Selected exercises.

David Khachatrian
davidkhachatrian.com

August 12, 2019 - ?

1 Introduction.

These exercise solutions refer to questions from the MIT OpenCourseWare course titled "Topics in Mathematics with Applications in Finance". I chose the questions that I believed would be most beneficial to work through/write up. Sections are labeled by the title of the Problem Set. This writeup and any mistakes/typos in it are mine – feel free to contact me if you notice something to correct!

$\mathcal{R}$

2 Linear Algebra.

Problem A-1

a. True; the row-rank and column-rank of a matrix are always equal.

b. True; we're transforming the basis vectors of B by the transformation A. All the basis vectors of AB will be linear combinations of the columns of B, and so are bounded by $rank(B)$.

c. True; for any rectangular matrix $C(m \times n), m \neq n$, assume its left-inverse L and right-inverse R exist. By necessity, L must be $n \times m$ (so that $LC = I_n$, the identity matrix of n dimensions) and R must be $m \times n$ (so that $CR = I_m$, the identity matrix of m dimensions). So just considering dimensions, $L \neq R$.

d. False; rectangular matrices can be full-rank but are never invertible.

e. False; for example, $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ is not invertible.

f. False; an eigenvalue of a matrix can be 0, in which case the matrix is not invertible.

g. False. By definition, A is diagonalizable if there exists a P such that $P^{-1}AP = D$ is diagonal. Consider that by the earlier equation, we also have $A = PDP^{-1}$. If A is full-rank, we can recognize this as an eigendecomposition of A, where P are the eigenvectors and D the corresponding eigenvalues.
However, there is no restriction on the values of the diagonal entries of D. Say some of those entries are 0. Then this suggests that 0 is an eigenvalue of A, and therefore A is not invertible.

Problem B-1

Note that with the singular-value decomposition $A = U\Sigma V^T$ (U and V orthogonal matrices, Σ rectangular-diagonal), we have that

$$AA^T = U(\Sigma\Sigma^T)U^T, \quad A^TA = V(\Sigma^T\Sigma)V^T$$

where we recognize the above as the eigendecompositions of the symmetric matrices AA^T and A^TA. Let us say A is $n \times m$ (so U is $n \times m$, V^T is $m \times m$, and Σ is $n \times m$).

a. True. Recall that $rank(AB) \leq min(rank(A), rank(B))$. This is also valid for more than two matrices at a time —at no point can a **linear** transformation spontaneously increase the span of a set of vectors. So

$$\begin{aligned} rank(U\Sigma V^T) &= \min(rank(U), rank(\Sigma), rank(V^T)) \\ &= \min(n, rank(\Sigma), m) \quad \text{(U and V are orthogonal and so have full rank.)} \\ &= rank(\Sigma) \quad \text{(A matrix's rank must be smaller than its number of rows and columns.)} \\ &= rank(A) \quad (U\Sigma V^T = A) \end{aligned}$$

b. True. We mentioned before that U and V^T correspond to the eigenvector matrices of AA^T and A^TA, so they are immediately determined once A is chosen.
Let's prove by contradiction. Suppose there existed a $\Sigma' \neq \Sigma$ such that $U\Sigma'V^T = A$. Then $U(\Sigma' - \Sigma)V^T = \mathbf{0}$. But $rank(\Sigma' - \Sigma) > 0$, so $rank(U(\Sigma' - \Sigma)V^T > 0)$ while $rank(\mathbf{0}) = 0$. But that means $U(\Sigma' - \Sigma)V^T \neq \mathbf{0}$. A contradiction! So our supposition is incorrect; in fact, there is no other $\Sigma' \neq \Sigma$ that corresponds to A (aside from permuting its rows/columns, which would involve permuting the rows and columns of U and V^T).

c. False. In particular, if A is not symmetric, then $U \neq V$. (Is explanation valid?)

Quick lemma: for a matrix M, MM^T and M^TM are symmetric. Proof: Say that m_j refers to the i'th column vector of M. Then we can describe $MM^T = \sum_i m_i m_i^T$ as a sum of outer products. An outer product

$P = vv^T$ is always symmetric $P_{ij} = v_i v_j = v_j v_i = P_{ij}$, and the sum of symmetric matrices are symmetric (you're performing the same operations to each pair of "matched" entries). So MM^T is symmetric (and therefore so is $(MM^T)^T = M^T M$).

Now for the actual proof of the claim. Let's prove by contradiction. Suppose that A ($n \times n$) is not symmetric and also $U = V$. We can always write out $\Sigma = KK^T$ (for a square Σ, the values of K would be the square-roots of the entries of Σ.) Then $A = UKK^T U^T = (UK)(UK)^T$. In general, MM^T is symmetric (see lemma above). So $(UK)(UK)^T = A$ is symmetric. But we said A is not symmetric. Contradiction! So a part of our initial assumption is incorrect. In this case, we can resolve the contradiction by making $U \neq V$.

d. True. Now that we force A to be symmetric, we have that $AA^T = A^T A$, so they have the same eigenvectors, so in fact we must have that $U = V$.

Problem B-3

We want to find for any symmetric positive (semi-)definite matrix A an orthonormal matrix U such that $U^T A U = D$ for a diagonal matrix with non-negative entries D.(A being positive semi-definite implies $v^T A v \geq 0 \; \forall v \neq \mathbf{0}$.)

Since A is symmetric, we can use singular-value decomposition to write $A = U\Sigma V^T$ with $U = V$ being orthogonal. Multiplying both sides on the left and on the right by U^{-1} (which, because U is orthogonal, has the relation $U^{-1} = U^T$) gives us $\Sigma = U^T A U$.

Note that since A is symmetric, we can read the above as the eigendecomposition of A: $A = U\Lambda U^T$ and $U^T A U = \Lambda$. So by choosing the unit-norm eigenvectors of A to form U, we can get the diagonal matrix of eigenvalues Λ. The eigenvalues of A must be non-negative; if any of them weren't, then we could take its corresponding eigenvector as v_i and have $v_i^T A v_i < 0$, which violates the assumption that A is positive semi-definite. So $\Lambda = \Sigma$ is our diagonal matrix with non-negative entries D.

Problem B-4

With $v_1, \cdots, v_n \in \mathbb{R}^m$ forming the columns of A (an $m \times n$ matrix), and $w \in \mathbb{R}^m$, we want $x \in \mathbb{R}^n$ that minimizes $\|L\|$ with $L = Ax - w$.

a. If A were diagonal, we can minimize L pretty simply. In particular, let $A_t(r \times c)$ be the truncated matrix that removed the rows or columns of A that contain just 0 (more specifically, $r = c = min(n, m)$). Truncate $x_t \in \mathbb{R}^c$ and $w_t \in \mathbb{R}^r$ accordingly. Now A_t is invertible, so we can set $x_t^* = A_t^{-1} w_t$ and have $L_t = A_t x_t^* - w_t := 0$. The "leftover" dimensions of w cannot be captured by Ax and so characterize the minimum value of $\|L\|$.

More simply: we have differing numbers of equations and coordinates, but each equation we do have says "coordinate i of your answer should be $z_i = w_i / A_{i,i}$". So you set your $x_i = z_i$ for all the equations you actually have, and that would minimize the error as much as you can using the information you have.

b. Now suppose we don't have A diagonal, but we know its SVD: $A = U\Sigma V^T$. No problem! We can write

$$Ax - w = U\Sigma V^T x - w = U\Sigma V^T x - UU^T w = U(\Sigma V^T x - U^T w)$$

Then we can say

$$L = U(\Sigma V^T x - U^T w)$$
$$U^T L = \Sigma(V^T x) - U^T w$$

where we're now minimizing $\|U^T L\|$, but since U is orthogonal, $\|U^T L\| = \|L\|$. Using part (a), our diagonal matrix $A \to \Sigma$, $x \to V^T x$, $w \to U^T w$. So we have $(x_t^* = A_t^{-1} w_t) \to ((V^T x)_t^* = \Sigma_t^{-1} U_t^T w_t)$. Finally, we can "unrotate" our optimal solution and get $x_t^* = V_t \Sigma_t^{-1} U_t^T w_t$.

Note that $(U_t \Sigma_t V_t)^{-1} = V_t \Sigma_t^{-1} U_t^T$, so we have that $x_t^* = A_t^{-1} w_t$ —we're still essentially "inverting a truncated version of A" to get our optimal solution.

Note: Rather than performing truncation, one could solve Part (a) and (b) by using the Moore-Penrose pseudoinverse of A. But in most contexts, it's computationally more feasible/convenient to truncate out uninformative rows/columns to greatly reduce space-complexity.

c. Say we instead have $(x_1, y_1), \cdots, (x_n, y_n) \in \mathbb{R}^2$ and we want to find a, b such that $\sum_{i=1}^n (ax_i + b - y_i)^2$ is minimized. This is just a case of the above:

- $m := 2, n := n$
- $x = [b, a]$.
- $A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$.
- $w = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix}^T \in \mathbb{R}^n$.

Technically one is minimizing $\|L\|^2$, but if you minimize $\|L\|$ you've minimized $\|L\|^2$.

d. A simple extension of (c) can find the quadratic equation $y = ax^2 + bx + c$ of best fit.

- Now $m = 3$.
- Change $x := [c, b, a]$.
- Change $A = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix}$.

- 8/12/19

3 Probability Theory and Stochastic Processes

3.1 Problem A-1

(f). We're assuming the time it takes each of $k = 3$ clerks to service a customer follows an exponential distribution. They're currently busy with customers and you're the only person in line. Perhaps shockingly, the probability you're last to leave is $1/k$ – no worse than the other people who had been with clerks for much longer by the time you get your turn!

This is due to the *memoryless property* of exponential distributions, namely that if $X \sim Exp(\lambda)$, then $Pr[X > (s+t) \mid X > t] = Pr[X > s]$ — the probability you'll have to wait a duration of length s is the same no matter how long you've waited before.

So, once you arrive at the clerk's booth, there is symmetry between you and the other customers/clerks. So by symmetry, the probability that you're last to leave is the probability that Person 2 is last to leave, etc, so $Pr[\text{you leave last}] = 1/3$.