\top \mathbf{x})^{\mathbf{2}}] \geq 0
\end{aligned}$$