Engr108: Introduction to Matrix Methods J. Costacurta
July 2024

Midterm Examination

This is a 90 minute in-class examination.

You may use one piece of paper (double-sided), but otherwise this examination is closed book. You may ask us questions if you find a problem confusing.

Please respect the honor code.

All problems have equal weight. Some are (quite) straightforward. Others, not so much.

Some problems involve applications. But you do not need to know *anything* about the problem area to solve the problem; the problem statement contains everything you need.

The problems do *not* appear in order of increasing difficulty.

Name: ______________________________

Stanford ID #: ______________________________

M.1 A *binary vector* is a vector with entries only 0 or 1 (for example, the 4-vector $(1, 0, 1, 0)$ is a binary vector). Such vectors are commonly used in *binary codes* such as the ASCII code, which is used to represent text on computers.

(a) Describe in words what is calculated by the inner product of two binary vectors.

(b) One distance metric commonly used with these vectors is called the *Hamming distance*. The Hamming distance between two n-vectors is defined as the number of entries (or positions) in which the two vectors differ. As an example, the Hamming distance between $x = (1, 0, 1)$ and $y = (1, 1, 0)$ is 2.

Show that the Hamming distance between any two binary vectors x and y is equivalent to $\|x - y\|^2$.

M.2 Suppose a and b are n-vectors with $\mathbf{avg}(a) = \mathbf{avg}(b) = 0$ and $\mathbf{rms}(a)$=5. For (a)-(c), either write the following quantities, or state that they can't be computed with the information given. For (d), your answer must be *True* or *False.* Provide a brief justification for your answers.

(a) $\mathbf{1}^T b$

(b) $\left\| \begin{bmatrix} a \\ b \end{bmatrix} \right\|$

(c) $\mathbf{std}(a)$

(d) If ρ, the correlation coefficient between a and b, is negative, the angle between them *must* be obtuse.

M.3 Recall the k-means clustering objective

$$J^{\text{clust}} = \frac{1}{N}\sum_{i=1}^{N} \|x_i - z_{c_i}.\|^2$$

In this problem, assume we are working with *distinct* datapoints $x_1, .., x_N$.

(a) Suppose we have a set of cluster centroids $z_1, ..., z_k$, where $k < N$. Is it always possible to add a new centroid z_{k+1} which will decrease J^{clust}? Justify your answer.

(b) What are the conditions under which a k-means clustering model has $J^{\text{clust}} = 0$? Would you consider such a model to provide the "best" possible clustering?

M.4 Recall the definition and first two derivatives of a third-order polynomial:

$$\begin{aligned} f(x) &= a + bx + cx^2 + dx^3 \\ f'(x) &= b + 2cx + 3dx^2 \\ f''(x) &= 2c + 6dx \end{aligned}$$

(a) Write the matrix that computes the first three derivatives of the above polynomial for a given x, i.e. give A such that

$$\begin{bmatrix} f(x) \\ f'(x) \\ f''(x) \end{bmatrix} = A \begin{bmatrix} 1 \\ x \\ x^2 \\ x^3 \end{bmatrix}$$

(b) Is it possible for the *columns* of A to be linearly independent? If yes, give an example of conditions on a, b, c, d that make the columns independent. If no, provide a brief justification.

(c) Is it possible for the *rows* of A to be linearly independent? If yes, give an example of conditions on a, b, c, d that make the rows independent. If no, provide a brief justification.

M.5 Suppose you run the Gram-Schmidt process on vectors v_1, v_2, v_3 (in order) which results in the nonzero orthonormal vectors q_1, q_2, q_3.

(a) Are v_1, v_2, v_3 linearly dependent or independent? Provide a brief explanation.

For each of the following changes to the original set of vectors, state whether the set of resulting orthonormal vectors will also change. If they won't change, briefly explain why not. If they will change, concisely describe how the new set of orthonormal vectors differs from q_1, q_2, q_3. Assume that Gram-Schmidt is run on the vectors in the order they're written.

(b) $v_1, 2v_2, v_3$

(c) $v_1, v_2, -2v_3$

(d) $v_1, v_2, v_2 + v_3$