STOR 565 Homework

Show all work. Note: all logarithms are natural logarithms.

1. Let $\langle x, y\rangle = x^t y = \sum_{i=1}^d x_i\, y_i$ be the usual inner product in $\mathbb{R}^d$. Recall that the norm of a vector $x \in \mathbb{R}^d$ is defined by $||x|| = \langle x, x\rangle^{1/2}$

   a. Show that $||x|| = 0$ if and only if $x = 0$.

   b. Use the definition of the norm to show that $||x + y||^2 = ||x||^2 + 2\langle x, y\rangle + ||y||^2$.

   c. Use this equation and the Cauchy Schwarz inequality to establish the triangle inequality for the vector norm, namely $||x + y|| \leq ||x|| + ||y||$.

   d. The standard Euclidean distance between two vectors $x, y \in \mathbb{R}^d$ is defined by $d(x, y) = ||x - y||$. Use part (c) to establish that $d(x, y) \leq d(x, z) + d(z, y)$ for any vectors $x, y, z \in \mathbb{R}^d$. Draw a picture illustrating this result.

2. Let $x = (x_1, \ldots, x_d)^t$ be a vector in $\mathbb{R}^d$.

   a. Show that $||x|| \leq |x_1| + \cdots + |x_d|$. Hint: use the fact that for $a, b \geq 0$ one has $a \leq b$ if and only if $a^2 \leq b^2$. Give an example where the bound holds with equality, and an example where one has strict inequality.

   b. Use Cauchy-Schwarz to get the upper bound $|x_1| + \cdots + |x_d| \leq ||x||\, d^{1/2}$. Find an example where the bound holds with equality.

3. Let $a_1, \ldots, a_n$ be positive numbers. Use the Cauchy-Schwartz inequality for inner products to show that $(\sum_{k=1}^n a_k)(\sum_{k=1}^n a_k^{-1}) \geq n^2$. Hint: Begin with the identity $1 = a_k\, a_k^{-1}$.

4. (Inequalities from Calculus) Use calculus to establish the following inequalities.

   a. $(1 + u/3)^3 \geq 1 + u$ for every $u \geq 0$

   b. $x + x^{-1} \geq 2$ for $x \geq 1$

   c. $\log(1 + x) \geq x - x^2/2$ for $x \geq 0$. Note that this inequality requires taking a second derivative to show that the first derivative is increasing.

5. Let X, X' be independent random variables with the same distribution. In this case we say that X' is an independent copy of X. Show that $\text{Var}(X) = \frac{1}{2}\mathbb{E}(X - X')^2$

6. Let $x = x_1, \ldots, x_n$ be a univariate sample of n numbers. It is a standard, and important, fact that the quantity $h(a) = \sum (x_i - a)^2$ is minimized when (and only when) a is the sample mean $m(x) = n^{-1} \sum_{i=1}^n x_i$. Here we show this in two different ways.

a. Take a derivative to find the number a that minimizes or maximizes the function h, and then take another derivative to show that the number you found minimizes the function.

b. Add and subtract $m(x)$ inside the parentheses, expand the square, take the sum, and examine the terms you find.

7. Let $x = x_1, \ldots, x_n$ be a univariate sample, and let $\tilde{x} = \tilde{x}_1, \ldots, \tilde{x}_n$ be the standardized version of x with $\tilde{x}_i = (x_i - m(x))/s(x)$. Show that $m(\tilde{x}) = 0$ and $s(\tilde{x}) = 1$.

8. Let $r(x, y)$ be the sample correlation of a bivariate data set $(x, y) = (x_1, y_1), \ldots, (x_n, y_n)$.

a. Let $ax + b$ denote the data set $ax_1 + b, \ldots, ax_n + b$ and define $cy + d$ similarly. Show that $r(ax + b, cy + d) = r(x, y)$ if $a, c > 0$.

b. Use the Cauchy-Schwarz inequality to show that $r(x, y)$ is always between -1 and $+1$.