CS57800 Statistical Machine Learning
Homework 1 Solutions

1 Foundations

1. Consider the planes $x_1+x_2+3x_3 = 4$ and $x_1+2x_2+4x_3 = 5$ in $\mathbb{R}^3$. Find parametric equations for the line of intersection of these two planes.

   The planes have the normal vectors $\mathbf{a} = (1, 1, 3)$ and $\mathbf{b} = (1, 2, 4)$, respectively. Let L be the line of the intersection. Then, $\mathbf{v} = \mathbf{a} \times \mathbf{b} = (-2, -1, 1)$ is parallel to L. The next step is to find a point in the line of intersection. To do that, let any variable be 0. If $x_3 = 0$,

   $$x_1 + x_2 = 4$$
   $$x_1 + 2x_2 = 5$$

   Thus, we get $x_1 = 3, x_2 = 1$, and so the equations of the lines are

   $$x_1 = 3 - 2t$$
   $$x_2 = 1 - t$$
   $$x_3 = t$$

2. Given three points $P(0, 0, 0)$, $Q(1, -1, 1)$, $R(4, 3, 7)$, find a vector which is orthogonal to the plane through P, Q, and R.

   Let a and b be the vectors defined below. Then the orthogonal vector is the cross product $a \times b$.

   $$\begin{aligned}
   a &= PQ = (1, -1, 1) - (0, 0, 0) = (1, -1, 1) \\
   b &= PR = (4, 3, 7) - (0, 0, 0) = (4, 3, 7) \\
   a \times b &= \begin{vmatrix} i & j & k \\ 1 & -1 & 1 \\ 4 & 3 & 7 \end{vmatrix} \\
   &= \begin{vmatrix} -1 & 1 \\ 3 & 7 \end{vmatrix} i - \begin{vmatrix} 1 & 4 \\ 1 & 7 \end{vmatrix} j + \begin{vmatrix} 1 & -1 \\ 4 & 3 \end{vmatrix} k \\
   &= -10i - 3j + 7k = (-10, -3, 7)
   \end{aligned}$$

   Therefore, the vector orthogonal to the plane is $(-10, -3, 7)$.

3. Differentiate the following equations.

   (a) $f(x) = (3x^2)(x^{\frac{1}{2}})$
   $f'(x) = 3 \cdot \frac{5}{2}x^{\frac{3}{2}}$

   (b) $f(x) = (e^{2x} + e)^{\frac{1}{2}}$
   $f'(x) = \frac{1}{2}(e^{2x} + e)^{\frac{-1}{2}}(e^{2x} \cdot 2 + 0)$

   (c) $f(x) = [\ln(5x^2 + 9)]^3$
   $f'(x) = \frac{1}{(5x^2+9)^3} \cdot 30x[\ln(5x^2 + 9)]^2$

4. Find $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$.

   (a) $f(x, y) = xy^3 + x^2y^2$
   $\frac{\partial f}{\partial x} = y^3 + 2xy^2$
   $\frac{\partial f}{\partial y} = 3xy^2 + 2x^2y$

   (b) $f(x, y) = xe^{2x+3y}$
   $\frac{\partial f}{\partial x} = 2xe^{2x+3y} + e^{2x+3y}$
   $\frac{\partial f}{\partial x} = 3xe^{2x+3y}$

5. We say that $f(n) \prec g(n)$ if $g(n)$ grows faster than $f(n)$. Order the following functions by $\prec$ from the the lowest to the highest:
   $(\frac{5}{3})^{2n}$, 10^8, $\sqrt{n^3}\log^2 n$, $2^{\log_2 n}$, $\log^4 \sqrt{n}$, $2^{3\log_2 n}$, 2^n

   The ascending order is
   $10^8 \prec \log^4 \sqrt{n} \prec 2^{\log_2 n} \prec \sqrt{n^3}\log^2 n \prec 2^{3\log_2 n} \prec 2^n \prec (\frac{5}{3})^{2n}$

6. Suppose you roll three dice. Compute the followings: (a) the expected value of the sum of the rolls, (b) the expected value of the product of the rolls, and (c) the variance of the sum of the rolls.

   (a) Let X_1, X_2, X_3 be the rolls.
   We have $E(X_i) = \frac{7}{2}$ for $i = 1, 2, 3$.
   Then $E(X_1 + X_2 + X_3) = E(X_1) + E(X_2) + E(X_3) = 3 \times \frac{7}{2} = \frac{21}{2}$

   (b) Since the rolls are independent, we have
   $E(X_1X_2X_3) = E(X_1)E(X_2)E(X_3) = (\frac{7}{2})^3 = \frac{343}{8}$

   (c) Since the rolls are independent, we have
   $Var(X_1 + X_2 + X_3) = Var(X_1) + Var(X_2) + Var(X_3) = 3Var(X_1)$.
   To compute the variance of X_1, we compute
   $Var(X_1) = E(X_1^2) - E(X_1)^2 = \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) - (\frac{7}{2})^2 = \frac{35}{12}$
   Hence, the variance of the sum is $\frac{35}{4}$.