COT 5615 Math for Intelligent Systems, Fall 2021

Homework 3

5.2 (15 pts) The supervisor is wrong, and the intern is most likely correct. She is examining 400 250-vectors, each representing the daily returns of a particular stock over one year. By the independence-dimension inequality, any set of $n+1$ or more n vectors is linearly dependent, so we know that this set of vectors is linearly dependent. That is, there exist some vectors that can be expressed as a linear combination of the others. It is quite likely that the returns of any particular stock, such as GOOG, can be expressed as a linear combination of the returns other stocks.

Even if Google's return can be expressed as a linear combination of the others, by the independence-dimension inequality, this fact is not as useful as it might seems. For example, Google's future returns would not be given by the same linear combination of other asset returns. So although the intern is right, and the supervisor is wrong, the observation cannot be monetized.

5.5 (15 pts) Two vectors are orthogonal if their inner product is zero. So we want to find γ such that $(a-\gamma b)^T b = a^T b - \gamma b^T b = 0$. If $b = 0$, then we can choose any γ, and we have $(a-\gamma b) \perp b$, since all vectors are orthogonal to 0. If $b \neq 0$, then $b^T b = \|b\|^2 \neq 0$, and we can take $\gamma = a^T b / b^T b$.

5.9 (15 pts) First we can find (or really, estimate) the speed of the computer, which we will call S flop/s. Gram-Schmidt requires $2nk^2$ flops. From $2 \approx S/\left(2kn^2\right)$, with $k = 10^3$ and $n = 10^4$, we find that $S \approx 10^{10}$ flop / s, which is 10 Gflop/s. Using this we can guess how long it will take to carry out Gram-Schmidt on the second problem, as $\left(2\tilde{n}\tilde{k}^2\right)/10^{10} = 0.05$ seconds.

Here's another way to get to the same answer. The time to carry out Gram-Schmidt scales linearly with n and quadratically by k. So if we reduce n by a factor of 10, we save a factor of 10. When we reduce k by a factor of two, we save an additional factor of 4. So with the new values of n and k(i.e., $\tilde{n}$ and $\tilde{k}$), we should be able to carry out Gram-Schmidt $40\times$ faster, and $2/40 = 0.05$ seconds, as above.

6.17 (15 pts)

(a) S *always* has linearly independent columns. To see this, suppose that $Sx = 0$. Since $Sx = (Ax, x)$, this means that $x = 0$.

(b) S *never* has linearly independent rows. S has $m+n$ rows, and each row has dimension n, so by the independence-dimension inequality, the rows are dependent.

A6.8 (5 pts) Not sure about A.

This is a trick problem and needs to be differentiated from 6.17.

6.18 (15 pts) The columns of an $m \times n$ Vandermonde matrix V are linearly dependent if there exists an n-vector c, with not all entries 0, such that

$$Vc = \sum_{i=1}^{n} c_i v_i = 0$$

The entries of the vector Vc are the values of the polynomial $p(t) = c_1 + c_2 t + \cdots + c_n t^{n-1}$ at $t_1, \cdots, t_m$:

$$Vc = \begin{bmatrix} c_1 & c_2 t_1 & c_3 t_1^2 & \cdots & c_n t_1^{n-1} \\ c_1 & c_2 t_2 & c_3 t_2^2 & \cdots & c_n t_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 t_m & c_3 t_m^2 & \cdots & c_n t_m^{n-1} \end{bmatrix} = \begin{bmatrix} p(t_1) \\ p(t_2) \\ \vdots \\ p(t_n) \end{bmatrix}$$

If $Vc = 0$, then $p(t_i) = 0$ for $i = 1, \cdots, m$, so $p(t)$ has at least m distinct roots $t_1, \cdots, t_m$. Using the fact from algebra mentioned in the hint this is only possible if the coefficients of p are zero, i.e., $c = 0$. We have shown that $Vc = 0$ implies $c = 0$, which means that the columns of V are linearly independent.

A6.2 (20 pts) All methods that can generate Vandermonde matrix are accepted.