Problem 7.27 from the textbook.

Problem 7.31 from the textbook.

Problem 7.39 from the textbook.

As you might recall from past statistics courses, the multivariate linear regression problem is defined as follows. Your goal is to find the best linear “fit” to a number of data points. Each of m data points consists of the values $x_1, \ldots, x_n$ of n variables $X_1, \ldots, X_n$, as well as the function value $f(x_1, \ldots, x_n)$ at that point. Your goal is to approximate $f$ with a linear function of $X_1, \ldots, X_n$, as “closely” as possible. That is, you want to find coefficients $a_0, a_1, \ldots, a_n$ such that $a_0 + \sum_{i=1}^{n} a_i x_i$ is “close” to $f(X_1, \ldots, X_n)$ for all your data points. For simplicity of notation, we will denote the $j$th data point by the vector $\vec{x}(j) := (x_1^{(j)}, \ldots, x_n^{(j)})$.

“Close” is usually defined in terms of the sum of squared distances, i.e., the goal in “standard” regression is to minimize $\sum_{j=1}^{m} (f(\vec{x}(j)) - a_0 - \sum_{i=1}^{n} a_i x_i^{(j)})^2$. It is not too difficult to show that the best coefficients $a_i$ can be found by solving a certain system of linear equations.

In this problem, our goal will instead be to minimize the sum of absolute differences

$$\sum_{j=1}^{m} |f(\vec{x}(j)) - a_0 - \sum_{i=1}^{n} a_i x_i^{(j)}|.$$  \hspace{1cm} (1)

That is, given the data points $\vec{x}(j)$ and function values $f(\vec{x}(j))$, your goal is to find $a_0, a_1, \ldots, a_n$ to minimize (1). Show that this problem can be solved in polynomial time using linear programming. State what your variables are, formulate a linear program, and justify its correctness.

Suppose that you have two companies (think Microsoft, Apple), each trying to sell their own copy of n different products (think Word Processor, Operating System, Spread Sheet, ...). For each product $i = 1, \ldots, n$, we know the following:

(a) the price $p_i \geq 0$ that company 1 charges and the price $p'_i \geq 0$ that company 2 charges.
(b) the quality $q_i \geq 0$ of company 1’s version, and the quality $q'_i \geq 0$ of company 2’s version.

For instance, Apple may make a better operating system, but Microsoft a better Word Processor. A user would like to assemble a system, consisting of exactly one copy each of the n products, e.g., the user wants to buy one Word Processor, one Operating System, etc. The goal is to maximize total quality minus total price.

In principle, the user can combine the products of the two companies. But there’s a catch: the products of different companies may not be perfectly compatible. For each pair $(i, j)$, we have a penalty term $\pi_{ij} \geq 0$, the loss in utility from combining company 1’s version of product i with company 2’s version of product j. Notice that it is possible that $\pi_{ij} \neq \pi_{ji}$: just because Apple’s Word Processor doesn’t run well on Microsoft Windows doesn’t mean that MS-Word doesn’t run well on Mac-OS. We assume that
products are always compatible internally, i.e., there’s no penalty for combining two products from the same company. All pairwise penalty terms are subtracted from the a priori utility of the system.

Give a polynomial-time algorithm for computing which components to purchase from which of the two companies to maximize the total system utility (including penalties) minus the total price.

(6) [0 points]

**Chocolate Problem (2 chocolate bars):** In the previous problem, you were solving the user’s problem of finding the right system to purchase. Now, let’s turn the tables and maximize the profit for company 1. Company 1 knows (and cannot change) all the qualities of the products (both its own and its competitor’s), as well as the competitor’s prices and all incompatibilities. The only thing that the company can change are the prices of its own products. Once company 1 sets its prices, we assume that the user chooses a subset to optimize exactly the objective function from the previous problem.

Company 1 then gets paid exactly the prices for the products it sells to the user. Thus, higher prices means more profits from selling that particular product, but also, it might mean that the user buys the competitor’s version instead, meaning no profit at all. In fact, not only for that particular product, but there might be chain effects of other products being bought from the competitor as well. So the right choice of prices is important.

Give a polynomial-time algorithm for finding optimal prices to charge to maximize overall revenue. (Hint: You might want to prove first that the set of products that company 1 sells at the optimum prices is the same as the set of products it would sell if it priced all products at price 0.)

(7) [0 points]

**Research Question:** If you look carefully in the chocolate problem, you will notice that we assume that there is only one user (or, equivalently, that all users have the same quality values for all products). Here is an interesting questions, which — as far as I know — is open. Suppose that there are just two users. For both users, the incompatibility penalties $\pi_{ij}$ are the same, and of course, they have to be charged the same prices. But the two users may have completely different $q_i, q'_i$ values, because they value different features in their products. Your goal is to choose prices $p_i$ to maximize the joint revenue from products sold to both users.

If you make any progress on this — either by finding a polynomial-time algorithm or proving some hardness result — I’d be interested to hear about it.