Math 43 Review Notes - Chapters 6 & 8.3

Symmetric Matrices and Orthogonal Diagonalization

A symmetric matrix has only real eigenvalues. Its eigenvectors can always be chosen to be orthonormal. A symmetric matrix can *always* be diagonalized, unlike other matrices. When we diagonalize a symmetric matrix we get a special diagonalization called an *orthogonal diagonalization.*

We diagonalize A as $A = Q\Lambda Q^T$, where Λ is the diagonal matrix with the eigenvalues of A on the diagonal, and Q is the matrix whose columns are *unit* eigenvectors.* Make sure that the unit eigenvectors in Q line up with the corresponding eigenvalues in Λ.

This diagonalization is very similar to the usual $S\Lambda S^{-1}$ diagonalization. The difference here is that we use *unit* eigenvectors, and since the columns of Q are orthonormal, $Q^{-1} = Q^T$, so we don't have to compute an inverse.

Example: Orthogonally diagonalize $\begin{pmatrix} 8 & 6 \\ 6 & -8 \end{pmatrix}$.

Solution: First find the eigenvalues.

$$\begin{vmatrix} 8-\lambda & 6 \\ 6 & -8-\lambda \end{vmatrix} = \lambda^2 - 100 = (\lambda - 10)(\lambda + 10).$$

Thus the eigenvalues are 10 and -10. Now find the eigenvectors and unit eigenvectors.

$$\lambda = 10 : \begin{pmatrix} -2 & 6 \\ 6 & -18 \end{pmatrix} \overset{rref}{\to} \begin{pmatrix} 1 & -3 \\ 0 & 0 \end{pmatrix} \Rightarrow \begin{bmatrix} 3 \\ 1 \end{bmatrix} \text{ is an eigenvector, and } \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix} \text{ is the unit eigenvector.}$$

$$\lambda = -10 : \begin{pmatrix} 18 & 6 \\ 6 & 2 \end{pmatrix} \overset{rref}{\to} \begin{pmatrix} 1 & 1/3 \\ 0 & 0 \end{pmatrix} \Rightarrow \begin{bmatrix} -1 \\ 3 \end{bmatrix} \text{ is an eigenvector, and } \begin{bmatrix} -1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix} \text{ is the unit eigenvector.}$$

Finally, we write

$$A = Q\Lambda Q^T = \begin{pmatrix} 3/\sqrt{10} & -1/\sqrt{10} \\ 1/\sqrt{10} & 3/\sqrt{10} \end{pmatrix} \begin{pmatrix} 10 & 0 \\ 0 & -10 \end{pmatrix} \begin{pmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ -1/\sqrt{10} & 3/\sqrt{10} \end{pmatrix}.$$

Markov Matrices

A *Markov matrix* is a matrix whose entries satisfy:

1. All the entries are ≥ 0.
2. The entries in each column add up to 1.

For example, $\begin{pmatrix} .2 & .5 \\ .8 & .5 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1/3 \\ 0 & 2/3 \end{pmatrix}$ are 2×2 Markov matrices.

A Markov matrix always has 1 as an eigenvalue. The eigenvectors corresponding to $\lambda = 1$ are called *steady-state eigenvectors.* To see why they are called this, recall the equation $A\mathbf{x} = \lambda\mathbf{x}$ that defines eigenvalues and eigenvectors. With $\lambda = 1$ this says $A\mathbf{x} = \mathbf{x}$; in other words multiplication by A doesn't change $\mathbf{x}$. Many real-life phenomena are modelled by the equation $\mathbf{x}_{n+1} = A\mathbf{x}_n$, with A a Markov matrix. The long range behavior of such a system is determined by the steady state eigenvectors.

Example: Find the eigenvalues and a steady-state eigenvector for $A = \begin{pmatrix} .3 & .4 \\ .7 & .6 \end{pmatrix}$.

Solution: We know right away that 1 is an eigenvalue for A, since A is a Markov matrix. To find the other, remember that the sum of the eigenvalues is equal to the sum of the diagonal entries of A. So we solve $1 + \lambda = .3 + .6$ to get the other eigenvalue $-.1$.

The steady-state eigenvectors are eigenvectors for $\lambda = 1$:

$$A - \lambda I = \begin{pmatrix} -.8 & .5 \\ .8 & -.5 \end{pmatrix} \overset{rref}{\to} \begin{pmatrix} 1 & -5/8 \\ 0 & 0 \end{pmatrix}.$$

Thus the steady state eigenvectors are multiples of $\begin{bmatrix} 5/8 \\ 1 \end{bmatrix}$. For instance, $\begin{bmatrix} 5 \\ 8 \end{bmatrix}$ or $\begin{bmatrix} 5/13 \\ 8/13 \end{bmatrix}$ are two examples.

*If one eigenvalue gives more than one linearly independent eigenvector, then you will have to orthogonalize the vectors using Gram-Schmidt or something else. We didn't consider anything like this in class, however, so don't worry about it.

More about Steady-State Eigenvectors

Markov matrices are useful for modelling many things. Below is a simple example.

Example: In any year, 92% of deer in the forest remain there, while 8% find their way into the suburbs (and people's backyards, where they eat their shrubbery). In addition, 88% of the deer in the suburbs remain there, while 12% are caught and returned into the forest. (Note that this is a simplified example; it doesn't take into account a lot of factors. Can you think of any assumptions in this model?)

Let f_n and s_n denote the number of deer in the forest and suburbs, respectively, in year n. Then we can write the above paragraph mathematically as

$$\begin{aligned} f_{n+1} &= \ .92f_n + .12s_n \\ s_{n+1} &= \ .08f_n + .88s_n \end{aligned}.$$

We can write this system of equations in matrix form as

$$\begin{bmatrix} f_{n+1} \\ s_{n+1} \end{bmatrix} = \begin{pmatrix} .92 & .12 \\ .08 & .88 \end{pmatrix} \begin{bmatrix} f_n \\ s_n \end{bmatrix}.$$

Let's call the matrix A. It is a Markov matrix, and thus it has $\lambda = 1$ as an eigenvalue. The steady-state eigenvectors (eigenvectors for $\lambda = 1$) are found by computing the nullspace of $A - \lambda I$.

$$A - \lambda I = \begin{pmatrix} -.08 & .12 \\ .08 & -.12 \end{pmatrix} \overset{rref}{\longrightarrow} \begin{pmatrix} 1 & -3/2 \\ 0 & 0 \end{pmatrix}.$$

Thus the steady state eigenvectors are multiples of $(3/2, 1)$. One multiple without fractions is $(3, 2)$. To express this in terms of percents, divide each entry by the sum of the two entries to get $(3/5, 2/5) = (.6, .4)$.

Thus we expect that after many years (i.e., when n is large) 60% of the total deer population will be in the forest and 40% will be in the suburbs. To see precisely why this is true, notice from the matrix equation above that

$$\begin{bmatrix} f_2 \\ s_2 \end{bmatrix} = A \begin{bmatrix} f_1 \\ s_1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} f_3 \\ s_3 \end{bmatrix} = A \begin{bmatrix} f_2 \\ s_2 \end{bmatrix} = A \left(A \begin{bmatrix} f_1 \\ s_1 \end{bmatrix} \right) = A^2 \begin{bmatrix} f_1 \\ s_1 \end{bmatrix}.$$

In general, we see that $\begin{bmatrix} f_{n+1} \\ s_{n+1} \end{bmatrix} = A^n \begin{bmatrix} f_1 \\ s_1 \end{bmatrix}$.

So the population is closely related to A^n. Recall that we can use diagonalization to find A^n. We diagonalize A as $A = S\Lambda S^{-1}$, and from there we get $A^n = S\Lambda^n S^{-1}$.

To diagonalize A we need the other eigenvalue and its eigenvector. As in the previous example, to find the other eigenvalue we solve $\lambda + 1 = .92 + .88$ to get $\lambda = .8$. An eigenvector for $\lambda = .8$ turns out to be $(-1, 1)$. Thus we can write

$$A^n = S\Lambda^n S^{-1} = \begin{pmatrix} 3 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & .8^n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -2 & 3 \end{pmatrix} \frac{1}{5} \quad = \quad \frac{1}{5} \begin{pmatrix} 3 + 2(.8^n) & 3 - 3(.8^n) \\ 2 - 2(.8^n) & 2 + 3(.8^n) \end{pmatrix}.$$

The last equality comes from multiplying the three matrices together. Notice that for fairly large values of n, $.8^n$ is very small. For example, $.8^{25} \approx .004$ and $.8^{50} \approx .000014$. So after a number of years we can ignore the $.8^n$ term and say

$$A \approx \begin{pmatrix} 3/5 & 3/5 \\ 2/5 & 2/5 \end{pmatrix}.$$

Notice that the columns are both the steady-state eigenvector of A that we found above. Moreover,

$$\begin{bmatrix} f_{n+1} \\ s_{n+1} \end{bmatrix} = A^n \begin{bmatrix} f_1 \\ s_1 \end{bmatrix} \approx \begin{pmatrix} 3/5 & 3/5 \\ 2/5 & 2/5 \end{pmatrix} \begin{bmatrix} f_1 \\ s_1 \end{bmatrix} = \begin{bmatrix} \frac{3}{5}f_1 + \frac{3}{5}s_1 \\ \frac{2}{5}f_1 + \frac{2}{5}s_1 \end{bmatrix} = (f_1 + s_1) \begin{bmatrix} 3/5 \\ 2/5 \end{bmatrix}.$$

So we see that in the long run (in this case after roughly 20-50 years) 3/5 (60%) of the deer population will be in the forest and 2/5 (40%) in the suburbs. In addition, the equation above says that the percentages will remain the same for all later years. This is why we call the vector $(3/5, 2/5)$ a *steady*-state.

Finally, notice that the percentage of deer in the forest versus the suburbs during year 1 had *absolutely no effect* on the outcome in the long run. Whether all the deer start out in the forest during year 1, or if it was 50/50, or whatever, makes no difference in the long run.

Positive Definite Matrices

A symmetric matrix with positive eigenvalues is called *positive definite*. The following gives a test for positive definiteness.

Let A be a symmetric matrix. If any one of the following is true, then the others are also true.

1. Every eigenvalue is positive.
2. Every upper left determinant is positive.
3. Every pivot is positive.
4. $\mathbf{x}^T A\mathbf{x}$ is positive except at $\mathbf{x} = \mathbf{0}$.

Remember, by positive we mean strictly greater than zero. Zero is *not* positive.

By "upper left determinant" we mean the determinant of the upper left part of the matrix. The four upper left determinants of this matrix are indicated on the right.

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 6 & 8 \\ 2 & 5 & 8 & 1 \\ 1 & 3 & 5 & 7 \end{pmatrix} \qquad 1, \quad \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = -2, \quad \begin{vmatrix} 1 & 2 & 3 \\ 3 & 4 & 6 \\ 2 & 5 & 8 \end{vmatrix} = -1, \quad \begin{vmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 6 & 8 \\ 2 & 5 & 8 & 1 \\ 1 & 3 & 5 & 7 \end{vmatrix} = -10$$

Example: Are all the eigenvalues of $\begin{pmatrix} 3 & 4 \\ 4 & 2 \end{pmatrix}$ positive?

Solution: The upper left determinants are 3 and $\begin{vmatrix} 3 & 4 \\ 4 & 2 \end{vmatrix} = -10$. They are not both positive, so neither are the eigenvalues.

Quadratic Forms

The term $\mathbf{x}^T A\mathbf{x}$ in (4) is called a *quadratic form*. Let's compute it for the 2×2 symmetric matrix $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$.

$$\mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x & y \end{bmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by & bx + cy \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bxy + cy^2.$$

Example: Using this formula, we can easily convert between a 2-dimensional quadratic form and its corresponding matrix. For example, the symmetric matrix $\begin{pmatrix} 3 & 2 \\ 2 & 7 \end{pmatrix}$ has quadratic form $3x^2 + 4xy + 7y^2$, and conversely, the quadratic form $5x^2 + 3xy + 2y^2$ corresponds to the matrix $\begin{pmatrix} 5 & 3/2 \\ 3/2 & 2 \end{pmatrix}$.

Completing the Square

You can check that we can write $ax^2 + 2bxy + cy^2 = a\left(x + \frac{b}{a}y\right)^2 + \left(\frac{ac - b^2}{a}\right)y^2$.

This is a lot to remember; however, notice that the coefficients of the square terms are actually *the pivots* of the matrix corresponding to this quadratic form. Breaking the quadratic form into these two square terms can be helpful sometimes.

Example: Are there any values of x and y that make $3x^2 + 12xy + 5y^2$ negative? If so, what are they?

Solution: The matrix corresponding to $3x^2 + 12xy + 5y^2$ is $\begin{pmatrix} 3 & 6 \\ 6 & 5 \end{pmatrix}$. Subtract $2\,row\ 1$ from $row\ 2$ to see what the pivots are. We get $\begin{pmatrix} \mathbf{3} & 6 \\ 0 & \mathbf{-7} \end{pmatrix}$. Now complete the square:

$$3x^2 + 12xy + 5y^2 = \mathbf{3}(x + 2y)^2 + \mathbf{-7}y^2.$$

The negative pivot indicates there are values of x and y that make this negative. To find them, choose x and y to make the first term 0, say $x = -2$, $y = 1$. With this x and y the quadratic form $3x^2 + 12xy + 5y^2$ evaluates to -7.

Finding the Axes of a Tilted Ellipse

This is a nice example of how quadratic forms and orthogonal diagonalization can help with an algebra problem. The ellipse on the left is an ordinary ellipse. Its equation is given by $ax^2 + by^2 = 1$. Its major and semi-major axes are the x and y axis, respectively.

However, we could have a "tilted" ellipse, one whose axes lie on diagonal lines, instead of the x and y axes. Its equation will have an additional term, an xy term. We want to be able to find its major and semi-major axes.

Example: Find the axes of the ellipse $11x^2 - 6xy + 19y^2 = 1$.

Solution: Notice that $11x^2 - 6xy + 19y^2$ is the quadratic form associated to $A = \begin{pmatrix} 11 & -3 \\ -3 & 19 \end{pmatrix}$.

Now orthogonally diagonalize A. Calculating the eigenvalues and eigenvectors, we find the eigenvalues are 20 and 10 with corresponding eigenvectors $(-1, 3)$ and $(3, 1)$ and corresponding unit eigenvectors $(-1, 3)/\sqrt{10}$ and $(3, 1)/\sqrt{10}$. Thus we can orthogonally diagonalize A as

$$A = Q\Lambda Q^T = \frac{1}{\sqrt{10}} \begin{pmatrix} -1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 20 & 0 \\ 0 & 10 \end{pmatrix} \begin{pmatrix} -1 & 3 \\ 3 & 1 \end{pmatrix} \frac{1}{\sqrt{10}}.$$

Now simplifying and substituting $Q\Lambda Q^T$ for A, we can write

$$11x^2 - 6xy + 19y^2 = \begin{bmatrix} x & y \end{bmatrix} A \begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{10} \begin{bmatrix} x & y \end{bmatrix} \begin{pmatrix} -1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 20 & 0 \\ 0 & 10 \end{pmatrix} \begin{pmatrix} -1 & 3 \\ 3 & 1 \end{pmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Simplify this a bit by multiplying the first two parts together and multiplying the last two together:

$$\begin{bmatrix} x & y \end{bmatrix} \begin{pmatrix} -1 & 3 \\ 3 & 1 \end{pmatrix} = \begin{bmatrix} -x + 3y & 3x + y \end{bmatrix}, \qquad \begin{pmatrix} -1 & 3 \\ 3 & 1 \end{pmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x + 3y \\ 3x + y \end{bmatrix}.$$

By doing this, the big messy term ends up looking just like a quadratic form:

$$\frac{1}{10} \begin{bmatrix} -x + 3y & 3x + y \end{bmatrix} \begin{pmatrix} 20 & 0 \\ 0 & 10 \end{pmatrix} \begin{bmatrix} -x + 3y \\ 3x + y \end{bmatrix} = \frac{1}{10} \begin{bmatrix} X & Y \end{bmatrix} \begin{pmatrix} 20 & 0 \\ 0 & 10 \end{pmatrix} \begin{bmatrix} X \\ Y \end{bmatrix},$$

where $X = -x + 3y$ and $Y = 3x + y$. Use the formula $ax^2 + 2bxy + cy^2$ on the previous page to write this as

$$\frac{1}{10}(20X^2 + (2)(0)XY + 10Y^2)$$

Plugging back in for X and Y and simplifying, we see that we have rewritten the original ellipse equation as

$$2(-x + 3y)^2 + (3x + y)^2 = 1.$$

The axes of the ellipse are given by the setting the terms being squared equal to 0. So the equations of the axes are

$$\begin{matrix} -x + 3y = 0 \\ 3x + y = 0 \end{matrix} \quad \text{or} \quad \begin{matrix} y = x/3 \\ y = -3x \end{matrix} .$$

Similar Matrices

We say B is *similar* to A if there is an M so that $B = M^{-1}AM$. For example, $\begin{pmatrix} 4 & 2 \\ -1 & 1 \end{pmatrix}$ is similar to $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ since $\begin{pmatrix} 4 & 2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix}$. This example comes from diagonalization.

Fact: Similar matrices have the same eigenvalues.

Example: $\begin{pmatrix} 2 & 5 \\ 0 & 3 \end{pmatrix}$ is not similar to $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ because they have different eigenvalues.

Example: Find an M to show that $B = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}$ is similar to $A = \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix}$.

Solution: We want to find an M so that $B = M^{-1}AM$. Multiply both sides of this equation by M to get $MB = AM$. Now let M be a generic 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and compute MB and AM:

$$MB = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} a & 3a + 2b \\ c & 3c + 2d \end{pmatrix}, \qquad AM = \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -a + 6c & -b + 6d \\ -a + 4c & -b + 4d \end{pmatrix}.$$

Since $MB = AM$, the corresponding entries in each must be equal, so we get the following system of equations:

$$\begin{array}{lcl} a &=& -a+6c \\ 3a+2b &=& -b+6d \\ c &=& -a+4c \\ 3c+2d &=& -b+4d \end{array} \quad \overset{\text{Simplify}}{\longrightarrow} \quad \begin{array}{l} a = 3c \\ a+b = 2d \\ a = 3c \\ 3c-b = 2d \end{array}.$$

This has an infinite number of solutions. Pick any value of c and d, say $c = d = 1$. Plugging into the equations, we get $a = 3$ and $b = -1$. Thus $M = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}$.

The only thing to be careful about when choosing c and d is to make sure that M is invertible. For instance $c = d = 0$ would be a bad choice. In general, it might not be this easy to solve the equations, but you can always move all the terms to the left hand side, write the system in matrix form and row reduce to find the solution. If there is no solution, or the only solution is all zeroes, then there is no M to be found, so the matrices are not similar. Note also that this same procedure (with a bigger M) works for larger matrices.

Singular Value Decomposition (SVD)

The singular value decomposition is a way to kind of diagonalize matrices of any shape. It uses the fact that $A^T A$ is a square symmetric matrix which can be orthogonally diagonalized.

Let A be an $m \times n$ matrix. The SVD is $A = U\Sigma V^T$, where

Σ — This is an $m \times n$ matrix whose diagonal entries are the square roots of the eigenvalues of $A^T A$. (These are called the *singular values* of A. All the other entries are 0.

V — This is an $n \times n$ matrix whose columns are the unit eigenvectors of $A^T A$.* Its columns are orthonormal.

U — This is an $m \times m$ matrix. For each nonzero eigenvector $\mathbf{v}$ of $A^T A$, compute $A\mathbf{v}$ and find the corresponding unit vector. These are the columns of U.† Its columns are also orthonormal.

Order matters. Arrange the singular values in Σ from largest to smallest so that the largest is the first entry, the next largest is the second entry, etc. The vectors in V and U must line up with their corresponding eigenvalues in Σ just like in diagonalization.

Helpful Hints:

(1) If A has more columns than rows, then AA^T will be a smaller matrix than $A^T A$ and thus its eigenvalues may be easier to find. In this case, any eigenvalue of AA^T is also an eigenvalue of $A^T A$, and the rest of the eigenvalues of $A^T A$ are 0.

(2) When finding a unit vector in the direction of a given vector, factor out anything you can from the vector and then ignore the number you factored out. For example, to find a unit vector in the same direction as $(4, 8)$, factor out a 4 to get $4(1, 2)$, ignore the 4 and just find the unit vector in the direction of $(1, 2)$, which is $(1, 2)/\sqrt{5}$. Since $(1, 2)$ and $(4, 8)$ both point in the same direction, the unit vector in the direction of either will be the same.

Example: Find the SVD of $\begin{pmatrix} 2 & 0 \\ 3 & 2 \end{pmatrix}$.

Solution: First compute $A^T A = \begin{pmatrix} 13 & 6 \\ 6 & 4 \end{pmatrix}$. Its eigenvalues are 16 and 1 with corresponding eigenvectors $(2, 1)$ and $(-1, 2)$. The corresponding unit eigenvectors are $(2, 1)/\sqrt{5}$ and $(-1, 2)/\sqrt{5}$. Thus

$$\Sigma = \begin{pmatrix} \sqrt{16} & 0 \\ 0 & \sqrt{1} \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}.$$

Now multiply each eigenvector by A and find the corresponding unit vectors to find the columns of U:

$$\begin{pmatrix} 2 & 0 \\ 3 & 2 \end{pmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix} \longrightarrow \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \qquad \begin{pmatrix} 2 & 0 \\ 3 & 2 \end{pmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \longrightarrow \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

Thus $A = U\Sigma V^T = \begin{pmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}$.

*Footnote on first page applies here, too.

†If $n < m$, then there will not be enough vectors to fill up U. More work is needed as the columns U must be extended to an orthonormal basis of $\mathbb{R}^m$. However, we didn't consider this in class, so don't worry about it.

Example: Find the SVD of $\begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$.

Solution: First compute $A^TA = \begin{pmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{pmatrix}$. Use helpful hint (1) to compute its eigenvalues.

$AA^T = \begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix}$ has eigenvalues 25 and 9, thus A^TA has eigenvalues 25, 9, and 0.

The corresponding eigenvectors are $(1,1,0)$, $(1,-1,4)$, and $(-2,2,1)$ with corresponding unit vectors $(1,1,0)/\sqrt{2}$, $(1,-1,4)/\sqrt{18}$, and $(-2,2,1)/3$. Thus

$$\Sigma = \begin{pmatrix} \sqrt{25} & 0 & 0 \\ 0 & \sqrt{9} & 0 \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{18} & -2/3 \\ 1/\sqrt{2} & -1/\sqrt{18} & 2/3 \\ 0 & 4/\sqrt{18} & 1/3 \end{pmatrix}.$$

Remember that Σ is the same shape as A, 3×2, so we can only fit the first two singular values in. Now multiply each eigenvector by A and find the corresponding unit vectors to find the columns of U:

$$\begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix} \longrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix} \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 9 \\ -9 \end{bmatrix} \longrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Notice that we did not use the third eigenvector, the one corresponding to the eigenvalue 0. This is because since A is 3×2, U is going to be 2×2 with vectors corresponding to the largest two eigenvalues.

So we have $A = U\Sigma V^T = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \\ -2/3 & 2/3 & 1/3 \end{pmatrix}$.

Orthonormal Bases for the Four Fundamental Spaces

The SVD gives orthonormal bases for the fundamental spaces. Let A be a matrix with rank r (i.e., A has r nonzero pivots).

(1) The first r columns of V are an orthonormal basis for the row space of A.

(2) The remaining columns of V are an orthonormal basis for the nullspace of A.

(3) The first r columns of U are an orthonormal basis for the column space of A.

(4) The remaining columns of U are an orthonormal basis for the nullspace of A^T (called the left nullspace).

Example: In the second example above A has rank 2. Thus the orthonormal bases are the following:

Row space — $\left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{18}} \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} \right\}$

Nullspace — $\left\{ \frac{1}{3} \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} \right\}$

Column space — $\left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$

Left nullspace — None: only $\mathbf{0}$ is in the left nullspace.