REVIEW FOR EXAM III

The exam covers sections 4.4, the portions of 4.6 on systems of differential equations and on Markov chains, and 5.1–5.4.

SIMILARITY AND DIAGONALIZATION

1. Two matrices A and B are similar if there is an invertible matrix P such that $P^{-1}AP = B$. A matrix D is diagonal if $d_{i,j} = 0$ whenever $i \neq j$. A matrix is diagonalizable if it is similar to a diagonal matrix. Note that every diagonal matrix is diagonalizable (just take $P = I$), but a diagonalizable matrix may or may not be diagonal.

2. Suppose A is diagonalizable, with $P^{-1}AP = D$. Let $d_1, \ldots, d_n$ be the diagonal entries of D, i.e. $d_j = d_{j,j}$. Let $\mathbf{p}_1, \ldots, \mathbf{p}_n$ be the columns of P. Rewriting our similarity equation as $AP = PD$ we have $[A\mathbf{p}_1|A\mathbf{p}_2|\cdots|A\mathbf{p}_n] = [d_1\mathbf{p}_1|d_2\mathbf{p}_2|\cdots|d_n\mathbf{p}_n]$. This shows that $A\mathbf{p}_j = d_j\mathbf{p}_j$, i.e. d_j is an eigenvalue of A and $\mathbf{p}_j$ is an eigenvector belonging to d_j. Note that these eigenvectors and eigenvalues occur in the same order in the matrices P and D. For example if

$$P = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

we have that $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$ are eigenvectors belonging to $\lambda_1 = 5$ and $\begin{bmatrix} 7 \\ 8 \\ 0 \end{bmatrix}$ is an eigenvector belonging to $\lambda_2 = 7$. We have $d_1 = 5$, $d_2 = 5$, and $d_3 = 7$. The eigenvalue $\lambda_1 = 5$ appears twice since exactly two of the columns are eigenvectors belonging to it. The eigenvalue $\lambda_2 = 7$ appears once since exactly one of the columns is an eigenvector belonging to it.

Since P is an invertible matrix its columns form a basis for $\mathbb{R}^n$. Thus we have the important fact that if A is diagonalizable, then there is a basis for $\mathbb{R}^n$ consisting of eigenvectors of A.

Each group of columns of P which are eigenvectors belonging to the same eigenvalue λ_i is a linearly independent subset of E_{λ_i} and so the number of elements in the group must be less than or equal to the dimension m_i of E_{λ_i}. But if we toss all the groups together we get n vectors (the columns of P), so $n \leq m_1 + \cdots m_k \leq n_1 + \cdots n_k = n$. The only way this can happen is for the number of vectors in each group to be exactly m_i and for m_i to equal n_i.

3. The argument above works in reverse. Let B_{λ_i} be the eigenbasis corresponding to λ_i. Toss all of these sets together to get a set B. It can be shown that B is linearly independent. If each $m_i = n_i$, then the number of elements of B is $m_1 + \cdots m_k = n_1 + \cdots n_k = n$, so B is a basis for $\mathbb{R}^n$ consisting of eigenvectors of A. Let P be the matrix whose columns are the elements of B. Then P is invertible. Let D be the diagonal matrix whose diagonal entries are the eigenvalues of A *repeated with the appropriate multiplicities and in the order corresponding to the columns of* P. (Look at the example in the previous section.) Then $AP = PD$, hence $P^{-1}AP = D$, and we have diagonalized A.

For example, if instead of being given P and D we are given that the eigenvalues of A are $\lambda_1 = 5$ and $\lambda_2 = 7$ and that $B_5 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \right\}$ and $B_7 = \left\{ \begin{bmatrix} 7 \\ 8 \\ 0 \end{bmatrix} \right\}$, then we would have

$$P = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

Note that in order to find D you *do not need to find* P^{-1}*!* Just follow the procedure above to construct P and D at the same time.

Note also that if some $m_i < n_i$, then A is *not* diagonalizable, so don't try to diagonalize it! You will not have enough eigenvectors in B to create P, so don't write down some silly matrix in an attempt to do so.

For example, if we had $\lambda_1 = 5$ and $\lambda_2 = 7$ and that $B_5 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}$ and $B_7 = \left\{ \begin{bmatrix} 7 \\ 8 \\ 0 \end{bmatrix} \right\}$ we would just say that A is not diagonalizable.

4. There are some shortcuts for showing that a matrix is diagonalizable.

   If A has n distinct eigenvalues, then since each $m_i \geq 1$ you are guaranteed of having enough vectors in B, so A is diagonalizable. (Of course the converse is false; if A does not have n distinct eigenvalues it may or may not be diagonalizable. You then have to do the hard work to find out which is true.)

5. Suppose A is diagonalizable, with $P^{-1}AP = D$. Then $A = PDP^{-1}$, from which it follows that for any positive integer k we have $A^k = PD^kP^{-1}$. Now D^k is easy to compute; it is just the diagonal matrix whose diagonal entries are the k^{th} powers of the diagonal entries of D. This can be used to give you a general formula for A^k.

   This in turn can sometimes be used to find the limit A^∞ of A^k as $k \to \infty$. (The limit does not always exist.) This can be applied to situations where you have some initial vector $\mathbf{x}_0$ and successive vectors $\mathbf{x}_{k+1} = A\mathbf{x}_k$ and want to find the limit $\mathbf{x}_\infty$ as $k \to \infty$; it is just $A^\infty \mathbf{x}_0$. The matrix A often arises in word problems where the next value of a certain vector is a linear function of its current value. Writing down A requires common sense and a careful reading of the problem. See the later section on Markov chains.

ORTHOGONALITY

1. Recall that the dot product of two column vectors $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ in $\mathbb{R}^n$ is the number $x_1y_1 + x_2y_2 + \cdots + x_ny_n$. Note that this is the same number you get by taking the

product of the $1 \times n$ matrix (row vector) $\mathbf{x}^T = [x_1, x_2, \ldots, x_n]$ and the $n \times 1$ matrix (column vector) $\mathbf{y}$ to get a 1×1 matrix:

$$\mathbf{x}^T\mathbf{y} = [x_1, x_2, \ldots, x_n] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = [x_1y_1 + x_2y_2 + \cdots + x_ny_n]$$

2. Two vectors $\mathbf{x}$ and $\mathbf{y}$ are orthogonal or perpendicular if $\mathbf{x} \cdot \mathbf{y} = 0$. Note that if these are column vectors in $\mathbb{R}^n$ then this is the same as saying that $\mathbf{x}^T\mathbf{y} = 0$.

ORTHOGONAL COMPLEMENTS

1. Given a subspace W of $\mathbb{R}^n$, the set $W^\perp$ of all vectors $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{x}$ is orthogonal to *every* vector in W is called the orthogonal complement of W. It is a subspace of $\mathbb{R}^n$. If the dimension of W is k, then the dimension of $W^\perp$ is $n-k$. $(W^\perp)^\perp = W$ and $W \cap W^\perp = \{\mathbf{0}\}$.

2. If $W = \text{span}(\mathbf{w}_1, \ldots, \mathbf{w}_k)$, then $W^\perp$ is the set of all $\mathbf{x}$ such that $\mathbf{w}_1 \cdot \mathbf{x} = 0, \ldots, \mathbf{w}_k \cdot \mathbf{x} = 0$. This is equivalent to saying that $\mathbf{w}_1^T\mathbf{x} = 0, \ldots, \mathbf{w}_k^T\mathbf{x} = 0$, and so $W^\perp$ is just the nullspace of the matrix A whose rows are the vectors $\mathbf{w}_1^T, \ldots, \mathbf{w}_k^T$. Thus the usual technique gives you a basis for $W^\perp$. If $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ is a basis for W, then the basis for $W^\perp$ will have $n-k$ elements $\{\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$, and the union of these two sets will be a basis for $\mathbb{R}^n$.

   Here is an example. Suppose $W = \text{span}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}\right)$. To find $W^\perp$ we must solve the equations $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$ and $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$. Thus we are finding the nullspace of the matrix $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$. It has REF $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. From this we get that $x_3 = 0$, $x_2 = t$, and $x_1 = -t$, so $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -t \\ t \\ 0 \end{bmatrix} = t\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$. Thus $W^\perp = \text{span}\left(\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right)$. We have a basis for $\mathbb{R}^3$ consisting of $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$, and $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$.

3. Suppose that we are given an $m \times n$ matrix A.

   Recall that the row space $\text{row}(A)$ is the subspace of $\mathbb{R}^n$ spanned by the rows of A. From the discussion above we see that $\text{row}(A)^\perp = \text{null}(A)$, where $\text{null}(A)$ is the nullspace of A.

   Recall that the column space $\text{col}(A)$ is the subspace of $\mathbb{R}^m$ spanned by the columns of A. Then $\text{col}(A)^\perp = \text{null}(A^T)$, the nullspace of A^T.

row(A), col(A), null(A), and null(A^T) are called the four fundamental subspaces associated to A.

ORTHOGONAL PROJECTIONS

1. Given a vector $\mathbf{v}$ in $\mathbb{R}^n$, there are unique vectors $\mathbf{w} \in W$ and $\mathbf{w}^\perp \in W^\perp$ such that $\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp$. To find them you first find bases for W and $W^\perp$ as in the previous paragraph. You solve the equation $r_1\mathbf{w}_1 + \cdots + r_k\mathbf{w}_k + r_{k+1}\mathbf{v}_{k+1} + \cdots + r_n\mathbf{v}_n = \mathbf{v}$ for $r_1, \ldots, r_n$ in the usual way. Then

$$\mathbf{w} = r_1\mathbf{w}_1 + \cdots + r_k\mathbf{w}_k, \quad \mathbf{w}^\perp = r_{k+1}\mathbf{v}_{k+1} + \cdots r_n\mathbf{v}_n$$

Here is an example. Let W be as in the previous example, and let $\mathbf{v} = \begin{bmatrix} 2 \\ 6 \\ 5 \end{bmatrix}$. We solve the equation

$$r_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + r_2 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + r_3 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \\ 5 \end{bmatrix}$$

to get $r_1 = 3$, $r_2 = 1$ and $r_3 = 2$. This gives us that $\mathbf{w} = 3\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 5 \end{bmatrix}$ and $\mathbf{w}^\perp = 2\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix}$.

$\mathbf{w}$ is called the orthogonal projection of $\mathbf{v}$ onto W. Note that if $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is an *orthogonal* basis for W (defined below), then there is a much easier procedure, described below, which you should use. If it is not orthogonal, then you have to fall back on the method just described.

ORTHOGONAL AND ORTHONORMAL SETS

1. A set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ of vectors is orthogonal if $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for all $i \neq j$. A set $\{\mathbf{q}_1, \ldots, \mathbf{q}_k\}$ of vectors is orthonormal if it is orthogonal and each vector has length 1, i.e. $||\mathbf{q}_i|| = 1$. Any orthogonal set of *non-zero* vectors can be transformed into an orthonormal set by dividing each vector by its length, i.e. $\mathbf{q}_i = \mathbf{v}_i/||\mathbf{v}_i||$.

2. Orthogonal sets of *non-zero* vectors are automatically linearly independent. So if you have an orthogonal spanning set of non-zero vectors for a subspace W of $\mathbb{R}^n$ it is automatically a basis for W. There is a bonus. It is very easy to find the coordinates of a vector with respect to an orthogonal basis. If $\mathbf{w} \in W$, then

$$\mathbf{w} = \left(\frac{\mathbf{w} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 + \cdots + \left(\frac{\mathbf{w} \cdot \mathbf{v}_k}{\mathbf{v}_k \cdot \mathbf{v}_k}\right)\mathbf{v}_k$$

For an orthonormal basis things are even simpler.

$$\mathbf{w} = (\mathbf{w} \cdot \mathbf{q_1})\mathbf{q_1} + \cdots + (\mathbf{w} \cdot \mathbf{q}_k)\mathbf{q}_k$$

3. If you have an orthogonal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ for W, then it is easy to compute the orthogonal projection $\mathbf{w}$ of a vector $\mathbf{v}$. Recall that $\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp$, where $\mathbf{w}^\perp$ is in $W^\perp$. Then

$$\mathbf{v} \cdot \mathbf{v}_i = (\mathbf{w} + \mathbf{w}^\perp) \cdot \mathbf{v}_i = \mathbf{w} \cdot \mathbf{v}_i + \mathbf{w}^\perp \cdot \mathbf{v}_i = \mathbf{w} \cdot \mathbf{v}_i + 0 = \mathbf{w} \cdot \mathbf{v}_i$$

So, substituting $\mathbf{v} \cdot \mathbf{v}_i$ for $\mathbf{w} \cdot \mathbf{v}_i$ into the formula above for $\mathbf{w}$ we get

$$\mathbf{w} = \left(\frac{\mathbf{v} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 + \cdots + \left(\frac{\mathbf{v} \cdot \mathbf{v}_k}{\mathbf{v}_k \cdot \mathbf{v}_k}\right)\mathbf{v}_k$$

Here is an example. Let $W = \text{span}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}\right)$ and $\mathbf{v} = \begin{bmatrix} 2 \\ 6 \\ 5 \end{bmatrix}$. Then

$$\mathbf{w} = \left(\frac{\begin{bmatrix} 2 \\ 6 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}\right)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \left(\frac{\begin{bmatrix} 2 \\ 6 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}}{\begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}}\right)\begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} =$$

$$\left(\frac{13}{3}\right)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \left(\frac{2}{6}\right)\begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 13/3 \\ 13/3 \\ 13/3 \end{bmatrix} + \begin{bmatrix} -1/3 \\ -1/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 12/3 \\ 12/3 \\ 15/3 \end{bmatrix} =$$

$$\begin{bmatrix} 4 \\ 4 \\ 5 \end{bmatrix}$$

The formula for $\mathbf{w}$ using an orthonormal basis for W is likewise obtained by replacing $\mathbf{w} \cdot \mathbf{q}_i$ by $\mathbf{v} \cdot \mathbf{q}_i$ in the formula from the previous section.

THE GRAM-SCHMIDT PROCESS

1. Suppose $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ is any basis for W. It can be transformed to an orthogonal basis for W by the following Gram-Schmidt process.

$\mathbf{v}_1 = \mathbf{x}_1$.

$\mathbf{p}_2$ is the orthogonal projection of $\mathbf{x}_2$ onto the subspace $W_1 = \text{span}(\mathbf{v}_1)$.

$\mathbf{v}_2 = \mathbf{x}_2 - \mathbf{p}_2$.

$\mathbf{p}_3$ is the orthogonal projection of $\mathbf{x}_3$ onto the subspace $W_2 = \text{span}(\mathbf{v}_1, \mathbf{v}_2)$.

$\mathbf{v}_3 = \mathbf{x}_3 - \mathbf{p}_3$.

Continue in this pattern until you have used up all the original vectors. The projections $\mathbf{p}_i$ are computed using the formula for $\mathbf{w}$ given earlier, applied to the vector $\mathbf{x}_i$ and the orthogonal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}\}$.

Here is an example. Let $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$, and $\mathbf{x}_3 = \begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix}$.

$$\mathbf{v}_1 = \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

$$\mathbf{p}_2 = \left(\frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right) \mathbf{v}_1 = \frac{4}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \mathbf{p}_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} - \frac{4}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/3 \\ -1/3 \\ 2/3 \end{bmatrix}.$$

$$\mathbf{p}_3 = \left(\frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right) \mathbf{v}_1 + \left(\frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right) \mathbf{v}_2 = \frac{7}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{2/3}{2/3} \begin{bmatrix} -1/3 \\ -1/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}.$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \mathbf{p}_3 = \begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

The arithmetic can be made easier by rescaling the $\mathbf{v}_i$; as soon as you create a $\mathbf{v}_i$, see whether or not it contains fractions. If it does, then replace it by a new $\mathbf{v}_i$ obtained by multiplying by some non-zero integer to cancel the denominators; then proceed to use the new $\mathbf{v}_i$ in further calculations.

In our example we would thus multiply the old $\mathbf{v}_2$ by 3 to obtain a new $\mathbf{v}_2 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$. Then we would have

$$\mathbf{p}_3 = \left(\frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right) \mathbf{v}_1 + \left(\frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right) \mathbf{v}_2 = \frac{7}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{2}{6} \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}.$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \mathbf{p}_3 = \begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

Do not rescale the $\mathbf{p}_i$! This would result in the $\mathbf{v}_i$ not being orthogonal.

2. Sometimes one wants to ultimately get an orthonormal basis $\{\mathbf{q}_1, \ldots, \mathbf{q}_k\}$. To do this you take the $\mathbf{v}_i$ above and divide by their lengths to get $\mathbf{q}_i = \mathbf{v}_i/||\mathbf{v}_i||$. Be sure to read the problem carefully to determine whether an orthogonal or orthonormal basis is required as the final answer.

ORTHOGONAL MATRICES

1. A matrix Q is orthogonal if the columns of Q form an orthonormal set. This is equivalent to $Q^{-1} = Q^T$. It is also equivalent to the set of rows of Q being an orthonormal set.

   Warning: In order to be an ortho*gonal* matrix it is not enough for the columns (respectively rows) to be an ortho*gonal* set; they must satisfy the stronger condition of being ortho*normal*. This is an unfortunate inconsistency in the terminology of linear algebra, but it has been firmly established for many decades, and we are stuck with it.

2. If Q is an orthogonal matrix then multiplication by it preserves dot products, lengths, and angles, i.e. $Q\mathbf{x} \cdot Q\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$, $||Q\mathbf{x}|| = ||\mathbf{x}||$, and the angle between $Q\mathbf{x}$ and $Q\mathbf{y}$ is equal to the angle between $\mathbf{x}$ and $\mathbf{y}$.

THE QR FACTORIZATION

1. Let A be an $n \times n$ matrix whose set of columns $\{\mathbf{a}_1, \ldots \mathbf{a}_n\}$ is linearly independent. Then there is an orthogonal matrix Q and an upper triangular matrix R such that $A = QR$. For example,

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{2} & 3/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{bmatrix} = QR$$

2. The way this arises is as follows: Gram-Schmidt the set $B = \{\mathbf{a}_1, \ldots \mathbf{a}_n\}$ to get an orthogonal set $B' = \{\mathbf{v}_1, \ldots \mathbf{v}_n\}$. Then normalize this new set to get an orthonormal set $B'' = \{\mathbf{q}_1, \ldots \mathbf{q}_n\}$. Each $\mathbf{q}_i$ is a multiple of $\mathbf{v}_i$, which in turn is a linear combination of the $\mathbf{a}_j$'s, so each $\mathbf{q}_i$ is a linear combination of the $\mathbf{a}_j$'s. This process can be reversed to express each $\mathbf{a}_i$ as a linear combination of the $\mathbf{q}_j$'s. In matrix form this gives $A = QR$ where the entries of R are the coefficients in these linear combinations.

3. How do you find Q? Use $\{\mathbf{q}_1, \ldots \mathbf{q}_n\}$ as the columns of Q, so $Q = [\mathbf{q}_1|\mathbf{q}_2|\cdots|\mathbf{q}_n]$.

4. How do you find R? Since you want $A = QR$, multiply both sides of this equation on the left by Q^{-1}. Since Q is orthogonal we have that $Q^{-1} = Q^T$, and so $R = Q^T A$.

5. So, if we were given $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$, we would first Gram-Schmidt the columns $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ of A to get the orthogonal set $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1/2 \\ 1/2 \end{bmatrix} \right\}$. To make the arithmetic easier it would probably be a good idea to rescale the second vector by multiplying by 2 to get the orthogonal set $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$. Each of these vectors has length $\sqrt{2}$, so dividing by the length gives

the orthonormal set $\left\{\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}\right\}$. We then use these as the columns of $Q = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$. We next transpose Q and multiply matrices to get

$$R = Q^T A = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 3/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{bmatrix}$$

6. Note that if you were given Q you do not need to do the Gram-Schmidt and normalization part of the process; just compute $R = Q^T A$ as above.

7. Finally, when you are asked for a factorization GIVE A FACTORIZATION! This means actually write out the two matrices side by side in the correct order. Your final answer should look like

$$\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{2} & 3/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{bmatrix}$$

ORTHOGONAL DIAGONALIZATION

1. A matrix A is orthogonally diagonalizable if there is an orthogonal matrix Q and a diagonal matrix D such that $Q^{-1}AQ = D$. The basic fact is that A is orthogonally diagonalizable if and only if A is symmetric. So, notice that while every orthogonally diagonalizable matrix is diagonalizable, not every diagonalizable matrix is orthogonally diagonalizable, since there are examples of diagonalizable matrices which are not symmetric.

2. If A is symmetric, hence orthogonally diagonalizable, then the eigenspaces E_{λ_i} and E_{λ_j} corresponding to distinct eigenvalues $\lambda_i \neq \lambda_j$ are perpendicular to each other, i.e. every vector in one eigenspace is perpendicular to every vector in the other eigenspace (i.e. their dot product is zero).

3. Here is how you orthogonally diagonalize a symmetric matrix A. Find the eigenvalues $\lambda_1, \dots, \lambda_k$ and the eigenbases $B_{\lambda_1}, \dots, B_{\lambda_k}$ in the usual way. Now you do something new. First use the Gram-Schmidt process to transform each basis B_{λ_i} to an orthogonal basis B'_{λ_i}, then divide each of the resulting vectors by its length to get an orthonormal basis B''_{λ_i}. Note that if the basis has just one element, there is nothing to do in the first of these steps, so you proceed to the second step. If there is more than one element in the basis, then you do have something to do in the first step; don't skip it and go to the second step. Also, note that you are doing this two step procedure on each of the eigenbases individually. Don't toss all of the eigenbases together and then apply the procedure; this would still give the right answer, but it would make the computation much longer and more complicated! So apply the two-step procedure to B_{λ_1} to get an orthonormal eigenbasis B''_{λ_1} for E_{λ_1}, then apply the procedure to B_{λ_2} to get an orthonormal eigenbasis B''_{λ_2} for E_{λ_2}, and so on until you have done this for all k eigenvalues. Then you create the matrix Q using these bases the same way as before; you toss all the B''_{λ_i} together and use these vectors as the n columns of Q. You create the diagonal matrix D the same way as before, putting the λ_i down the diagonal, having as many copies of λ_i as there are vectors in B''_{λ_i} and having them in the same order as the columns of Q.

4. Here is an example. Suppose the symmetric matrix A has eigenvalues $\lambda_1 = 3$ and $\lambda_2 = 6$ and that the corresponding eigenbases are $B_3 = \left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}\right\}$ and $B_6 = \left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right\}$.

We first apply the Gram-Schmidt procedure to B_3.

We start with $B_3 = \{\mathbf{x}_1, \mathbf{x}_2\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \right\}$.

$$\mathbf{v}_1 = \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

$$\mathbf{p}_2 = \left(\frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \right) \mathbf{v}_1 = \frac{4}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \mathbf{p}_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} - \frac{4}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix}.$$

We now have an orthogonal basis $B_3' = \{\mathbf{v}_1, \mathbf{v}_2\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -\frac{1}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix} \right\}$ for E_3.

We next divide each of these vectors by its length to get a unit vector.

$||\mathbf{v}_1|| = \sqrt{1+1+1} = \sqrt{3}$, so we get

$$\mathbf{q_1} = \frac{\mathbf{v}_1}{||\mathbf{v}_1||} = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}.$$

$||\mathbf{v}_2|| = \sqrt{\frac{1}{9} + \frac{1}{9} + \frac{4}{9}} = \sqrt{\frac{6}{9}} = \sqrt{\frac{2}{3}}$, so we get

$$\mathbf{q_2} = \frac{\mathbf{v}_2}{||\mathbf{v}_2||} = \sqrt{\frac{3}{2}} \begin{bmatrix} -\frac{1}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}.$$

We now have an orthonormal basis

$$B_3'' = \{\mathbf{q_1}, \mathbf{q_2}\} = \left\{ \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix} \right\}.$$

We make two observations at this point.

First, note that as soon as we obtained the vector $\mathbf{v}_2 = \begin{bmatrix} -\frac{1}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$, we could have rescaled it by multiplying by 3 to get the new $\mathbf{v}_2 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$, so that we would have $B_3' = \{\mathbf{v}_1, \mathbf{v}_2\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} \right\}$. We then have $||\mathbf{v}_1|| = \sqrt{3}$ and $||\mathbf{v}_2|| = \sqrt{1+1+4} = \sqrt{6}$, and so get

$$B_3'' = \{\mathbf{q}_1, \mathbf{q}_2\} = \left\{ \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix} \right\}.$$

Second, note that we did NOT go on to toss the vector $\mathbf{x}_3 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ into our Gram-Schmidt machine along with $\mathbf{x}_1$ and $\mathbf{x}_2$. Remember that we are processing ONE eigenbasis at a time, NOT all of them together. The vector $\mathbf{x}_3$ is not an element of our basis B_3 so we do not include it in our processing of B_3 to get B_3' and B_3''.

Now we move on to the eigenbasis $B_6 = \{\mathbf{x}_3\} = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$. Since there is only one vector here the Gram-Schmidt process reduces just to setting $\mathbf{v}_3 = \mathbf{x}_3 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, so we have $B_6' = B_6 = \{\mathbf{v}_3\} = \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$.

Next we compute $||\mathbf{v}_3|| = \sqrt{1+1} = \sqrt{2}$ and set $\mathbf{q}_3 = \dfrac{\mathbf{v}_3}{||\mathbf{v}_3||} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$. Thus we have the orthonormal basis $B_6'' = \{\mathbf{q}_3\} = \left\{ \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \right\}$.

Finally we assemble the matrices Q and D.

$$Q = [\mathbf{q}_1\, \mathbf{q}_2\, \mathbf{q}_3] = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \end{bmatrix} \quad D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

SYSTEMS OF LINEAR DIFFERENTIAL EQUATIONS

1. Suppose the variable x is a function of the variable t. We write this as $x = x(t)$ and let $x' = x'(t)$ be the derivative $\dfrac{dx}{dt}$ of x with respect to t.

2. The differential equation $x' = ax$ has general solution $x = ce^{at}$, where c is a constant.

3. Now suppose that we have several variables $x_1 \ldots, x_n$ which are functions of t. We can assemble these into a vector $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and let $\mathbf{x}' = \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix}$.

4. If we have several differential equations of the form $x_1' = d_1x_1, \ldots, x_n' = d_nx_n$, we can write them in the form $\mathbf{x}' = D\mathbf{x}$ where $D = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}$ is a diagonal matrix. Note that these equations have nothing to do with each other, and we have soutions $x_1 = c_1e^{d_1t}, \ldots, x_n = c_ne^{d_nt}$. This is called an uncoupled system.

5. Now suppose that we have a system of the form $\mathbf{x}' = A\mathbf{x}$, where A is not a diagonal matrix. This is called a coupled system. We wish to solve it by changing variables to get an uncoupled system. We can do this if A is diagonalizable.

We have that $P^{-1}AP = D$, where P is an invertible matrix and D is a diagonal matrix. We make a change of variables by setting $\mathbf{x} = P\mathbf{y}$, where $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$. It turns out that $\mathbf{x}' = P\mathbf{y}'$, and substituting these into $\mathbf{x}' = A\mathbf{x}$ gives $P\mathbf{y}' = AP\mathbf{y}$ and so $\mathbf{y}' = P^{-1}AP\mathbf{y} = D\mathbf{y}$, which is the desired uncoupled system.

6. For example, suppose we have the system $x_1' = -4x_1 + 3x_2$, $x_2' = -18x_1 + 11x_2$. This has matrix form

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} -4 & 3 \\ -18 & 11 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

The matrix $A = \begin{bmatrix} -4 & 3 \\ -18 & 11 \end{bmatrix}$ has eigenvalues 2 and 5 with associated eigenvectors $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$, respectively. Therefore A is diagonalizable with $D = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$ and $P = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}$. We set $\mathbf{x} = P\mathbf{y}$. This gives us the uncoupled system $y_1' = 2y_1$, $y_2' = 5y_2$, which has solution $y_1 = c_1e^{2t}$, $y_2 = c_2e^{5t}$. To get the solutions to the original system we use $\mathbf{x} = P\mathbf{y}$ to get

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} c_1e^{2t} \\ c_2e^{5t} \end{bmatrix} = \begin{bmatrix} (c_1e^{2t} + c_2e^{5t}) \\ (2c_1e^{2t} + 3c_2e^{5t}) \end{bmatrix}$$

Note that this can be written in the convenient form

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = c_1e^{2t} \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2e^{5t} \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

7. Sometimes one wants a solution which satisfies certain initial conditions, such as $x_1(0) = 1$ and $x_2(0) = 1$. Setting $t = 0$, $x_1 = 1$ and $x_2 = 1$ gives

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

This is just a matter of solving the system of linear equations with augmented matrix $\left[\begin{array}{cc|c} 1 & 1 & 1 \\ 2 & 3 & 1 \end{array}\right]$. This has RREF $\left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & -1 \end{array}\right]$, so we have $c_1 = 2$ and $c_2 = -1$, so

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2e^{2t} \begin{bmatrix} 1 \\ 2 \end{bmatrix} - e^{5t} \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

MARKOV CHAINS

1. A good way to understand Markov chains is to recall the example we used in class. See the lecture notes #32 and #33 for some of the details omitted below.

Suppose the population of the U.S. is constant, with no people leaving or entering the country. Start with x_0 being the population of California and y_0 being the population of the rest of the country (outside California). Suppose that each year 10% of the population outside CA moves into CA, while 20% of the population inside CA moves out.

Let x_1 and y_1 be the populations inside and outside CA one year later. Since 20% of the population of CA has moved out, we have that 80% remain. Thus one part of x_1 is $0.8x_0$. The other part consists of the 10% of the people outside who have moved in, i.e. $0.1y_0$. Thus the total is $x_1 = 0.8x_0 + 0.1y_0$.

Since 20% of the CA population has moved out, they give one part $0.2x_0$ of the population y_1 outside CA. The other part consists of the 90% of the population outside who have remained outside, giving $0.9y_0$. Thus the total population outside is $y_1 = 0.2x_0 + 0.9y_0$.

We can describe the situation by the matrix equation $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$. You should be able to translate a word problem like this into this matrix form.

Note that each column of this matrix sums to 1. Such a matrix is called a <u>stochastic</u> matrix.

Let $\mathbf{x}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$, $\mathbf{x}_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$, and $A = \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix}$. Then we can desribe this transition as $\mathbf{x}_1 = A\mathbf{x}_0$. Since we are assuming that the same pattern happens year after year we can describe the population two years after the start by $\mathbf{x}_2 = A\mathbf{x}_1 = AA\mathbf{x}_0 = A^2\mathbf{x}_0$. In general, after k years we have $\mathbf{x}_k = A^k\mathbf{x}_0$.

2. In general, computing a high power of a matrix is difficult, but there is one case in which it is easy. Suppose A is diagonalizable, with $P^{-1}AP = D$. Then $A = PDP^{-1}$, from which it follows that for any positive integer k we have $A^k = PD^kP^{-1}$. (We have the pairs $P^{-1}P$ cancelling out.) Now D^k is easy to compute; it is just the diagonal matrix whose diagonal entries are the k^{th} powers of the diagonal entries of D. This can be used to give you a general formula for A^k.

3. This in turn can be used to try to find the limit A^∞ of A^k as $k \to \infty$. One would have $A^\infty = PD^\infty P^{-1}$. Of course for this to work we need for the powers d_i^k which occur on the diagonal of D^k to have limits. If $|d_i| < 1$, then $d_i^k \to 0$, while if $d_i = 1$, then $d_i^k \to 1$.

4. We return to our example. If you do the usual eigenstuff for our A you get $p_A(\lambda) = \lambda^2 - 1.7\lambda + 0.7 = (\lambda - 1)(\lambda - 0.7)$, so the eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 0.7$. We find the eigenbases to be $B_1 = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ and $B_{0.7} = \left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$.

   This gives $P = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 1 & 0 \\ 0 & 0.7 \end{bmatrix}$. Thus as $k \to \infty$ we have $1^k \to 1$ and $0.7^k \to 0$. Thus $D^\infty = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

   We will have $A^\infty = PD^\infty P^{-1} = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/3 & 1/3 \\ -2/3 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1/3 & 1/3 \\ 0 & 0 \end{bmatrix} =$

$$\begin{bmatrix} 1/3 & 1/3 \\ 2/3 & 2/3 \end{bmatrix}$$

So in the long run we have the population vector $\mathbf{x}_\infty = A^\infty \mathbf{x}_0 = \begin{bmatrix} 1/3 & 1/3 \\ 2/3 & 2/3 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} (1/3)(x_0 + y_0) \\ (2/3)(x_0 + y_0) \end{bmatrix}$.

This says that in the long run one third of the total population is inside CA and the other two thirds is outside.