MODULE 9: DIAGONALIZATION AND PAGE RANK

Linear Algebra 3: Determinants and Eigenvalues

TOPIC 1: Diagonalization

Powers of Matrices

Motivation: it can be useful to take large powers of matrices, for example A^k, for large k.
But: multiplying 2 $n \times n$ matrices requires roughly n^3 computations. Is there a more efficient way to compute A^k?

Diagonal Matrices

DEFINITION — Diagonal Matrices

↳ A matrix is diagonal if the only nonzero elements, if any, are on the main diagonal.

The following are all diagonal matrices.

$$I_n, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

We will only be working with diagonal square matrices in the course.

Powers of Diagonal Matrices

If A is diagonal, then A^k is easy to compute. For example,

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 5 \end{pmatrix}$$

$$A^2 = \begin{pmatrix} 3 & 0 \\ 0 & 5 \end{pmatrix}\begin{pmatrix} 3 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 3^2 & 0 \\ 0 & 5^2 \end{pmatrix}$$

$$A^k = \begin{pmatrix} 3^k & 0 \\ 0 & 5^k \end{pmatrix}$$

But what if A is not diagonal?

Diagonalization

DEFINITION — Diagonalization

↳ Suppose $A \in \mathbb{R}^{n \times n}$. We say that A is diagonalizable if it is similar to a diagonal matrix, D. That is, we can write $A = PDP^{-1}$.

Also note that $A = PDP^{-1}$ iff: $A = (\vec{v}_1, \vec{v}_2 \cdots \vec{v}_n)\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}(\vec{v}_1 \; \vec{v}_2 \cdots \vec{v}_n)^{-1}$

$\vec{v}_1, \cdots, \vec{v}_n$ are linearly independent eigenvectors, and $\lambda_1, \cdots, \lambda_n$ are the corresponding eigenvalues (in order).

Proof: We construct $P = (\vec{v}_1 \ \vec{v}_2 \ \cdots \ \vec{v}_n)$. Then

$$\begin{aligned} AP &= A(\vec{v}_1 \ \vec{v}_2 \ \cdots \ \vec{v}_n) \\ &= (A\vec{v}_1 \ A\vec{v}_2 \ \cdots \ A\vec{v}_n) \\ &= (\lambda_1\vec{v}_1 \ \lambda_2\vec{v}_2 \ \cdots \ \lambda_n\vec{v}_n) \end{aligned}$$

$$AP = (\vec{v}_1 \ \vec{v}_2 \ \cdots \ \vec{v}_n)\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

$$= PD$$

Or, $A = PDP^{-1}$. (QED)

Theorem: If A is diagonalizable $\Leftrightarrow$ A has n linearly independent vectors.

Note: $\Leftrightarrow$ means if and only if (iff)

Example 1:

Diagonalize if possible. $\begin{pmatrix} 2 & 6 \\ 0 & -1 \end{pmatrix} = A$.

$\lambda_1 = 2: (A - 2I|0) = \left(\begin{array}{cc|c} 0 & 6 & 0 \\ 0 & -3 & 0 \end{array}\right)$ $\quad 6x_2 = 0 \Rightarrow x_2 = 0. \Rightarrow \vec{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$\lambda_2 = -1: (A + I|0) = \left(\begin{array}{cc|c} 3 & 6 & 0 \\ 0 & 0 & 0 \end{array}\right)$ $\quad 3x_1 + 6x_2 = 0 \Rightarrow \vec{v}_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$

$$\begin{aligned} A &= PDP^{-1} \\ &= \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} \end{aligned}$$

$$P^{-1} = \frac{1}{-1}\begin{pmatrix} -1 & -2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}.$$

Example 2:

Diagonalize if possible. $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = A$.

$\lambda = 1: (A - I|0) = \left(\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right)$ $\quad x_2 = 0. \Rightarrow \vec{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$

$\Rightarrow$ A is not diagonalizable.

Exercise:
Suppose matrix A is 2x2 and has the eigenvectors and eigenvalues below.

$$\lambda_1 = -1, \quad \vec{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}; \quad \lambda_2 = 0, \quad \vec{v}_2 = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$$

If $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, what is a_{11} equal to?

$$P = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}, \quad P^{-1} = \frac{1}{2(3)-5(1)} \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$$

$$D = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

$$A = PDP^{-1}$$

$$= \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$$

$$A = \begin{pmatrix} -6 & 10 \\ -3 & 5 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

$$\therefore a_{11} = -6$$

Diagonalization Theorems

Motivation:
* how can we determine whether a given nxn matrix can be diagonalized?
* can we determined whether a square matrix can be diagonalized if we know:
  - ↳ the algebraic or geometric multiplicities of the eigenvalues? (we can)
  - ↳ whether the matrix is invertible? (we cannot)
* if a matrix has a repeated eigenvalue, how can we diagonalize the matrix?

Theorem: Distinct Eigenvalues and Diagonalizability
If A is nxn and has n distinct eigenvalues, then A is diagonalizable

Why does this theorem hold?
↳ For an nxn matrix to be diagonalizable it must have n linearly independent eigenvectors.
↳ Eigenvectors corresponding to distinct eigenvalues are independent.

Is it necessary for an nxn matrix to have n distinct eigenvalues for it to be diagonalizable?
↳ NO. The identitity matrix is diagonalizable.

Diagonalization Example 1:
Give an example of a nonzero square matrix that is in RREF, is diagonalizable, and is singular.

Solution: Any matrix that has distinct eigenvalues can be diagonalized.
This matrix below can be diagonalized:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

This matrix can be diagonalized:

$$A = PDP^{-1}, \quad P = P^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Note: if we know that a matrix is not invertible, we cannot conclude that the matrix is not diagonalizable

How can we tell whether a matrix with repeated eigenvalues is diagonalizable?

Theorem: Diagonalizability

Suppose → A is any $n \times n$ real matrix
↳ A has distinct eigenvalues $\lambda_1, \dots, \lambda_k$, $k \leq n$
↳ a_i = algebraic multiplicity of λ_i
↳ g_i = dimension of λ_i eigenspace, or the geometric multiplicity

Then
↳ A is diagonalizable $\Leftrightarrow \sum g_i = n \Leftrightarrow g_i = a_i \ \forall i$
↳ A is diagonalizable $\Leftrightarrow$ the eigenvectors, $\forall$ eigenvalues, together form a basis for $\mathbb{R}^n$

Diagonalization Example 2:

The following are equivalent:
* $n \times n$ matrix A is diagonalizable
* $\sum g_i = n$
* $g_i = a_i \ \forall i$
* the eigenvectors, $\forall$ eigenvalues, together form a basis for $\mathbb{R}^n$

True or False: If A is not invertible, then A is not diagonalizable.
False. If $A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, then A is not invertible and can be diagonalized:

$$A = PDP^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Note → some matrices that are not invertible can be diagonalized
↳ some matrices that have a repeated eigenvalue can be diagonalized

Diagonalization Example 3:
For what value of k is $A = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ diagonalizable?

Case 1: $k = 0$. Then $A = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and can be diagonalized:

$$A = PDP^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Case 2: $k \neq 0$. Then $A = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$, and $\lambda = 1$. Obtain eigenvectors:

$$(A - I \mid 0) = \left(\begin{array}{cc|c} 0 & k & 0 \\ 0 & 0 & 0 \end{array}\right) \Rightarrow \vec{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

A can only be diagonalized when $k = 0$.

matrix invertibility does not tell if matrix is diagonalizable, have to look at multiplicities of the eigenvalues or whether the eigenvectors span all of our n

Note: Matrix A is invertible $\forall$ values of k.
Matrix A is diagonalizable $\exists$ values of k.
The invertibility of a matrix does not tell us anything about whether the matrix is diagonalizable.

<u>Diagonalization a Matrix with Repeated Eigenvalues</u>

How can we diagonalize a matrix that has a repeated eigenvalue?

The only eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = \lambda_3 = 3$. If possible, construct P and D such that $AP = PD$.

$$A = \begin{pmatrix} 7 & 4 & 16 \\ 2 & 5 & 8 \\ -2 & -2 & -5 \end{pmatrix}$$

$\lambda_1 = 1$. Identifying corresponding eigenvectors:

$$A - \lambda_1 I = A - I = \begin{pmatrix} 6 & 4 & 16 \\ 2 & 4 & 8 \\ -2 & -2 & -6 \end{pmatrix} \xrightarrow[-\frac{1}{2}R_3 \to R_3]{\frac{1}{2}R_1 \to R_1,\ \frac{1}{2}R_2 \to R_2} \begin{pmatrix} 3 & 2 & 8 \\ 1 & 2 & 4 \\ 1 & 1 & 3 \end{pmatrix} \xrightarrow{R_2 - R_3 \to R_2} \begin{pmatrix} 3 & 2 & 8 \\ 0 & 1 & 1 \\ 1 & 1 & 3 \end{pmatrix} \xrightarrow{R_3 - R_2 \to R_3} \begin{pmatrix} 3 & 2 & 8 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix}$$

$$\xrightarrow{R_1 - 2R_2 - 3R_3 \to R_1} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$x_2 + x_3 = 0 \Rightarrow x_2 = -x_3$

$x_1 = -2x_3 = 2x_2$

A vector in the null space of $A - \lambda_2 I$ is $\vec{v}_1 = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$.

$\lambda_2 = 3$. $A - \lambda_2 I = A - 3I$

$$= \begin{pmatrix} 4 & 4 & 16 \\ 2 & 2 & 8 \\ -2 & -2 & -8 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$x_1 + x_2 + 4x_3 = 0$

Eigenvectors corresponding to $\lambda_2 = 3$ must satisfy this relation.

x_2 and x_3 are free variables.

$\Rightarrow x_1 = -x_2 - 4x_3$

Parametric vector form:

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_2 - 4x_3 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -4 \\ 0 \\ 1 \end{pmatrix}$$

$$\therefore P = (\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3) = \begin{pmatrix} 2 & -1 & -4 \\ 1 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Exercise: Are the following matrices diagonalizable?

$A = \begin{pmatrix} 7 & 1 \\ 0 & 5 \end{pmatrix}$ ✓. The eigenvalues of A can be determined by inspection. $7 \neq 5$. [the eigenvalues are distinct,] and [eigenvectors corresponding to distinct eigenvalues are linearly independent.] So we can form an invertible matrix P with the eigenvectors of A. So, A is diagonalizable.

$\begin{pmatrix} 7 \\ 0 \end{pmatrix}$ & $\begin{pmatrix} 1 \\ 5 \end{pmatrix}$

$A = \begin{pmatrix} 7 & 1 \\ 0 & 7 \end{pmatrix}$ ✗. The eigenvalues of A can be determined by inspection, the eigenvalues are not distinct. An eigenvector we can find for this matrix is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, but we cannot construct another eigenvector that will allow us to create an invertible matrix P. because the dimension of the eigenspace is only 1. So we cannot form an invertible matrix P. So A is not diagonalizable.

Matrix Powers

Motivation: Suppose A is an $n \times n$ matrix. Recall that:

↳ In some applications we need to compute A^k for large k

↳ computing A^k directly could require many computations, especially if n is large and many of the elements in A are nonzero

Using the concept of similar matrices, we can obtain a more efficient approach.

Example: Matrix Powers

Suppose A is a 2×2 matrix whose eigenvalues and associated eigenvectors are as below. Compute A^{100}.

$$\lambda_1 = -\frac{1}{2}, \vec{v}_1 = \begin{pmatrix}2\\1\end{pmatrix}; \quad \lambda_2 = \frac{1}{2}, \vec{v}_2 = \begin{pmatrix}1\\-2\end{pmatrix}$$

Because the eigenvalues of A are distinct, we can diagonalize A.

$$A = PDP^{-1}$$
$$A^2 = PDP^{-1}PDP^{-1} = PD^2P^{-1}$$
$$A^3 = PDP^{-1}PDP^{-1}PDP^{-1} = PD^3P^{-1}$$
$$\vdots \qquad \vdots$$
$$A^k = PD^kP^{-1}$$

$$\begin{pmatrix}2&1\\1&-2\end{pmatrix}^{-1} = \frac{1}{-4-1}\begin{pmatrix}-2&-1\\-1&2\end{pmatrix} = \frac{1}{-5}\begin{pmatrix}-2&-1\\-1&2\end{pmatrix} = \frac{1}{5}\begin{pmatrix}2&1\\1&-2\end{pmatrix}$$

Thus, to compute A^{100}, we can instead compute $PD^{100}P^{-1}$.
Using these values, A^{100} becomes

$$A^{100} = PD^{100}P^{-1} = \begin{pmatrix}2&1\\1&-2\end{pmatrix}\begin{pmatrix}-1/2&0\\0&1/2\end{pmatrix}^{100}\left(\frac{1}{5}\begin{pmatrix}2&1\\1&-2\end{pmatrix}\right)$$
$$= \frac{1}{5}\begin{pmatrix}2&1\\1&-2\end{pmatrix}\begin{pmatrix}2^{-100}&0\\0&2^{-100}\end{pmatrix}\begin{pmatrix}2&1\\1&-2\end{pmatrix}$$
$$= \frac{1}{5(2^{100})}\begin{pmatrix}2&1\\1&-2\end{pmatrix}\begin{pmatrix}1&0\\0&1\end{pmatrix}\begin{pmatrix}2&1\\1&-2\end{pmatrix} = \frac{1}{5(2^{100})}\begin{pmatrix}2&1\\1&-2\end{pmatrix}\begin{pmatrix}2&1\\1&-2\end{pmatrix} = \frac{1}{5(2^{100})}\begin{pmatrix}5&0\\0&5\end{pmatrix}$$
$$= \frac{1}{2^{100}}\begin{pmatrix}1&0\\0&1\end{pmatrix}$$
$$A^{100} = 2^{-100}I_2$$

Exercise:

Suppose matrix A is 2×2 and has the following eigenvectors and eigenvalues:

$$\lambda_1 = -1, \vec{v}_1 = \begin{pmatrix}2\\1\end{pmatrix}; \quad \lambda_2 = 0, \vec{v}_2 = \begin{pmatrix}5\\3\end{pmatrix}$$

Now suppose we want to compute A^{10}.
If $A^{10} = \begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}$, what is a_{11} equal to?

$$A^{10} = PD^{10}P^{-1}, \quad P = \begin{pmatrix}2&5\\1&3\end{pmatrix}, \quad D = \begin{pmatrix}-1&0\\0&0\end{pmatrix}$$

$$P^{-1} = \frac{1}{2(3)-5(1)}\begin{pmatrix}3&-5\\-1&2\end{pmatrix} = \begin{pmatrix}3&-5\\-1&2\end{pmatrix}$$

$$= \begin{pmatrix}2&5\\1&3\end{pmatrix}\begin{pmatrix}-1&0\\0&0\end{pmatrix}^{10}\begin{pmatrix}3&-5\\-1&2\end{pmatrix} = \begin{pmatrix}2&5\\1&3\end{pmatrix}\begin{pmatrix}-1^{10}&0\\0&0^{10}\end{pmatrix}\begin{pmatrix}3&-5\\-1&2\end{pmatrix}$$
$$= \begin{pmatrix}2&5\\1&3\end{pmatrix}\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}3&-5\\-1&2\end{pmatrix} = \begin{pmatrix}2&0\\1&0\end{pmatrix}\begin{pmatrix}3&-5\\-1&2\end{pmatrix} = \begin{pmatrix}6&-10\\3&-5\end{pmatrix} \therefore a_{11} = 6$$

TOPIC 2: Complex Eigenvalues

Review of Complex Numbers

Set of imaginary (or complex) numbers, $\mathbb{C} = \{a+bi \mid a,b \in \mathbb{R}\}$, $i = \sqrt{-1}$

Addition and Subtraction Examples:

$$(2-3i)+(-1+i) = (2-1)+(-3+1)i = 1-2i$$

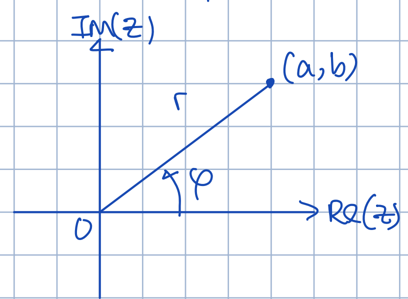
$$\begin{aligned}(2-3i)(-1+i) &= -2+2i+3i-3i^2 \\ &= -2+5i+3 \qquad\qquad i^2=-1. \\ &= 1+5i\end{aligned}$$

We can conjugate complex numbers: $\overline{a+bi} = a-bi$

The absolute value of a complex number: $|a+bi| = \sqrt{a^2+b^2}$

We can write complex numbers in polar form: $a+ib = r(\cos\varphi + i\sin\varphi)$, where:

$$r = |a+bi| \qquad \tan\varphi = \frac{b}{a}$$



Complex Conjugate Properties:

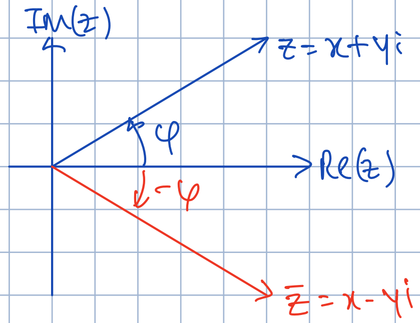
If x and y are complex numbers, $\vec{v} \in \mathbb{C}^n$, it can be shown that:

* $\overline{(x+y)} = \bar{x}+\bar{y}$
* $\overline{A\vec{v}} = A\overline{\vec{v}}$, $A \in \mathbb{R}^{n\times n}$
* $\mathrm{Im}(x\bar{x}) = 0$

Example: (True or False) If x and y are complex numbers, then $\overline{(xy)} = (\bar{x}\bar{y})$ TRUE

Let $x = a+bi$, $y = c+di$.
$$\begin{aligned}\overline{xy} &= \overline{(a+bi)(c+di)} \\ &= \overline{(ac-bd)+i(ad+bc)} \\ &= ac-bd-i(ad-bc) \\ &= (a-bi)(c-di) = \bar{x}\,\bar{y} \text{ (QED).}\end{aligned}$$

Conjugation reflects points across the real axis.



Exercise: (True or False) If $k \in \mathbb{R}$ and $z \in \mathbb{C}$, then $\overline{kz} = k\overline{z}$. TRUE.

Let $z = a+bi$. $\overline{kz} = \overline{k(a+bi)} = \overline{ak+bki}$
$= ak - bki$
$= k(a-bi) = k\overline{z}$ (QED).

Exercise: (True or False) If A is a real 2×2 matrix and $\vec{v}$ is a vector with two complex entries, then $\overline{A\vec{v}} = A\overline{\vec{v}}$. TRUE.

If $x_1, x_2 \in \mathbb{C}$, then $\overline{x_1+x_2} = \overline{x_1} + \overline{x_2}$.
If $x_1, x_2 \in \mathbb{C}$ and $a_{11}, a_{12} \in \mathbb{R}$, then $\overline{a_{11}x_1 + a_{12}x_2} = \overline{a_{11}x_1} + \overline{a_{12}x_2}$
$= a_{11}\overline{x_1} + a_{12}\overline{x_2}$.

Suppose $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, and $\vec{v} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ is a vector with complex entries, then $A\vec{v}$ is a vector with two entries.

$$A\vec{v} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}$$
$$\overline{A\vec{v}} = \begin{pmatrix} a_{11}\overline{x_1} + a_{12}\overline{x_2} \\ a_{21}\overline{x_1} + a_{22}\overline{x_2} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} \overline{x_1} \\ \overline{x_2} \end{pmatrix}$$
$$\overline{A\vec{v}} = A\overline{\vec{v}} \text{ (QED)}$$

Complex Numbers and Polynomials

Theorem: Fundamental Theorem of Algebra
Every polynomial of degree n has exactly n complex roots, counting multiplicity.

For example, $(x-2)^2$ is a 2nd order polynomial. It has two roots
$(x-2)^2(x-1)$ is a 3rd order polynomial, it has three roots

Theorem:
If $\lambda \in \mathbb{C}$ is a root of a real polynomial $p(x)$, then the conjugate $\overline{\lambda}$ is also a root of $p(x)$.

Because of this theorem, if λ is an eigenvalue of real matrix A with eigenvector $\vec{v}$, then $\overline{\lambda}$ is an eigenvector of A with eigenvector $\overline{\vec{v}}$.

$$A\vec{v} = \lambda\vec{v}$$
$$\overline{A\vec{v}} = \overline{\lambda\vec{v}}$$
$$A\overline{\vec{v}} = \overline{\lambda}\,\overline{\vec{v}}$$

Examples:
① If A is a 2×2 matrix and one of its eigenvectors is $\vec{v} = \begin{pmatrix} 1 \\ i \end{pmatrix}$, give another eigenvector of A that is not a multiple of $\vec{v}$. $\overline{\vec{v}} = \begin{pmatrix} 1 \\ -i \end{pmatrix}$

② Four of the eigenvalues of a 7×7 matrix are -2, $4+i$, $-4-i$, and i. What are the other eigenvalues? $\Rightarrow 4-i,\ -4+i,\ -i$

Exercises:

① Suppose A is a 2×2 matrix and one of its eigenvectors is $\vec{v}_1 = \begin{pmatrix} 2 \\ 1+i \end{pmatrix}$.
Which of the following vectors is also an eigenvector of A?

$\begin{pmatrix} 2 \\ 1-i \end{pmatrix}$ ✓ $\because$ conjugate of $\vec{v}_1$

$\begin{pmatrix} 4 \\ 2+2i \end{pmatrix}$ ✓ $\because$ multiple of $\vec{v}_1$

② Suppose A is a square matrix and $\lambda = 1+i$ is an eigenvalue of A.
Which of the following would have to be an eigenvalue of A?

$1-i$ ✓ $\because$ If λ is an eigenvalue of A, then the conjugate of λ is also an eigenvalue of A.

i ✗ $\because$ If λ is an eigenvalue of A, then the conjugate of λ is also an eigenvalue of A. Otherwise, we do not know enough about A to determine what its other eigenvalues could be.

Rotations, Dilations and Eigenvalues

$\varphi : \phi\ \varphi$

The standard matrix for the transform that rotates vectors by $\varphi = \pi/4$ radians about the origin, and then scales (or dilates) vectors by $r = \sqrt{2}$, is

$$A = \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix}\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

What are the eigenvalues of A? Express them in polar form.

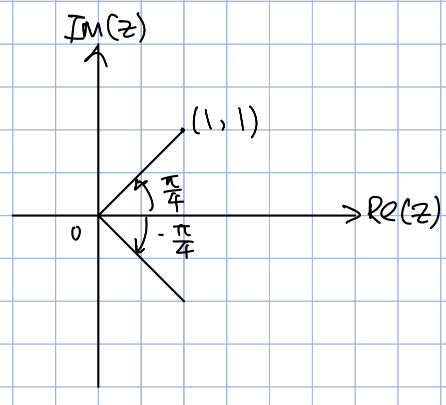
$$0 = \det(A - \lambda I) = (1-\lambda)^2 + 1 = \lambda^2 - 2\lambda + 2$$

$$\lambda = \frac{2 \pm \sqrt{4 - 4(1)(2)}}{2}$$

$$= 1 \pm \tfrac{1}{2}\sqrt{-4} = 1 \pm i$$

$$= \sqrt{1^2+1^2}\left(\cos\tfrac{\pi}{4} \pm i \sin\tfrac{\pi}{4}\right)$$

$$= r(\cos\varphi \pm i \sin\varphi)$$



General Rotation-Dilation Case

The matrix in the previous example is an example of a rotation-dilation matrix.

A rotation-dilation matrix has the form:

$$C = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

Calculate the eigenvalues of C and express them in polar form.

$$0 = \det(C - \lambda I) = (a-\lambda)^2 + b^2 : \lambda^2 - 2a\lambda + (a^2+b^2)$$

$$\lambda = \frac{2a \pm \sqrt{4a^2 - 4(1)(a^2+b^2)}}{2} = a \pm \frac{\sqrt{-4b^2}}{2}$$

$$= a \pm bi = r(\cos\varphi \pm i \sin\varphi) \text{ where } r^2 = a^2+b^2,\ \tan\varphi = \frac{b}{a}.$$

Rotation-Dilation Matrices

DEFINITION — Rotation-Dilation Matrix

↳ A matrix of the form $C = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ is a rotation-dilation matrix because it is the composition of a rotation by φ and dilation by r, where

$$r^2 = a^2 + b^2, \quad \tan\varphi = \frac{b}{a}$$

Moreover, the eigenvalues of C are $\lambda = a \pm bi$.

Example: Determine the eigenvalues of $A = \begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}$

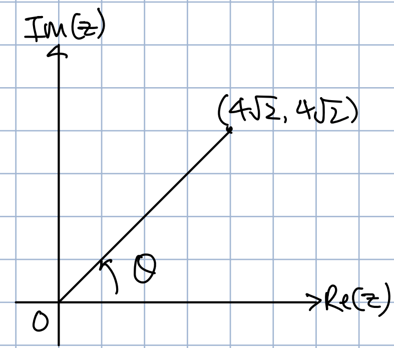
Because this is a rotation-dilation matrix, we do not need to determine the characteristic polynomial. The eigenvalues of this matrix are

$$\lambda = 2 \pm 3i$$

Exercise:
Suppose $T(\vec{x}) = A\vec{x}$ is a linear transform that maps vectors in $\mathbb{R}^2$ to vectors in $\mathbb{R}^2$. Matrix A is, therefore, 2×2. Suppose also that A is a rotation-dilation matrix, so T scales vectors in $\mathbb{R}^2$ by a factor of k, and rotates them by an angle θ. Assume $k \geq 0$, the rotation is counter-clockwise, and $0 \leq \theta \leq \pi$.

If the eigenvalues of A are $\lambda = \sqrt{2}(4 \pm 4i)$,

(a) what must k be equal to?



Eigenvalues of A, $\lambda = \sqrt{2}(4 \pm 4i) = 4\sqrt{2} \pm 4\sqrt{2}i$

$r = k$. $a = b = 4\sqrt{2}$

$k^2 = (4\sqrt{2})^2 + (4\sqrt{2})^2$

$= 64$

$k = \pm 8.$

$\therefore$ Since $k \geq 0$, $k = 8$ ✱

(b) what must θ be equal to?

$\tan\theta = \frac{4\sqrt{2}}{4\sqrt{2}} = 1 \Rightarrow \text{ref } \theta = \tan^{-1} 1 = \pi/4$ rad.

$\therefore$ Given $0 \leq \theta \leq \pi$, $\theta = \pi/4$ rad. ✱

The PCP^{-1} Decomposition

Theorem:
If A is a real 2×2 matrix with eigenvalue $\lambda = a - bi$ (where $b \neq 0$) and associated eigenvector $\vec{v}$, then we may construct the decomposition

$$A = PCP^{-1}$$

→ Find proof from other sources.

where

$$P = (\mathrm{Re}\,\vec{v} \quad \mathrm{Im}\,\vec{v}) \text{ and } C = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

* C is referred to as a rotation dilation matrix, because it is the composition of a rotation by φ and dilation by r.
* the $A = PCP^{-1}$ decomposition allows us to compute large powers of A efficiently.

Example: If possible, construct matrices P and C such that $AP = PC$. The eigenvalues of A are given.

$$A = \begin{pmatrix} 1 & -2 \\ 1 & 3 \end{pmatrix}, \quad \lambda = 2 \pm i$$

$$(\underbrace{A - \lambda I}|0) = \left(\begin{array}{cc|c} 1-\lambda & 2 & 0 \\ 1 & 3-\lambda & 0 \end{array}\right)$$

$A - \lambda I$ has to be singular, so these rows are expected to be multiples of each other

$(1-\lambda)x_1 - 2x_2 = 0.$

x_2 is free. Set $x_2 = 1 - \lambda$.

$(1-\lambda)x_1 - 2(1-\lambda) = 0 \Rightarrow x_1 = 2.$

$$\text{Eigenvector } \vec{v} = \begin{pmatrix} 2 \\ 1-\lambda \end{pmatrix} = \begin{pmatrix} 2 \\ 1-(2-i) \end{pmatrix} = \begin{pmatrix} 2 \\ -1+i \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix} + i\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$\because$ choose $2-i$ (either one is fine)

$$\therefore P = (\mathrm{Re}(\vec{v}) \quad \mathrm{Im}(\vec{v})) = \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$$

Exercise: Suppose A is the matrix below.

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$$

Which of the following vectors are eigenvectors of A?

$\begin{pmatrix} i \\ -1 \end{pmatrix}$ ✓ $\quad \begin{pmatrix} i \\ 1 \end{pmatrix}$ ✓ $\quad \begin{pmatrix} -i \\ 1 \end{pmatrix}$ ✓

By inspection, A is a rotation-dilation matrix. $\Rightarrow \lambda = 2 \pm i$

$$(A - \lambda I|0) = \left(\begin{array}{cc|c} 2-\lambda & -1 & 0 \\ \cdot & \cdot & \cdot \end{array}\right)$$

$(2-\lambda)x_1 - x_2 = 0. \Rightarrow (2-\lambda)x_1 = x_2$

$$\vec{v} = \begin{pmatrix} 1 \\ 2-\lambda \end{pmatrix} = \begin{pmatrix} 1 \\ 2-(2\pm i) \end{pmatrix} = \begin{pmatrix} 1 \\ \mp i \end{pmatrix}$$

TOPIC 3: Google Page Rank

Example: Car Rental company (from Module 3)

A car rental company has 3 rental locations, A, B and C. Cars can be returned at any location. The table below gives the pattern of rental and returns for a given week.

		rented from		
		A	B	C
returned to	A	0.8	0.1	0.2
	B	0.2	0.6	0.3
	C	0	0.3	0.5

There are 1000 cars at each location today. What happens to the distribution of cars after a long time?

Can use the transition matrix, P, to find the distribution of cars after 1 week.

$$\vec{x}_1 = P\vec{x}_0, \quad P = \frac{1}{10}\begin{pmatrix} 8 & 1 & 2 \\ 2 & 6 & 3 \\ 0 & 3 & 5 \end{pmatrix}$$

The distribution of cars after n weeks is

$$\vec{x}_n = P^n \vec{x}_0.$$

Because P is regular stochastic, $\vec{x}_n$ tends to a steady-state, which we can find by solving

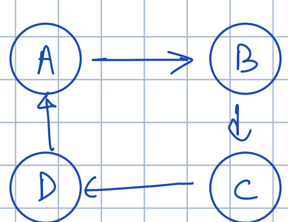
$$(P - I)\vec{q} = \vec{0}$$

(defn)
Recall: A stochastic matrix P is <u>regular</u> if there is some k such that P^k only contains strictly positive entries.

* We can determine whether a matrix, P, is regular stochastic by computing P^k for $k = 2, 3, 4, \ldots$
But sometimes we can see from inspection that a matrix will not be regular stochastic.

Determining Whether A Stochastic Matrix is Regular

Example:
By Inspection, is the corresponding stochastic matrix regular? Is there a steady-state?



$$P = \begin{array}{c} \begin{matrix} A & B & C & D \end{matrix} \\ \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}, \quad \vec{x}_k = P\vec{x}_{k-1}$$

$$\vec{x}_0 = \frac{1}{4}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \vec{q}$$

Theorem: Regular Stochastic Matrices

If P is a regular $m \times m$ stochastic matrix with $m \geq 2$, then:

↳ for any initial probability vector $\vec{x}_0$, $\lim_{n\to\infty} P^n \vec{x}_0 = \vec{q}$

↳ P has a unique eigenvector, $\vec{q}$, which has eigenvalue $\lambda = 1$

↳ there is a stochastic matrix Π such that $\lim_{n\to\infty} P^n = \Pi$

↳ each column of Π is the same probability vector $\vec{q}$

↳ the eigenvalues of P satisfy $|\lambda| \leq 1$

Long Term Behavior

To investigate the long-term behavior of a system that has a regular stochastic matrix P, we could:

* compute the steady-state vector, $\vec{q}$, by solving $(P-I)\vec{q} = \vec{0}$
* compute $P^n \vec{x}_0$ for large n
* compute P^n for large n, each column of the resulting matrix is the steady-state

Computing P^n for large n requires a computer. Students would not see much problems on exams, but they may appear on homework and other parts of the course.

Examples: Steady-State

(1) True or False: a steady-state vector for a stochastic matrix is an eigenvector

True. $P\vec{q} = \vec{q}$, $\lambda = 1$

(2) Give an example of a 2x2 stochastic matrix, A, that is in echelon form. A steady-state vector for the Markov chain $\vec{x}_{k+1} = A\vec{x}_k$ is $\vec{q} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

$$A = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A\vec{q} = \vec{q}.$$

Example: Convergence

If P is a regular stochastic matrix with steady state vector $\vec{r} = \frac{1}{6}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\vec{x}_0 = \frac{1}{10}\begin{pmatrix} 9 \\ 0 \\ 1 \end{pmatrix}$, what does the sequence $\vec{x}_k = P^k \vec{x}_0$ converge to?

$$\vec{x}_k \to \vec{r} \text{ as } k \to \infty$$

Example: Long Term Behavior

Consider the Markov chain

$$\vec{x}_k = A\vec{x}_{k-1} = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}\vec{x}_{k-1}, \quad k=1,2,3,\ldots \quad \vec{x}_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

The eigenvalues of A are 1 and 0.6. Analyze the long-term behavior of the system. In other words, determine what $\vec{x}_k$ tends to as $k \to \infty$.

$$(A - I \mid 0) = \left(\begin{array}{cc|c} -0.2 & 0.2 & 0 \\ 0.2 & -0.2 & 0 \end{array}\right) \Rightarrow \vec{v} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\text{as } k \to \infty, \quad \vec{x}_k \to \vec{v} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Exercise:

Suppose P is a regular stochastic matrix with steady state vector

$$\vec{w} = \frac{1}{10}\begin{pmatrix} 4 \\ 2 \\ 4 \end{pmatrix}$$

and

$$\vec{x}_0 = \frac{1}{10}\begin{pmatrix} 3 \\ 3 \\ 4 \end{pmatrix}$$

As k goes to infinity, the sequence $\vec{x}_k = P^k \vec{x}_0$ converges to the vector below.

$$\vec{q} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

What is x_1 equal to?

P is regular stochastic, so the sequence will converge to the steady-state no matter what $\vec{x}_0$ happens to be (as long as x_0 is a probability vector). $\therefore x_1 = 0.4$

Google PageRank

There are many search engines that we can use to find relevant information on the web.

* when searching for information on the Internet using any search engine, we can be presented with many search results
* for a search engine to give useful information to the user, it must quickly order search results in some way
* essentially: how can the engine quickly decide which results appear at the top of the list?

Mathematical Model for Web Traffic

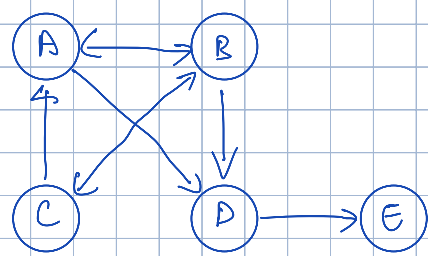
The PageRank algorithm is based on a mathematical model that assumes that we have:

* a collection of web pages that have links to each other
* users who are navigating the web
* a set of rules that govern how the users navigate the web

We impose assumptions about how the users navigate the web?
(a) A user on a web page is equally likely to go to any page that their page links to
(b) If a user is on a page that does not link to other pages, the user stays at their page.
(c) The distribution of users can be modeled using a Markov process, $\vec{x}_{k+1} = P\vec{x}_k$, where
   * $\vec{x}_k \in \mathbb{R}^n$ is a probability vector, gives the proportion of users on each page at iteration k
   * P is an $n \times n$ stochastic matrix — tells us how users transition from one iteration to the next
   * n is the number of pages in the web

Example Web with Five Pages

A set of web pages link to each other according to the diagram below. Use the assumptions on the previous slide to construct a Markov chain that represents how users navigate the web.



A → B or D (2 ways)
B → A, C or D (3 ways)
C → A or B (2 ways)
D → E (1 way)
E → ϕ (assumed all users stay on that page)

$$\vec{x}_{k+1} = P\vec{x}_k \ , k = 0, 1, \ldots$$

$$P = \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{pmatrix} 0 & 1/3 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/3 & 0 & 0 & 0 \\ 1/2 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

called a transition matrix

Transition Matrix, Importance and PageRank

* The square matrix we constructed in the previous example is a <u>transition matrix</u>. It describes how users transition between pages in the web.
* The steady-state vector, $\vec{q}$, for the Markov-chain, can characterize the long-term behavior of users in a given web.
* The <u>importance</u> of a page in a web are the entries of $\vec{q}$.
* The <u>PageRank</u> is the ranking assigned to each page based on its importance. The highest ranked page has PageRank 1, the second PageRank 2, and so on.
* Two pages with same importance receive the same PageRank (some other method would be needed to resolve ties).

Exercise:

A web consists of exactly three pages, A, B, and C.

* page A only links to B
* page B has links to pages A and C.
* page C does not link to the other pages

The transition matrix for this web, P, has the form

$$P = \begin{pmatrix} 0 & 1/2 & c_1 \\ a & b & c_2 \\ 0 & 1/2 & c_3 \end{pmatrix}$$

$\boxed{a=1, b=0. \quad c_1=c_2=0, c_3=1.}$

What are the values of a, b, c_1, c_2 and c_3?

Remaining Questions

Our simple mathematical model has some limitations that must be addressed for it to be useful.

* Will our transition be regular stochastic?
* What can we do to build a model that will give us a regular stochastic matrix?
* What can we do to better handle the pages that do not link to other pages?

Adjustments Needed

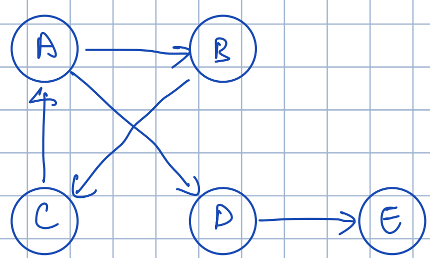
Our mathematical model for PageRank (PR) has two problems:

* the transition matrix is not regular: we do not have a unique steady-state
* pages that do not link to other pages can have the largest importance, or highest PageRank

Adjustment 1: If a user reaches a page that does not link to other pages, the user will choose any page in the web, with equal probability, and move to that page.

↳ we will denote this modified transition matrix as P_*.

Example:



P and P_* are as follows:

$$P = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}, \quad P_* = \begin{pmatrix} 0 & 0 & 1 & 0 & 1/5 \\ 1/2 & 0 & 0 & 0 & 1/5 \\ 0 & 1 & 0 & 0 & 1/5 \\ 1/2 & 0 & 0 & 0 & 1/5 \\ 0 & 0 & 0 & 1 & 1/5 \end{pmatrix}$$

Adjustment 2: A user at any page will navigate to any page among those that their page links to with equal probability p, and to any page in the web with equal probability $1-p$. The transition matrix becomes

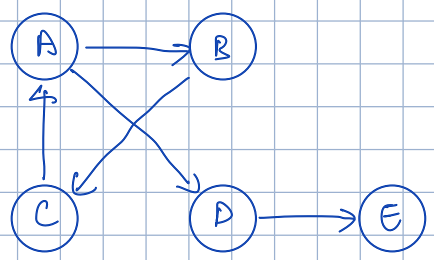
$$G = pP_* + (1-p)K.$$

All the elements of the $n \times n$ matrix K are equal to $\frac{1}{n}$.

Note: * p is referred to as the damping factor. Google is said to use $p = 0.85$.

* Adjustment 2 forces G to be regular stochastic when $0 < p \le 1$.

Google Matrix Example



The Google matrix for this web, with $p=0.85$ is $G = 0.85P_* + 0.15K$, where

$$P_* = \begin{pmatrix} 0 & 0 & 1 & 0 & 1/5 \\ 1/2 & 0 & 0 & 0 & 1/5 \\ 0 & 1 & 0 & 0 & 1/5 \\ 1/2 & 0 & 0 & 0 & 1/5 \\ 0 & 0 & 0 & 1 & 1/5 \end{pmatrix}, \quad K = \frac{1}{5}\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Computing Page Rank

Because G is stochastic, for any initial probability vector $\vec{x}_0$,

$$\lim_{n\to\infty} G^n \vec{x}_0 = \vec{q}.$$

We can obtain steady-state evaluating $G^n\vec{x}_0$ for large n, by solving $G\vec{q} = \vec{q}$, or by evaluating $\vec{x}_n = G\vec{x}_{n-1}$ for large n.

Elements of the steady-state vector give the importance of each page in the web, which can be used to determine PageRank.

Largest element in steady-state vector corresponds to page with PageRank 1, second largest with PageRank 2, and so on.

Exercise:
A web consists of exactly three pages, A, B and C.
* page A only links to B.
* page B has links to pages A and C
* page C does not link to the other web pages

The transition matrix for this web, P, is

$$P = \begin{pmatrix} 0 & 1/2 & 0 \\ 1 & 0 & 0 \\ 0 & 1/2 & 1 \end{pmatrix}$$

After making adjustments 1 and 2, and using a damping factor of 0.85, we can construct the Google Matrix, G. The Google matrix has the following form.

$$G = pP_* + (1-p)K \qquad G = 0.85P_* + 0.15K$$

What is every entry of matrix K equal to? $\boxed{\frac{1}{n} = \frac{1}{3}.}$

Matrix P_* has the form below.

$$P_* = \begin{pmatrix} 0 & 1/2 & a \\ 1 & 0 & b \\ 0 & 1/2 & c \end{pmatrix}$$

Find a, b, and c. $a = b = c = \frac{1}{3}$ *

With adjustment 2, a user will navigate to any page in the web. So every entry in the last column of the adjusted transition matrix will be equal to $\frac{1}{3}$.