Math 54, Spring 2009, Sections 109 and 112
Midterm 2 Review Exercises
Solutions

These exercises don't cover some of the **very important** computational-type problems, including many of the things listed under the "be able to" section of the review sheet. You can find examples of those types of problems on the sample exam and in the sections of the book (including the supplemental exercises at the end of each chapter). These problem are a little more theoretical, and are aimed at making sure you have a good grasp of the ideas underlying the algorithms.

1) (a) Assume that A is an $m \times n$ matrix and that B is an $n \times p$ matrix. Show that $AB = 0$ (the zero matrix) if and only if $\operatorname{Col} B \subseteq \operatorname{Nul} A$.

(b) What can you say about $\operatorname{Rank} A$ and $\operatorname{Rank} B$?

(a) First suppose that $AB = 0$. To show that $\operatorname{Col} B \subseteq \operatorname{Nul} A$, we take an arbitrary $y \in \operatorname{Col} B$ and show that it is in $\operatorname{Nul} A$. If $y \in \operatorname{Col} B$, then it is of the form $B\vec{x}$ for some $\vec{x}$. Then $A\vec{y} = AB\vec{x} = 0\vec{x} = \vec{0}$, so $\vec{y} \in \operatorname{Nul} A$ by the defintion of $\operatorname{Nul} A$.

Now suppose the converse, that $\operatorname{Col} B \subseteq \operatorname{Nul} A$, and we'll show that $AB = 0$. Since $B\vec{x} \in \operatorname{Col} B$ for every $\vec{x} \in \mathbb{R}^p$, our assumption says that $B\vec{x} \in \operatorname{Nul} A$. Thus $AB\vec{x} = \vec{0}$ for all $\vec{x}$. In particular, $AB\vec{e}_1 = \vec{0}$, so the first column of AB is $\vec{0}$. Continuing, we can show that all of the columns are $\vec{0}$ in this fashion, so $AB = 0$.

(b) Since $\operatorname{Col} B \subseteq \operatorname{Nul} A$, we must have $\dim \operatorname{Col} B \leq \dim \operatorname{Nul} A$. That is, $\operatorname{Rank} B \leq \dim \operatorname{Nul} A$. By the rank-nullity theorem, $\dim \operatorname{Nul} A = n - \operatorname{Rank} A$, so $\operatorname{Rank} B \leq n - \operatorname{Rank} A$. That is, $\operatorname{Rank} A + \operatorname{Rank} B \leq n$.

2) [p.371, #3] Suppose $\vec{x}$ is an eigenvector of A corresponding to an eigenvalue λ. Show that $\vec{x}$ is an eigenvector of $5I - 3A + A^2$. What is its eigenvalue?

Since $\vec{x}$ is an eigenvector of A, $\vec{x} \neq 0$. We can compute

$$\begin{aligned}(5I - 3A + A^2)\vec{x} &= 5I\vec{x} - 3A\vec{x} + A^2\vec{x} \\ &= 5\vec{x} - 3\lambda\vec{x} + A\lambda\vec{x} \\ &= 5\vec{x} - 3\lambda\vec{x} + \lambda^2\vec{x} \\ &= (5 - 3\lambda + \lambda^2)\vec{x}.\end{aligned}$$

So $\vec{x}$ is an eigenvector of $(5I - 3A + A^2)$ with eigenvalue $(5 - 3\lambda + \lambda^2)$.

3) Find a 2×2 matrix A such that $A^2 + 6I = 5A$. What if we require that A not be diagonal?

We can rearrange the given equation to read $A^2 - 5A + 6I = 0$, or alternatively $(A - 2I)(A - 3I) = 0$. One can now check that the following matrices satisfy the equation:

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}.$$

This works because diagonal matrices multiply entry-by-entry, which is not true in general. So what can we do to make this work with a non-diagonal matrix? Assume that P is some invertible matrix. If $A^2 - 5A + 6I = 0$, then

$$\begin{aligned}0 &= P^{-1}0P \\ &= P^{-1}(A^2 - 5A + 6I)P \\ &= P^{-1}A^2P - 5P^{-1}AP + 6P^{-1}P \\ &= (P^{-1}AP)^2 - 5(P^{-1}AP) + 6I.\end{aligned}$$

That is, $P^{-1}AP$ satisfies the same equation. So we can take one of our diagonal examples that worked, and find an invertible matrix P such that $P^{-1}AP$ is not diagonal. For instance,

$$A = \begin{bmatrix} 1 & 4 \\ 1 & 5 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} -2 & -20 \\ 1 & 7 \end{bmatrix}$$

has the desired property. Note: in finding a matrix that satisfies $A^2 - 5A + 6I = 0$, we found one whose eigenvalues satisfy $\lambda^2 - 5\lambda + 6 = 0$. This is not a coincidence; it is part of a theorem called the Cayley-Hamilton Theorem.

4) [p.371, #1] True or false? If true, explain why, and if false provide a counterexample. Assume all matrices are square.

(a) If A contains a row of zeros, then 0 is an eigenvalue of A.

(b) Every eigenvector of A is also an eigenvector of A^2.

(c) If A is diagonalizable, then the columns of A are linearly independent.

(d) If A and B are invertible $n \times n$ matrices, then AB is similar to BA.

(e) If A is an $n \times n$ diagonalizable matrix, then every vector in $\mathbb{R}^n$ can be written as a linear combination of eigenvectors of A.

(a) True. If A contains a row of all 0, then cofactor expansion across this row says that $\det A = 0$ thus A is not invertible. This means that 0 is an eigenvalue of A. Alternatively, A^T has a column of all zeros. If that is the k-th column of A, then $\vec{e}_k$ is an eigenvector of A^T with eigenvalue 0. A and A^T have the same eigenvalues, so 0 is an eigenvalue of A.

(b) True. If $x \neq 0$ and $A\vec{x} = \lambda\vec{x}$, then $A^2\vec{x} = \lambda^2\vec{x}$.

(c) False. $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is diagonal(izable), but has a column that is the zero vector. Any set that contains $\vec{0}$ is linearly dependent.

(d) True. $AB = B^{-1}(BA)B$ says that AB is similar to BA.

(e) True. A matrix is diagonalizable if and only if $\mathbb{R}^n$ has a basis of eigenvectors of A, in which case every vector in $\mathbb{R}^n$ can be written as a linear combination of the elements of this basis.

5) (a) Suppose that A is an $n \times m$ matrix. Show that $A^T A\vec{x} \cdot \vec{x} \geq 0$ for every $\vec{x} \in \mathbb{R}^m$.

(b) Show that if $\|\vec{x}\| \leq 1$, then $\|A\vec{x}\|^2 \leq \left\|A^T A\vec{x}\right\|$.

(a) Recall that if C is $n \times n$ and $\vec{x} \in \mathbb{R}^n$, we have $C\vec{x} \cdot y = \vec{x} \cdot C^T\vec{y}$. Thinking of $A^T A\vec{x}$ as $A^T(A\vec{x})$, this means that

$$A^T A\vec{x} \cdot \vec{x} = A\vec{x} \cdot A\vec{x} = \|A\vec{x}\|^2 \geq 0.$$

(b) Picking up where we left off, we get

$$\begin{aligned} \|A\vec{x}\|^2 &= A^T A\vec{x} \cdot \vec{x} \\ &\leq \left\|A^T A\vec{x}\right\| \|\vec{x}\| \\ &\leq \left\|A^T A\vec{x}\right\|. \end{aligned}$$

(The equality step follows from part (a), the first inequality is the Cauchy-Schwarz inequality (p.432), and the second inequality comes from the fact that $\|\vec{x}\| \leq 1$.)

6) Suppose that $\vec{y} \in \mathbb{R}^n$, that $\|y\| = 1$, and that W is a subspace of $\mathbb{R}^n$. Show that $\vec{y} = \mathrm{Proj}_W \vec{y} + \mathrm{Proj}_{W^\perp} \vec{y}$ and that $\|\vec{y}\|^2 = \|\mathrm{Proj}_W \vec{y}\|^2 + \|\mathrm{Proj}_{W^\perp} \vec{y}\|^2$.

The Orthogonal Decomposition Theorem (p.395) says that $\vec{y}$ can be written uniquely in the form $\vec{y} = \mathrm{Proj}_W \vec{y} + \vec{z}$, $\vec{z} \in W^\perp$. On the other hand, it also says that $\vec{y}$ can be written uniquely in the form $\vec{y} = \mathrm{Proj}_{W^\perp} \vec{y} + \vec{w}$ where $w \in (W^\perp)^\perp = W$. Since these decompositions are supposed to be unique, they must be the same, so $\vec{z} = \mathrm{Proj}_{W^\perp} \vec{y}$ and $\vec{w} = \mathrm{Proj}_W \vec{y}$. Hence $\vec{y} = \mathrm{Proj}_W \vec{y} + \mathrm{Proj}_{W^\perp} \vec{y}$. Since $\mathrm{Proj}_W \vec{y} \in W$ and $\mathrm{Proj}_{W^\perp} \vec{y} \in W^\perp$, we have $\mathrm{Proj}_W \vec{y} \perp \mathrm{Proj}_{W^\perp} \vec{y}$. The statement about norms now follows immediately from the Pythagoren Theorem.

(7) Let V be the inner product space $C[0,1]$ of all continuous functions defined on the interval $[0,1]$, with inner product

$$\langle f, g\rangle = \int_0^1 f(x)g(x)dx.$$

Find a non-zero function that is orthogonal to both $f_1(x) = 1$ and $f_2(x) = x^3$.

Let $W = \operatorname{Span}\{f_1, f_2\}$. The idea is to find non-zero something in $W^\perp$ (which will be orthogonal to f_1 and f_2). We do this by taking some function g which is *not* in W, and computing $g - \operatorname{Proj}_W g$, which will be in $W^\perp$. So we are free to start with any function $g \notin W$. Let's choose $g(x) = x^2$. To compute $\operatorname{Proj}_W g$, we need an orthogonal basis for W, so we use Gram-Schmidt. Set $h_1(x) = f_1(x) = 1$, and $h_2(x) = f_2(x) - \operatorname{Proj}_{\operatorname{Span} h_1} f_2(x) = x^3 - \frac{\langle h_1, f_2\rangle}{\langle h_1, h_1\rangle} = x^3 - \frac{1}{4}$. Now we compute

$$\operatorname{Proj}_W g(x) = \frac{\langle g, h_1\rangle}{\langle h_1, h_1\rangle} h_1(x) + \frac{\langle g, h_2\rangle}{\langle h_2, h_2\rangle} h_2(x) = \frac{1}{3} + \frac{28}{27}(x^3 - \frac{1}{4}).$$

Thus $g(x) - \operatorname{Proj}_W g(x) = x^2 - \frac{1}{3} - \frac{28}{27}(x^3 - \frac{1}{4}) = x^2 - \frac{28}{27}x^3 - \frac{2}{27}$. One can check that this is orthogonal to both f_1 and f_2. On an exam, the numbers would come out nicer...

(8) Let $\mathcal{B}$ be a basis for a vector space V. Use the fact that $[\cdot]_B$ is an isomorphism to prove that $\{\vec{v}_1, \ldots, \vec{v}_n\}$ spans V if and only if $\{[\vec{v}_1]_B, \ldots, [\vec{v}_n]_B\}$ spans $\mathbb{R}^m$.

First suppose that $\{\vec{v}_1, \ldots, \vec{v}_n\}$ spans V, and we will show that $\{[\vec{v}_1]_B, \ldots, [\vec{v}_n]_B\}$ spans $\mathbb{R}^m$. Fix $\vec{y} \in \mathbb{R}^m$, and we will show that this arbitrary vector is in the span of the given coordinate vectors. Since $[\cdot]_B$ is an isomorphism, it is onto. Thus, there is some $\vec{x} \in V$ such that $[\vec{x}]_B = \vec{y}$. Since $\{\vec{v}_1, \ldots, \vec{v}_n\}$ spans V, there are coefficients $c_1, \ldots, c_n$ such that $\vec{x} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$. Thus

$$\vec{y} = [\vec{x}]_B = [c_1\vec{v}_1 + \cdots + c_n\vec{v}_n]_B = c_1[\vec{v}_1]_B + \cdots + c_n[\vec{v}_n]_B.$$

Thus $\vec{y} \in \operatorname{Span}\{[\vec{v}_1]_B, \ldots, [\vec{v}_n]_B\}$, and since $\vec{y}$ was arbitrary, this set spans $\mathbb{R}^m$.

Conversely, suppose that $\{[\vec{v}_1]_B, \ldots, [\vec{v}_n]_B\}$ spans $\mathbb{R}^m$. We wish to show that $\{\vec{v}_1, \ldots, \vec{v}_n\}$ spans V, so fix $\vec{x} \in V$, and we will show that it is a linear combination of these vectors.

Since $[\vec{x}]_B \in \mathbb{R}^m$, our assumption says that there are coefficients $c_1, \ldots, c_n$ such that $[\vec{x}]_B = c_1[\vec{v}_1]_B + \cdots c_n[\vec{v}_n]_B = [c_1\vec{v}_1 + \cdots c_n\vec{v}_n]_B$. Since $[\cdot]_B$ is an isomorphism, it is one-to-one. Thus, the previous equality implies that $\vec{x} = c_1\vec{v}_1 + \cdots c_n\vec{v}_n$. Thus $\vec{x} \in \operatorname{Span}\{\vec{v}_1, \ldots, \vec{v}_n\}$. Since $\vec{x}$ was arbitrary, this set spans V.