CSM 16A Designing Information Systems and Devices

Week 11 Worksheet Metas

Term: **Spring 2020** *Name:*

Problem 1: Linear Algebra Review

Meta: Description: Meant to be a review of eigenvectors, eigenvalues and transformations.

1. Suppose λ is an eigenvalue for the matrix A. Consider the λ-eigenspace of A: the set of all vectors v satisfying the equation $\mathbf{A}\vec{v} = \lambda\vec{v}$. Show that this eigenspace is a subspace by directly checking the three conditions needed to be a subspace.

   Meta: This gives students some review of eigenvalues, eigenvectors and subspaces. It might help to start this question by prompting students to state what the three conditions are to prove that the eigenspace is a subspace.

   Solution: First, we have to check that $\vec{0}$ is in the subspace: this is true because $\mathbf{A}\vec{0} = \lambda\vec{0} = \vec{0}$ (regardless of what the eigenvalue λ is).

   Next, suppose $\vec{u}$ and $\vec{v}$ are in the subspace. This means that:

$$\mathbf{A}\vec{u} = \lambda\vec{u}$$
$$\mathbf{A}\vec{v} = \lambda\vec{v}$$

$$\mathbf{A}(\vec{u}+\vec{v}) = \mathbf{A}\vec{u} + \mathbf{A}\vec{v} = \lambda\vec{u} + \lambda\vec{v} = \lambda(\vec{u}+\vec{v})$$

   This means $\vec{u}+\vec{v}$ is also in the subspace.

   Finally, suppose $\vec{v}$ is in the subspace and r is a scalar. Then,

$$\mathbf{A}(r\vec{v}) = r(\mathbf{A}\vec{v}) = r(\lambda\vec{v}) = \lambda(r\vec{v})$$

   This means that $r\vec{v}$ is also in the subspace.

   Since the eigenspace satisfies all three conditions of being a subspace, we can say that it is a subspace.

2. Solve for the eigenvalue-eigenvector pairs for the following 2 by 2 matrix:

$$\mathbf{A} = \begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix}$$

   Meta: This is a mechanical question, if you feel that your students are comfortable with solving for eigenvalues and eigenvectors you can consider skipping this part.

   Solution: To solve for eigenvalues and eigenvectors, let's go back and review the definition of eigenvectors and eigenvalues:
   If $\vec{x}$ and λ are the eigenvector and eigenvalue of $\mathbf{A}$, respectively, then the following equation holds:

$$\mathbf{A}\vec{x} = \lambda\vec{x}$$

Since the (appropriately sized) identity matrix is analogous to multiplying by 1 in arithmetic, we can say:

$$\mathbf{A}\vec{x} = \lambda\mathbf{I}\vec{x}$$

Rearranging, we get:

$$\mathbf{A}\vec{x} - (\lambda\mathbf{I})\vec{x} = \vec{0}$$
$$(\mathbf{A} - \lambda\mathbf{I})\vec{x} = \vec{0}$$

What does this look like? It looks similar to solving for the nullspace of $(\mathbf{A} - \lambda\mathbf{I})$!
Assuming that there is a nontrivial nullspace, that also means that $\mathbf{det}(\mathbf{A} - \lambda\mathbf{I}) = 0$!
Let's solve for λ first:

$$\begin{aligned}(\mathbf{A} - \lambda\mathbf{I}) &= \begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \\ &= \begin{bmatrix} 4-\lambda & 1 \\ 3 & 2-\lambda \end{bmatrix} \\ \mathbf{det}(\mathbf{A} - \lambda\mathbf{I}) &= (4-\lambda)(2-\lambda) - 3 \\ &= 5 - 6\lambda + \lambda^2 \\ &= (\lambda - 5)(\lambda - 1)\end{aligned}$$

By factoring:

$$\lambda = 5, 1$$

Let's check: We've just solved for the eigenvalues. But what about the eigenvectors?
To do that, we plug in λ into $(\mathbf{A} - \lambda\mathbf{I})$ and solve for the nullspace!
For $\lambda = 5$:

$$(\mathbf{A} - \lambda\mathbf{I})\vec{x} = \vec{0}$$
$$\begin{bmatrix} -1 & 1 \\ 3 & -3 \end{bmatrix}\vec{x} = \vec{0}$$

We can see that eigenvector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ spans the nullspace of the above matrix.
So the first pair is

$$\lambda = 5, \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Repeating for $\lambda = 1$,

$$(\mathbf{A} - \lambda\mathbf{I})\vec{x} = \vec{0}$$
$$\begin{bmatrix} 3 & 1 \\ 3 & 1 \end{bmatrix}\vec{x} = \vec{0}$$

We can see that eigenvector $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$ spans the nullspace of the above matrix.
So, the second pair is

$$\lambda = 1, \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

3. Projection of a vector $\vec{u}$ onto $\vec{v}$ is given by:

$$\frac{\vec{u} \cdot \vec{v}}{||\vec{v}||^2}\vec{v}$$

Prove that projection onto a vector $\vec{v}$ is a linear transformation.

Meta: Students might ask where this formula is derived from, it helps to be prepared for this.

Solution: Let us represent this transformation using P.

$$P(\vec{a}) = \frac{\vec{a} \cdot \vec{v}}{||\vec{v}||^2}\vec{v}$$

Let's check if it satisfies the condition of linearity.

$$P(\vec{a} + \vec{b}) = \frac{(\vec{a} + \vec{b}) \cdot \vec{v}}{||\vec{v}||^2}\vec{v}$$

$$P(\vec{a} + \vec{b}) = \frac{\vec{a} \cdot \vec{v}}{||\vec{v}||^2}\vec{v} + \frac{\vec{b} \cdot \vec{v}}{||\vec{v}||^2}\vec{v}$$

$$P(\vec{a} + \vec{b}) = P(\vec{a}) + P(\vec{b})$$

Hence, the projection transformation satisfies additivity. Let's check if it satisfies the condition of scalar multiplication.

$$P(r\vec{a}) = \frac{(r\vec{a}) \cdot \vec{v}}{||\vec{v}||^2}\vec{v}$$

$$P(r\vec{a}) = r \cdot \frac{(\vec{a}) \cdot \vec{v}}{||\vec{v}||^2}\vec{v}$$

$$P(r\vec{a}) = r \cdot P(\vec{a})$$

Hence, the projection transformation is a linear transformation as it satisfies both the conditions - vector addition and scalar multiplication.

Problem 2: Introduction to Inner Products

Learning Goal:

Description: goes over definition, properties, and simple applications of inner products

Prereqs: basic linear algebra, i.e. what vectors are

1. What is an inner product?

   Solution: An inner product describes a way to multiply vectors, such that the result is a scalar. It is often used to describe properties such as the length of a vector, the angle between vectors, orthogonality of vectors, etc.

   An inner product must satisfy the following properties:

   1. Symmetry: $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$
   2. Homogeneity: $\langle c\vec{x}, \vec{y} \rangle = c \langle \vec{x}, \vec{y} \rangle$
   3. Additivity: $\langle \vec{x} + \vec{y}, \vec{z} \rangle = \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle$
   4. Positive-definiteness: $\langle \vec{x}, \vec{x} \rangle \geq 0$, and is $= 0$ iff $\vec{x} = \vec{0}$

2. What is the dot product between two vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$?

   Solution: The dot product is defined as the sum of element-wise products, i.e.

   $$x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$$

   In the next four parts, we prove that the dot product is an inner product. Do note that the dot product is simply a type of inner product, and other inner products are also possible.

3. Prove that the dot product satisfies symmetry, i.e. that $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$

   Solution: First, we write the definition of a dot product again:

   $$x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$$

   Since x_i and y_i are just scalars, and we know that scalar multiplication commutes, we can rewrite this as:

   $$y_1 x_1 + y_2 x_2 + \ldots + y_n x_n = \langle \vec{y}, \vec{x} \rangle$$

4. Prove that the dot product satisfies homogeneity, i.e. that $\langle c\vec{x}, \vec{y} \rangle = c \langle \vec{x}, \vec{y} \rangle$: $c \in \mathbb{R}$

   Solution: Writing out $\langle c\vec{x}, \vec{y} \rangle$, we have:

   $$cx_1 y_1 + cx_2 y_2 + \ldots + cx_n y_n$$

   Since there is a c in every term, we can pull it out, getting:

   $$c(x_1 y_1 + x_2 y_2 + \ldots + x_n y_n) = c \langle \vec{x}, \vec{y} \rangle$$

5. Prove that the dot product satisfies additivity, i.e. that $\langle \vec{x} + \vec{y}, \vec{z} \rangle = \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle$

   Solution: Writing out $\langle \vec{x} + \vec{y}, \vec{z} \rangle$, we get:

   $$(x_1 + y_1) z_1 + (x_2 + y_2) z_2 + \ldots + (x_n + y_n) z_n$$

   distributing we get

   $$x_1 z_1 + x_2 z_2 + \ldots + x_n z_n + y_1 z_1 + y_2 z_2 + \ldots + y_n z_n = \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle$$

6. Prove that the dot product satisfies positive-definiteness, i.e., that $\langle \vec{x}, \vec{x} \rangle \geq 0$, and is equal to 0 iff $\vec{x} = \vec{0}$

   Solution: $\langle \vec{x}, \vec{x} \rangle = x_1^2 + x_2^2 + \ldots + x_n^2$. Since each term in this sum is ≥ 0, $\langle \vec{x}, \vec{x} \rangle \geq 0$. Also, $\langle \vec{x}, \vec{x} \rangle$ is clearly 0 only when $x_1, x_2, \ldots, x_n = 0$, i.e., $\vec{x} = \vec{0}$

   We will now consider ways to use dot products to do neat things. For each of the following, assume that you're given a $\vec{x}$, and that you get a pick $\vec{y}$ of your choosing. Describe a $\vec{y}$, such that when you compute $\langle \vec{x}, \vec{y} \rangle$, you get:

7. The sum of every element in $\vec{x}$

   Solution: We can do this by setting $\vec{y} = \vec{1}$ taking the following dot product:

$$\left\langle \vec{1}, \vec{x} \right\rangle = 1x_1 + 1x_2 + \ldots + 1x_n$$

8. The sum of certain elements in $\vec{x}$

   Solution: We can do this by letting $\vec{y}$ be a vector of 1s and 0s, where the ones are in the positions corresponding to the desired elements.

9. The mean of all the items in $\vec{x}$ (for $\vec{x}$ in $\mathbb{R}^n$)

   Solution: For this case, we can have some vector $\vec{y}$, where every element is $\frac{1}{n}$, so we have:

$$\langle \vec{x}, \vec{y} \rangle = \frac{1}{n}x_1 + \frac{1}{n}x_2 + \ldots + \frac{1}{n}x_n = \frac{1}{n}(x_1 + x_2 + \ldots + x_n)$$

10. The sum of the elements of $\vec{x}$ squared

    Solution: For this case, we can just take the dot product of x with itself,

$$\langle \vec{x}, \vec{x} \rangle = x_1^2 + x_2^2 + \ldots + x_n^2$$

   We will conclude by making some observations based on that last case.

11. Consider that last case, where we summed the squares of the elements of a vector. Try doing that for a few 2-dimensional vectors (vectors of length 2). What do you notice about the resulting answer? What about for vectors of length 3, or for vectors of any length n?

    Solution: After trying out a few examples, you may notice that the dot product of a vector with itself is the square of the length of the vector! Another way to see this is to think of the normal euclidean distance equation:

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

    This can be generalized to any number of dimensions. Now, consider that $\vec{y}$ is a vector of all zeroes, we now have an equation which is exactly the square root of $\langle \vec{x}, \vec{x} \rangle$, which we also name the $\ell 2$-norm of $\vec{x}$, or $\|x\|_2$, or also $\|x\|$.

Problem 3: Eigenspace, Orthogonality, and Symmetric Matrices

Suppose we have a matrix $A \in \mathbb{R}^{n\times n}$.

Meta: Make sure to instruct the students to do these parts sequentially, because later parts rely on the proofs of previous parts. These parts are derivations for some fundamental properties of Eigenspaces and Orthogonality.

1. Show that if $\vec{v}$ is an eigenvector of A, then it must also be an eigenvector of A^2.

   Solution: Suppose $A\vec{v} = \lambda\vec{v}$. Left multiply both sides by A, we have $A^2\vec{v} = A \cdot \lambda\vec{v}$. Since λ is a constant, we can switch its position with A on the right side. This gives us:

   $$A^2\vec{v} = \lambda \cdot A\vec{v} = \lambda \cdot \lambda\vec{v} = \lambda^2\vec{v}.$$

2. Show that if $\vec{u}$ is an eigenvector of A with associated eigenvalue α, and $\vec{v}$ is an eigenvector of A^T with associated eigenvalue β, if $\alpha \neq \beta$, then $\vec{u}$ and $\vec{v}$ must be orthogonal to each other.

   Solution: From what's given in the question, we know that:

   $$A\vec{u} = \alpha\vec{u},$$

   $$A^T\vec{v} = \beta\vec{v}.$$

   To show $\vec{u}$ and $\vec{v}$ are orthogonal to each other, we must show that $\vec{u}^T\vec{v} = 0$.

   Since we have:

   $$A\vec{u} = \alpha\vec{u},$$

   Left multiply the first equation by $\vec{v}^T$. This gives us:

   $$\vec{v}^TA\vec{u} = \vec{v}^T\alpha\vec{u} = \alpha\vec{v}^T\vec{u}$$

   At the same time, note the following:

   $$\vec{v}^TA\vec{u} = (A^T\vec{v})^T\vec{u} = (\beta\vec{v})^T\vec{u} = \beta\vec{v}^T\vec{u}$$

   Therefore, we can see that:

   $$\alpha\vec{v}^T\vec{u} = \beta\vec{v}^T\vec{u}$$

   $$(\alpha - \beta)\vec{v}^T\vec{u} = 0$$

   Since $\alpha \neq \beta$, $\alpha - \beta \neq= 0$, then it must be that $\vec{v}^T\vec{u} = \vec{u}^T\vec{v} = 0$.

   Therefore, $\vec{u}$ and $\vec{v}$ must be orthogonal to each other.

   For the following parts, assume A is also symmetric.

3. Show that A has all real eigenvalues.

   Solution:

   Without loss of generality, let $(\lambda, \vec{v})$ be any eigenvalue-vector pair of A.

   We have $A\vec{v} = \lambda\vec{v}$.

   Consider the expression $\vec{v}^TA^TA\vec{v}$, we have:

   $$\vec{v}^TA^TA\vec{v} = (A\vec{v})^TA\vec{v} = \langle A\vec{v}, A\vec{v}\rangle = \|A\vec{v}\|^2$$

   Since A is also symmetric, $A = A^T$.

   At this point, using what we have shown in part 1 of problem 3, we also have:

   $$\vec{v}^TA^TA\vec{v} = \vec{v}^TA^2\vec{v} = \vec{v}^T\lambda^2\vec{v} = \lambda^2\vec{v}^T\vec{v} = \lambda^2\|\vec{v}\|^2$$

We can see:

$$\|A\vec{v}\|^2 = \lambda^2 \|\vec{v}\|^2$$

$$\lambda^2 = \frac{\|A\vec{v}\|^2}{\|\vec{v}\|^2}$$

Since $\|A\vec{v}\|^2 > 0, \|\vec{v}\|^2 > 0$, we can see that $\lambda^2 =$ some positive number.

Hence, λ must be real.

4. Using the result from part 2, explain why the eigenvectors of A are orthogonal to each other. (If the set of all eigenvectors are orthogonal to each other, we call the set an *orthogonal eigenbasis*)

Solution: Since A is symmetric, $A = A^T$. Suppose $\vec{u}$ is an eigenvector of A with associated eigenvalue α, and $\vec{v}$ is another eigenvector of A with associated eigenvalue β.

Slightly modifying the proof from part 2 of problem 3, we can see that

$$A\vec{u} = \alpha\vec{u},$$

$$A\vec{v} = A^T\vec{v} = \beta\vec{v}.$$

Now the rest of the proof from part 2 follows.

Problem 4: Robust Linear Systems

Up and till now, we have been extensively studying different examples of linear systems represented by the iconic matrix vector equation $A\vec{x} = \vec{v}$ and how to solve them.

However, we haven't looked much into the sensitivity of a linear system to external changes. In particular, how the solutions to such linear systems react to small changes (we call these changes *perturbations*) in A or b can be of great importance to designing a system *robust* to changes.

In this question, we will work toward deriving a well-know metric used to measure such sensitivity to *perturbations* within the system.

1. To get started, consider the following linear system:

$$\begin{bmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 32 \\ 23 \\ 33 \\ 31 \end{bmatrix}$$

First, find the solution to this system. Then, consider the following linear system with some slight *perturbation* to the right-hand side (i.e. $\vec{b}$).

$$\begin{bmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 32.1 \\ 22.9 \\ 33.1 \\ 30.9 \end{bmatrix}$$

Find its solution, and compare how much it has changed from the previous system to how much $\vec{b}$ has changed from the previous system. What did you notice? Is this linear system sensitive to *perturbations*?

Solution: If we solve the first system, we can find its solution to be:

$$\vec{x}_{original} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^T$$

If we solve the second system, we can find its solution to be:

$$\vec{x}_{perturbed} = \begin{bmatrix} 9.2 & -12.6 & 4.5 & -1.1 \end{bmatrix}^T$$

As we can see, if we look at how $\vec{b}$ changes, every entry has only either increased or decreased by 0.1, on an order of about 1/200 with respect to its original value.

However, if we look at how $\vec{x}_{perturbed}$ changes from $\vec{x}$. We can see most of them has changed by the order of about 10/1. Overall, this represents an amplification of a relative error between $\vec{b}$ and $\vec{x}$ on the order of 2000.

This linear system is clearly sensitive to *perturbations*!

2. Before moving forward, let us provide the following definition of **a norm that applies to matrices**.

We define the spectral norm on a matrix A as the greatest possible value of the vector norm $\|A\vec{v}\|$ for all unit-length vectors $\vec{v}$.

In other words,

$$\|A\| = \max_{\|\vec{v}\|=1} \|A\vec{v}\|$$

In addition, assume the following property holds:

$$\|A\vec{v}\| \leq \|A\| \, \|\vec{v}\|$$

Let's first study the case where we *perturb* $\vec{b}$ slightly. Specifically, given an **invertible** matrix A, we have the following pair of solutions to a linear system and a lightly perturbed one:

$$A\vec{v} = \vec{b}$$

$$A(\vec{v} + \delta\vec{v}) = \vec{b} + \delta\vec{b}$$

Here, $\delta\vec{v}$ and $\delta\vec{b}$ represents the slight perturbation in the system.

Show that we can find some constant c such that:

$$\frac{\|\delta\vec{v}\|}{\|\vec{v}\|} \leq c \cdot \frac{\left\|\delta\vec{b}\right\|}{\left\|\vec{b}\right\|}$$

For those interested, we call this constant the *condition number*.

Meta: Students might be confused with what $\delta\vec{v}$ and $\delta\vec{b}$ mean in the context of perturbation. An intuitive way to explain this is we are essentially just taking the equation $A\vec{v} = \vec{b}$ and multiplying both sides by δ (a really small number), like dx when we are integrating a function $f(x)$.

Solution: Starting with the given matrix-vector equations, we have:

$$A\vec{v} = \vec{b}$$

$$A\vec{v} + A\delta\vec{v} = \vec{b} + \delta\vec{b}$$

Subtracting the first equation from the second one, we have:

$$A\delta\vec{v} = \delta\vec{b}$$

$$\delta\vec{v} = A^{-1}\delta\vec{b}$$

Applying the matrix norm inequality, we notice that:

$$\left\|A^{-1}\delta\vec{b}\right\| = \|\delta\vec{v}\| \leq \left\|A^{-1}\right\| \left\|\delta\vec{b}\right\|$$

Applying the inequality to $A\vec{v} = \vec{b}$, we have:

$$\left\|\vec{b}\right\| \leq \|A\| \|\vec{v}\|$$

Hence, we can multiply the two inequalities:

$$\|\delta\vec{v}\| \left\|\vec{b}\right\| \leq \left\|A^{-1}\right\| \left\|\delta\vec{b}\right\| \|A\| \|\vec{v}\|$$

$$\frac{\|\delta\vec{v}\|}{\|\vec{v}\|} \leq (\|A\| \left\|A^{-1}\right\|) \cdot \frac{\left\|\delta\vec{b}\right\|}{\left\|\vec{b}\right\|}$$

Hence, we have shown that the relative error in the solution to a linear system ($\|\delta\vec{v}\| / \|\vec{v}\|$) can be bounded in terms of the relative error in our measurements for $\vec{b}$ ($\left\|\delta\vec{b}\right\| / \left\|\vec{b}\right\|$) as follows:

$$\frac{\|\delta\vec{v}\|}{\|\vec{v}\|} \leq (\|A\| \left\|A^{-1}\right\|) \cdot \frac{\left\|\delta\vec{b}\right\|}{\left\|\vec{b}\right\|}$$

In particular,

$$c = \|A\| \left\|A^{-1}\right\|$$

3. Now, instead of perturbing our measurement of the vector $\vec{b}$, we perturb the matrix A by some amount δA. In particular, we have the following pair of solutions to a linear system and a lightly perturbed one:

$$A\vec{v} = \vec{b}$$
$$(A + \delta A)(\vec{v} + \delta\vec{v}) = \vec{b}$$

Show that we can achieve a similar bound on the relative error of the solution to the perturbed linear system using the **same** condition number from the previous part:

$$\frac{\|\delta\vec{v}\|}{\|\vec{v} + \delta\vec{v}\|} \leq c \cdot \frac{\|\delta A\|}{\|A\|}$$

Solution: Expanding the second equation, we get:

$$A\vec{v} + A\delta\vec{v} + \delta A(\vec{v} + \delta\vec{v}) = \vec{b}$$

Subtracting the first equation from the equation above, we get:

$$\delta\vec{v} = -A^{-1}\delta A(\vec{v} + \delta\vec{v})$$

Applying the inequality on matrix-vector norms again, we have:

$$\|\delta\vec{v}\| \leq \left\|A^{-1}\right\| \|\delta A\| \|\vec{v} + \delta\vec{v}\|$$

Hence, we can rewrite it as:

$$\frac{\|\delta\vec{v}\|}{\|\vec{v} + \delta\vec{v}\|} \leq (\|A\| \left\|A^{-1}\right\|) \cdot \frac{\|\delta A\|}{\|A\|}$$