Math 219, Linear Algebra I — Fall 2020
Course website: `https://sites.aub.edu.lb/kmakdisi/`
Problem set 9, due Saturday, November 28 at 2pm via Moodle

Exercises from Corwin-Szczarba:
Section 7.3, exercises 3, 4, 5, 6.
Section 10.5, exercises 1ef, 2, 3.
Section 10.6, exercises 1abhij, 6. (In exercise 1, also find a basis for each eigenspace, and diagonalize if possible.)

Additional Exercises (also required):

Exercise A9.1: Consider the matrices

$$A = \begin{pmatrix} 1 & 1 & 1 & 3 \\ 1 & 1 & 2 & 4 \\ 1 & 2 & 0 & 3 \\ 3 & 4 & 3 & 10 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 1 & 1 & 3 \\ 1 & 1 & 2 & 4 \\ 1 & 2 & 0 & 3 \end{pmatrix}.$$

a) Find the determinants $\det A$ and $\det B$.
b) What are the ranks $\operatorname{rank} B$, $\operatorname{rank} C$, and $\operatorname{rank} A$ (in that order)? Justify.

Exercise A9.2: (Elementary matrices — see also Sections 7.6, 7.7 of the book) Let

$$E_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 77 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 88 \\ 0 & 0 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}.$$

The E_i are certain types of 3×3 elementary matrices, and analogs exist for all $n \times n$ matrices.
a) For $i = 1, 2, 3$ find $E_i A$ and AE_i, and interpret in terms of row or column operations on A.
b) Which operations preserve $\ker A$? What about $\operatorname{Image} A$? Why?
c) What do these operations do to $\det A$? (Solve in two ways, once by finding $\det E_i$, and once from the determinant being multilinear and alternating in the rows or columns of A.)
d) Let $\vec{\mathbf{b}} \in \mathbf{R}^3$. We know how to solve the linear system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ for unknown $\vec{\mathbf{x}} \in \mathbf{R}^3$ by Gaussian elimination. Show that each step in Gaussian elimination amounts to replacing A and $\vec{\mathbf{b}}$ by EA and $E\vec{\mathbf{b}}$ for a suitable elementary matrix E (in practice, we often do several steps at once).

Exercise A9.3: Let V be a finite-dimensional inner product space, and let $W \subset V$ be a subspace. Let $T : V \to V$ be the linear transformation "orthogonal projection onto W":

$$T(\vec{\mathbf{v}}) = \operatorname{Proj}_W \vec{\mathbf{v}}.$$

Show that T is diagonalizable, and find the characteristic polynomial of T.

Look at, but do not hand in:
Section 7.3, exercises 7, 10, 11, 12, 13, 14.
Section 7.4, exercises 1, 2, 3, 5, 6, 7, 8.

"Look At" Exercise L9.1: This exercise gives one proof of the "expansion by minors" formula for the determinant. For an $n \times n$ matrix A, and $1 \le i, j \le n$, define the $(n-1) \times (n-1)$ submatrix A_{ij} of A to be the matrix obtained by *removing* the ith row and jth column of A. (For example, A_{1n} is the submatrix in the lower left corner of A.)

a) Assume that $A\vec{\mathbf{e}}_1 = \vec{\mathbf{e}}_1$, so that A has the form $\begin{pmatrix} 1 & * \\ 0 & A_{11} \end{pmatrix}$. Show directly, from the expansion involving permutations σ, that $\det A = \det A_{11}$. This is a primitive version of expanding by minors along the first column.

b) Generalize to the statement that if $A\vec{\mathbf{e}}_j = \vec{\mathbf{e}}_i$, then $\det A = (-1)^{i+j} \det A_{ij}$ ("primitive expansion by minors along the jth column"). Hint: do some "exchange" row operations and column operations, keeping careful track of the sign. From the multilinearity of the determinant, conclude the general expansion by minors along the jth column. I.e., if $A = (a_{ij})_{1 \le i,j \le n}$, then, for any choice of j, $\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{ij}$. Because $\det A = \det A^{tr}$, this implies the expansion by minors along any row of A: $\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A_{ij}$.

c) Show that if $i \neq k$, then $\sum_{j=1}^n (-1)^{i+j} a_{kj} \det A_{ij} = 0$. (Interpret this as the determinant of a suitable matrix.) Conclude that if one defines $\tilde{A} = (b_{ij})_{1 \le i,j \le n}$ by $b_{ij} = (-1)^{i+j} \det A_{ji}$, then $A\tilde{A} = \tilde{A}A = (\det A)I$, where I is the identity matrix. Thus if A is invertible, $A^{-1} = (\det A)^{-1}\tilde{A}$. In particular, this gives a formula for each entry of A^{-1} as a polynomial in the entries of A, divided by the determinant of A, which is the same formula as the one from Cramer's rule (Section 10.7).

“Look At” Exercise L9.2: Find a polynomial $q(x) \in \mathcal{P}_3$ for which

$$\int_{x=0}^{1} \big(e^x - q(x)\big)^2 dx$$

is as small as possible. (Hint: show using exercise 4.7.15 from Corwin-Szczarba that this is a projection from $\mathcal{C}[0,1]$ to its subspace $\mathcal{P}_3$, with respect to the inner product

$$\langle f, g\rangle = \int_{x=0}^{1} f(x)g(x)\, dx.$$

You already found an orthogonal basis for $\mathcal{P}_3$ in Corwin-Szczarba, problem 4.7.2.)

“Look At” exercise L9.3: (Least-squares linear regression) In $\mathbf{R}^7$, define the vectors

$$\begin{aligned} \vec{\mathbf{u}} &= (1,1,1,1,1,1,1), \\ \vec{\mathbf{x}} &= (x_1,x_2,x_3,x_4,x_5,x_6,x_7) = (1,4,5,6,7,9,10), \\ \vec{\mathbf{y}} &= (y_1,y_2,y_3,y_4,y_5,y_6,y_7) = (0.9,1.6,2.2,1.9,2.8,2.9,3.8). \end{aligned}$$

a) Let $W = \operatorname{span}\{\vec{\mathbf{u}}, \vec{\mathbf{x}}\}$. Compute an orthogonal basis for W and use it to find $\vec{\mathbf{z}} = \operatorname{Proj}_W \vec{\mathbf{y}}$.

b) Find a, b such that $\vec{\mathbf{z}} = a\vec{\mathbf{x}} + b\vec{\mathbf{u}}$.

c) In $\mathbf{R}^2$, draw the points (x_1, y_1), (x_2, y_2), and so on (i.e., these are the points with coordinates $(1, 0.9), (4, 1.6), \ldots, (10, 3.8)$.) Also draw the line $y = ax + b$ corresponding to the values of a and b from part (b) above. Explain why the line passes very close to the points. (Hint: why is the “vector of errors” $\vec{\mathbf{e}} = \vec{\mathbf{y}} - a\vec{\mathbf{x}} - b\vec{\mathbf{u}}$ so short?)

“Look At” Exercise L9.4: Let V be a finite-dimensional inner product space. Let $\vec{\mathbf{z}} \neq \vec{\mathbf{0}}$, and consider the linear transformation $T : V \to V$, given by

$$T(\vec{\mathbf{x}}) = \vec{\mathbf{x}} - 2\frac{\langle \vec{\mathbf{x}}, \vec{\mathbf{z}}\rangle}{\langle \vec{\mathbf{z}}, \vec{\mathbf{z}}\rangle}\vec{\mathbf{z}}.$$

a) Show that $T(\vec{\mathbf{x}})$ is the reflection of $\vec{\mathbf{x}}$ with respect to the hyperplane $\{\vec{\mathbf{z}}\}^\perp$. Draw a simple picture as part of your explanation.

b) Show that for all $\vec{\mathbf{x}} \in V$, $\|T(\vec{\mathbf{x}})\| = \|\vec{\mathbf{x}}\|$. (If you prefer, show directly that T is an isometry.)

c) Describe a “nice” basis β of V with respect to which ${}_\beta[T]_\beta$ is “easy” to understand. (Hint: it will be a diagonal matrix.)

“Look At” Exercise L9.5: (Another proof that row rank = column rank) Let A be an $m \times n$ matrix, and view A as a linear transformation from $\mathbf{R}^n$ to $\mathbf{R}^m$ (with the standard bases on both vector spaces). Recall that we have defined the column space $CS(A) \subset \mathbf{R}^m$ to be the span of the columns of A, and the row space $RS(A) \subset \mathbf{R}^n$ to be the span of the rows of A. We are used to the fact that $CS(A) = \operatorname{Image} A$.

a) Show that $\ker A = \big(RS(A)\big)^\perp$. So the kernel and the row space are orthogonal complements.

b) From the formula $\dim W^\perp + \dim W = \dim V$ for general subspaces $W \subset V$ of an inner product space, deduce that the row rank of A is equal to its column rank:

$$\dim RS(A) = \dim CS(A).$$

We are thus justified in calling this common value the **rank** of A.

c) Show that (i) A is injective $\iff$ $\operatorname{rank} A = n$ $\iff$ A^{tr} is surjective; and that (ii) A is surjective $\iff$ $\operatorname{rank} A = m$ $\iff$ A^{tr} is injective.

Cultural note: statement (ii) in part (c) can be reinterpreted as saying that A is **not** surjective $\iff$ there exists a nonzero $1 \times m$ matrix $c = \begin{pmatrix} c_1 & c_2 & \cdots & c_m \end{pmatrix}$ with $cA = \vec{\mathbf{0}}$. The existence of c amounts to saying that the image of A is contained in the hyperplane $c_1x_1 + \cdots + c_mx_m = 0$.