L14Eigenvectors
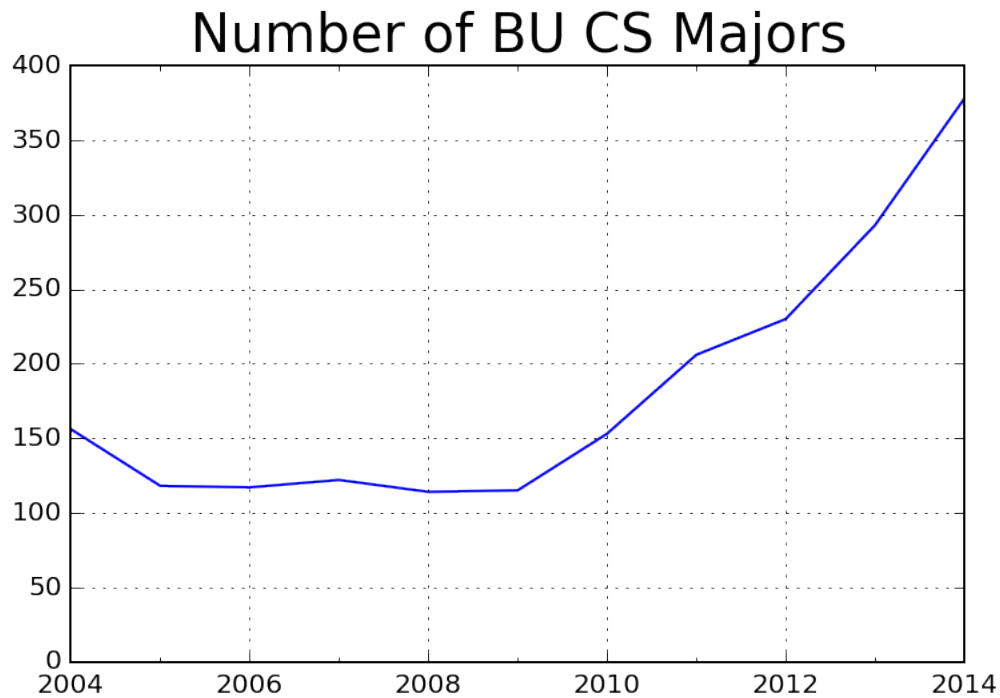
March 31, 2015

1 Eigenvectors and Eigenvalues

Currently the number of CS majors is growing very rapidly. We'd like to have some estimate of where the numbers might stabilize.

Modeling the number of students in the CS major is challenging for two reasons:

1. students enter and leave the major at various points during their undergraduate degree, and
2. there is some extrinsic "level of interest" in CS that influences the rates at which students enter and leave the major.

To address the first issue, we will model the population of CS majors in terms of cohorts – students in year 1, year 2, etc, of their education, and we will estimate how many students enter the program at each

stage. To address the second issue, we will use a proxy for "level of interest." Since there is a more or less fixed size Freshman class at BU, the number of students who are declared as CS majors as Freshmen is a reasonable indicator of the current "level of interest" in CS in that year.

We will use a simple linear model that we will train on historical data.

Our state variable is

$$\mathbf{x}_t = \begin{bmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \\ x_{4,t} \end{bmatrix}$$

where $x_{i,t}$ is the number of students who are declared CS majors and in their ith year at BU, in the fall of year t. Our model is a linear dynamical system (or linear recurrence):

$$\mathbf{x}_{t+1} = A\mathbf{x}_t.$$

Given historical data which are measurements of our state variable from past years, we can estimate A by a method called *least squares* (which we will cover later in the semester.)

To give you a sense of how the number of students of different classes varies depending on the situation, here are three examples, corresponding to shrinking, early growth, and steady growth:

```
In [4]: plt.subplot(1,3,1)
        cohorts.iloc[0].plot(kind='bar')
        plt.title(cohorts.index[0].strftime('%Y'))
        plt.subplot(1,3,2)
        cohorts.iloc[6].plot(kind='bar')
        plt.title(cohorts.index[6].strftime('%Y'))
        plt.subplot(1,3,3)
        cohorts.iloc[10].plot(kind='bar')
        plt.title(cohorts.index[10].strftime('%Y'))
        plt.subplots_adjust(wspace=0.7)
```
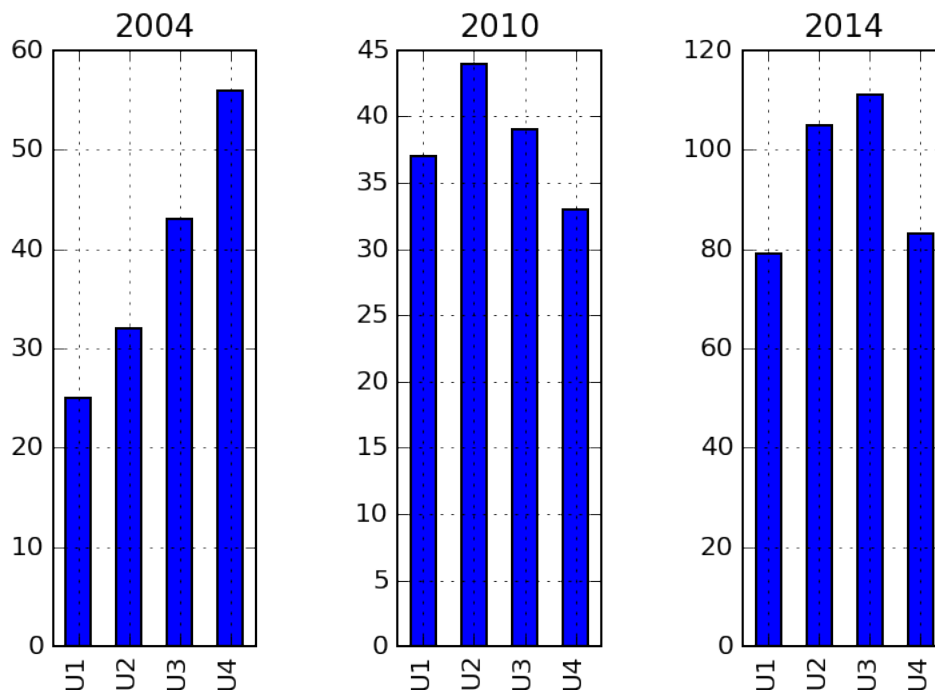
So, given enough information, in the form of this state vector, we can in fact make accurate predictions of number of CS majors:

```
In [5]: OneYearAhead = cohorts.dot(A.T).shift(1,freq=Year)
        OneYearAhead.columns=pd.Index(['U1','U2','U3','U4'])
        OneYearAhead.sum(axis=1).plot()
        cohorts.sum(axis=1).plot()
        plt.ylim(ymin=0)
        plt.legend(['Model','Actual'],loc='best')

Out[5]: <matplotlib.legend.Legend at 0x10e3a4d90>
```
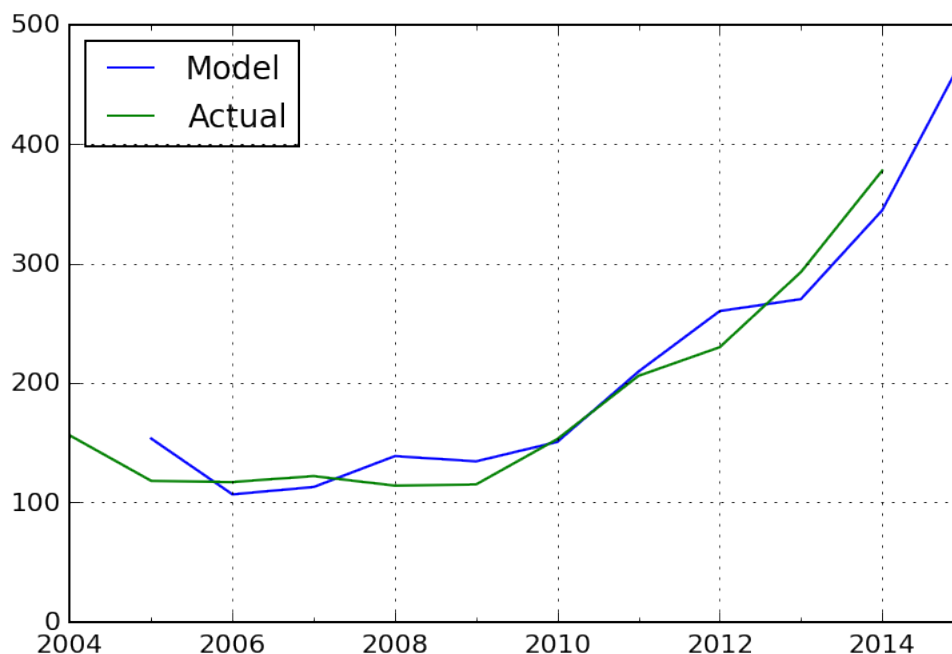

Besides predicting next year's class sizes, analyzing this model can help us answer other important questions:

1. If CS major enrollments were at steady-state, how many students will be Freshmen, Sophomores, etc? (This impacts course offerings).
2. If the CS major were at steady-state, how many students would be graduating each year?
3. What is the intrinsic growth rate predicted by this model?

Today's lecture will start to develop the tools to answer such questions (and more).

## 1.1 Eigenvectors and Eigenvalues

The system above is an example of a linear transformation $\mathbf{x} \mapsto A\mathbf{x}$.

As we've seen, such transformations, thinking geometrically, can move a vector in a variety of directions.

The starting point for understanding such systems is to look for special vectors which **do not change their direction** when multiplied by A.

Example.

Let $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. The images of $\mathbf{u}$ and $\mathbf{v}$ under multiplication by A are shown here:
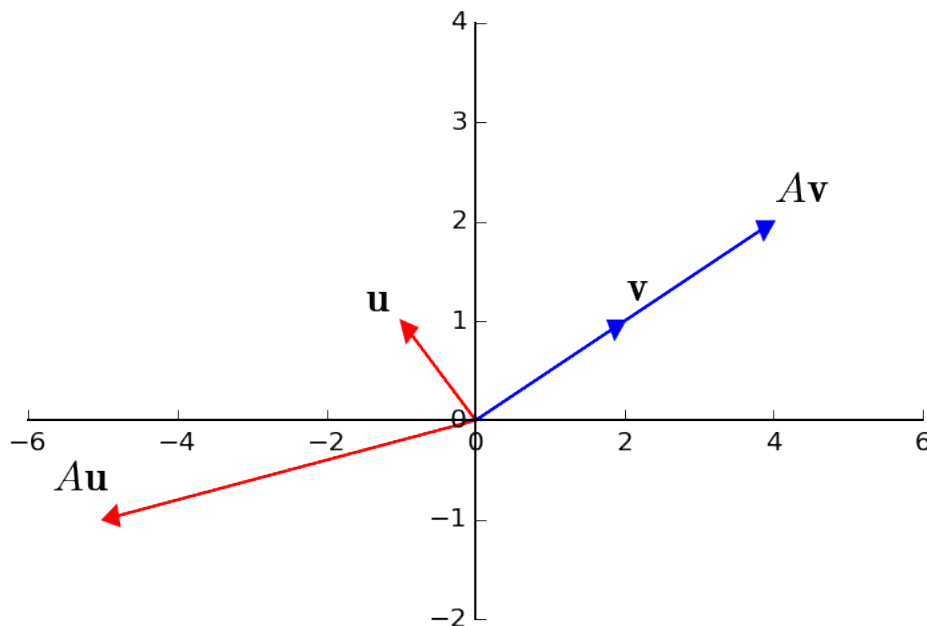
For a "typical" vector (like $\mathbf{u}$) the action of A is to scale and rotate it.
But for some special vectors (like $\mathbf{v}$) the action of A is only to "stretch" it (without any rotation).
In fact, $A\mathbf{v}$ is simply $2\mathbf{v}$.
Today we will by studying equations such as

$$A\mathbf{x} = 2\mathbf{x} \quad \text{or} \quad A\mathbf{x} = -4\mathbf{x}$$

where special vectors are transformed by A into scalar multiples of themselves.

Definition. An **eigenvector** of an $n \times n$ matrix A is a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ. A scalar λ is called an **eigenvalue** of A if there is a nontrivial solution $\mathbf{x}$ of $A\mathbf{x} = \lambda\mathbf{x}$; such an $\mathbf{x}$ is called an *eigenvector corresponding to λ*.

1.2 Checking Eigenvectors and Eigenvalues

Before considering how to compute eigenvectors, let's look at simply how to determine if a given vector is an eigenvector of a matrix.

Example. Let $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

Are $\mathbf{u}$ and $\mathbf{v}$ eigenvectors of A?

Solution.

$$A\mathbf{u} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} -24 \\ 20 \end{bmatrix} = -4 \begin{bmatrix} 6 \\ -5 \end{bmatrix} = -4\mathbf{u}.$$

$$A\mathbf{v} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 11 \end{bmatrix} \neq \lambda \begin{bmatrix} 3 \\ -2 \end{bmatrix}.$$

4

So **u** is an eigenvector corresponding to the eigenvalue -4, but **v** is not an eigenvector of A.

Example. Show that 7 is an eigenvalue of the matrix A, and find the corresponding eigenvectors.

For 7 to be an eigenvalue of A, it must be the case that

$$A\mathbf{x} = 7\mathbf{x}$$

has a nontrivial solution.

This is equivalent to:

$$A\mathbf{x} - 7\mathbf{x} = \mathbf{0}$$

or

$$(A - 7I)\mathbf{x} = \mathbf{0}.$$

To solve this homogeneous equation, form the matrix

$$A - 7I = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}.$$

The columns of $A - 7I$ are obviously linearly dependent, so $(A - 7I)\mathbf{x} = \mathbf{0}$ has a nontrivial solution.

So 7 is an eigenvalue of A.

To find the corresponding eigenvectors, we solve $(A - 7I)\mathbf{x} = \mathbf{0}$ using row operations:

$$\begin{bmatrix} -6 & 6 & 0 \\ 5 & -5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The general solution has the form $x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Each vector of this form with $x_2 \neq 0$ is an eigenvector corresponding to $\lambda = 7$.

1.3 Eigenspace

Clearly, λ is an eigenvalue of an $n \times n$ matrix A if and only if the equation

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

has a nontrivial solution.

Notice that the set of *all* solutions of that equation is just *the null space of the matrix $A - \lambda I$.*

So the set of all eigenvectors corresponding to a particular λ is a *subspace of* $\mathbb{R}^n$ and is called the **eigenspace** of A corresponding to λ.

Example. For matrix $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, we found that the general solution for the eigenvector corrsponding to $\lambda = 7$ is the expression $\mathbf{x}_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. This means that the eigenspace corresponding to $\lambda = 7$ consists of all multiples of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, which is the line through $(1, 1)$ and the origin.

We could also show that the eigenspace corresponding to $\lambda = -4$ is the line through $(-6, 5)$.
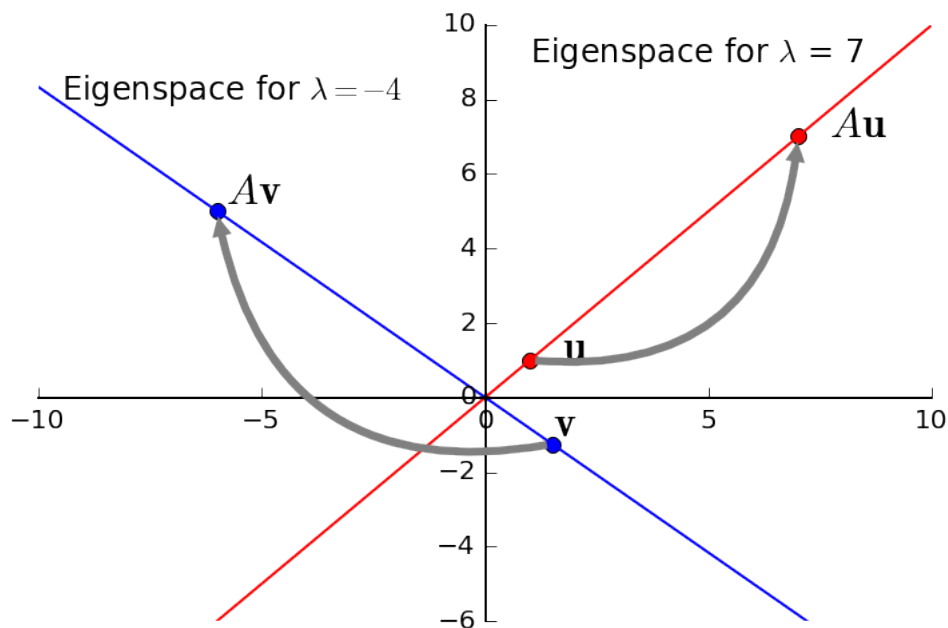
Example. Let $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. An eigenvalue of A is 2. Find a basis for the the corresponding eigenspace.

Solution. The eigenspace we are looking for is the nullspace of $A - 2I$.

We will use the strategy for finding a basis for a nullspace that we learned in the last lecture.

Form

$$A - 2I = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix}$$

and row reduce the augmented matrix for $(A - 2I)\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

At this point, it is clear that 2 is indeed an eigenvalue of A because the equation $(A - 2I)\mathbf{x} = \mathbf{0}$ has free variables.

The general solution is $x_1 = \frac{1}{2}x_2 + 3x_3$.
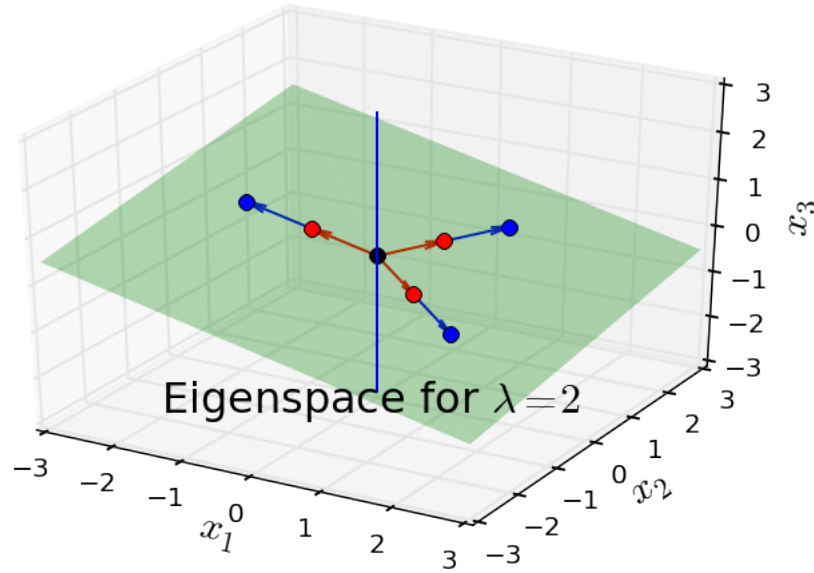
Expressed as a vector, the general solution is:

$$\begin{bmatrix} \frac{1}{2}x_2 + 3x_3 \\ x_2 \\ x_3 \end{bmatrix}.$$

Expressed as a vector sum, this is:

$$x_2 \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \quad \text{with } x_2 \text{ and } x_3 \text{ free.}$$

So this eigenspace is a two-dimensional subspace of $\mathbb{R}^3$, and a basis for this subspace is

$$\left\{ \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Eigenspace for $\lambda=2$

1.4   Eigenvalues of a Triangular Matrix

Theorem. The eigenvalues of a triangular matrix are the entries on its main diagonal.

   Proof. We'll consider the 3×3 case. If A is upper triangular, then $A - \lambda I$ has the form

$$A - \lambda I = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ 0 & a_{22} - \lambda & a_{23} \\ 0 & 0 & a_{33} - \lambda \end{bmatrix}$$

   Now, λ is an eigenvalue of A if and only if the equation $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has a nontrivial solution – in other words, if an only if the equation has a free variable.

   The special pattern that the zero entries have in A means that there will be a free variable if any diagonal entry is also zero, because then there will be a column without a pivot.

   That is, $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has a free variable if and only if at least one of the entries on the diagonal of $(A - \lambda I)$ is zero.

   This happens if and only if λ equals one of the entries a_{11}, a_{22}, or a_{33} on the diagonal of A.

Example. Let $A = \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & 3 & 4 \end{bmatrix}$.

Then the eigenvalues of A are 3, 0, and 2. The eigenvalues of B are 4 and 1.

1.5 Invertibility and Eigenvalues

So far we haven't probed what it means for a matrix to have an eigenvalue of 0.

This happens if and only if the equation $A\mathbf{x} = 0\mathbf{x}$ has a nontrivial solution.

But that equation is equivalent to $A\mathbf{x} = \mathbf{0}$ which has a nontrivial solution if and only if A is not invertible.

This means that *0 is an eigenvalue of A if and only if A is not invertible.*

This draws an important connection between invertibility and zero eigenvalues.

So we have **yet another** addition to the Invertible Matrix Theorem!

1.6 Eigenvectors and Difference Equations

Let's return to the problem we considered at the outset: predicting future values of $\mathbf{x}_t$ (the number of CS majors of each class in year t).

We modeled the future as $A\mathbf{x}_{t+1} = A\mathbf{x}_t \quad (t = 0, 1, 2, \dots)$.

A **solution** of such an equation is an explicit description of $\{\mathbf{x}_t\}$ whose formula for each $\mathbf{x}_t$ does not depend directly on A or on the preceding terms in the sequence other than the initial term $\mathbf{x}_0$.

The simplest way to build a solution is to take an eigenvector $\mathbf{x}_0$ and its corresponding eigenvalue λ and let

$$\mathbf{x}_t = \lambda^t \mathbf{x}_0 \quad (t = 0, 1, 2, \dots)$$

This sequence is a solution because

$$A\mathbf{x}_t = A(\lambda^t \mathbf{x}_0) = \lambda^t(A\mathbf{x}_0) = \lambda^k(\lambda \mathbf{x}_0) = \lambda^{t+1}\mathbf{x}_0 = \mathbf{x}_{t+1}.$$

We can extend this to combinations of eigenvectors.

If A has two eigenvectors $\mathbf{u}$ and $\mathbf{v}$ corresponding to eigenvalues λ and μ, then another solution is

$$x_t = \lambda^t \mathbf{u} + \mu^t \mathbf{v}$$

To address the problem of predicting the CS major population, let's assume that we can express the state vector $\mathbf{x}$ as a linear combination of eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p$. That is,

$$\mathbf{x}_0 = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_p\mathbf{u}_p.$$

Then at time t,

$$\mathbf{x}_t = \lambda_1^t a_1 \mathbf{u}_1 + \lambda_2^t a_2 \mathbf{u}_2 + \cdots + \lambda_p^t a_p \mathbf{u}_p.$$

Which gives us an explicit solution for the number of students in the major in any future year.

Note that the expression above is exponential in each of the $\lambda_1, \dots, \lambda_p$.

The resulting value will be essentially determined by the largest λ_i.

To see this, let's say that λ_1 is the largest eigenvalue, and λ_2 is the second largest.

Then the proportion of $\mathbf{u}_1$ to $\mathbf{u}_2$ contained in $\mathbf{x}_t$ will be given by $(\lambda_1/\lambda_2)^t$.

Since this ratio is larger than 1, the relative proportion of $\mathbf{u}_1$ will grow exponentially.

This shows the importance of the largest eigenvalue in the value of $\mathbf{x}_t$ for large t. In the long run, $\mathbf{x}_t$ will be very close to $\lambda_1^t \mathbf{u}_1$.

The largest eigenvalue shows the asymptotic rate of growth (or shrinkage) of the state vector $\mathbf{x}_t$.

For the matrix A that describes CS major opoulation growth, the largest eigenvalue is 1.19.

In fact, the incoming class (fall 2015) shows an increase in stated interest in CS of about 20%!