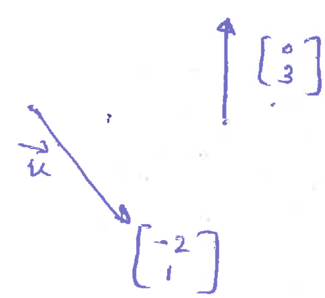
6.4. Eigenvalues: power method

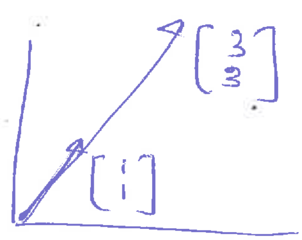
Let A be an $n \times n$ matrix. If $A\vec{x} = \lambda \vec{x}$ for some nonzero vector $\vec{x}$ and a scalar λ, then λ is an eigenvalue of A, and $\vec{x}$ is an eigenvector of A.

ex. $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ $\quad \vec{u} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ $\quad \vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$A\vec{u} = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$



$$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$



* Infinity norm

$$\|\vec{x}\|_\infty = \max_{1 \le i \le n} |x_i|$$

The characteristic polynomial of A is $\det(A-\lambda I)$, and its roots are the eigenvalues of A.

The eigenvectors are ~~the~~ nonzero solutions of

$$(A-\lambda^* I)\vec{x} = \vec{0}.$$

Def^n: A dominant eigenvalue of $A_{n\times n}$ is an eigenvalue λ, whose magnitude is greater than all other eigenvalues of A. (If it exists, an eigenvector associated to λ is called a dominant eigenvector).

$|\lambda_1| > |\lambda_i| \quad i = 2, \dots, n.$

λ_1 - dominant eigenvectors.

ex. Find the dominant eigenvalue of the matrix.

$$A = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}$$

$$\det(A-\lambda I) = 0$$

$$\begin{vmatrix} 2-\lambda & -12 \\ 1 & -5-\lambda \end{vmatrix} = 0$$

$$(2-\lambda)(-5-\lambda) + 12 = 0$$

$$\lambda^2 + 3\lambda + 2 = 0$$

$$(\lambda+1)(\lambda+2) = 0$$

$\lambda = -2 \leftarrow$ dominant.

$\lambda = -1$

Note Not every matrix has a dominant eigenvalue.

ex.
$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad \lambda_1 = 1 \quad |\lambda_1| = 1$$
$$\lambda_2 = -1 \quad |\lambda_2| = 1.$$

(*) Goal: Find the dominant eigenvector and the dominant eigenvalue(s)

→ "go to page 4"

The motivation behind power m/d is that multiplication by a matrix tends to move vectors toward the dominant eigenvector direction.

Skip!

- First we assume that the matrix A has a dominant eigenvalue with corresponding eigenvectors.
- Then, we choose an initial approximation $\vec{x_0}$ of one of the dominant eigenvectors of A.

$$\vec{x_1} = A\vec{x_0}$$
$$\vec{x_2} = A\vec{x_1} = A \cdot A\vec{x_0} = A^2\vec{x_0}$$
$$x_3 = A^3\vec{x_0}$$
$$\vdots$$
$$x_k = Ax_{k-1} = A\left(A^{k-1}\vec{x_0}\right) = A^k\vec{x_0}.$$

For large powers of k, we obtain a good approximation of the dominant eigenvector of A (by properly scaling).

ex. The matrix $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$ has a dominant eigenvalue of 4 with eigenvector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Consider the random starting vector: $\begin{bmatrix} -5 \\ 5 \end{bmatrix}$.

Then $\vec{x}_1 = A x_0 = \begin{bmatrix} 10 \\ 0 \end{bmatrix}$

$\vec{x}_2 = A\vec{x}_1 = \begin{bmatrix} 10 \\ 20 \end{bmatrix}$

$\vec{x}_3 = A\vec{x}_2 = \begin{bmatrix} 70 \\ 20 \end{bmatrix}$

$\vec{x}_4 = A\vec{x}_3 = \begin{bmatrix} 250 \\ 260 \end{bmatrix} = 260 \begin{bmatrix} \frac{25}{26} \\ 1 \end{bmatrix} = 260 \begin{bmatrix} 0.9615 \\ 1 \end{bmatrix}$

$\Rightarrow$ Multiplying a random starting vector repeatedly by the matrix A has resulted in moving the vector $\vec{x}_k$ very close to the dominant eigenvector of A

In general,

the power method constructs a seq. $\{ \vec{x}_k \}$ defined by

$$\vec{x}_k = \frac{A^k \vec{x}_0}{\| A^k \vec{x}_0 \|_\infty} \qquad k \geq 1$$

with an initial approximation $\vec{x}_0$ to approximate the dominant eigenvector.

Th$\underline{\text{m}}$

Convergence of Power method.

Let A be an $n \times n$ matrix with real eigenvalues $\lambda_1, \ldots, \lambda_n$ satisfying $|\lambda_1| > |\lambda_2| \geq \ldots |\lambda_n|$, and ~~assume that~~ λ_1 has multiplicity 1. Assume that there are n-linearly independent eigenvectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ of A. For almost every initial vector power method converges to an eigenvector associated to λ_1.

Proof. Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ be the eigenvectors that forms a basis for $\mathbb{R}^n$, with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$

Then, the initial vectors

$$\vec{x}_0 = c_1 v_1 + c_2 v_2 + \ldots + c_n v_n \quad \text{Such that } c_1 \neq 0.$$

Applying power method yields

$$\vec{x}_1 \; \cancel{\vec{x}} = A\vec{x}_0 = A(c_1 v_1 + c_2 v_2 + \dots + c_n v_n)$$

$$= c_1 A v_1 + c_2 A v_2 + \dots + c_n A v_n$$

$$\cancel{\vec{x}} \quad = c_1 \lambda_1 v_1 + c_2 \lambda v_2 + \dots + c_n \lambda_n v_n$$

$$\cancel{\vec{x}} = A\vec{x}_1 = A(c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \dots + c_n \lambda_n v_n)$$

$$A^2 \vec{x}_0 \quad \cancel{\vec{x}} = c_1 \lambda_1^2 \vec{v}_1 + c_2 \lambda_2^2 + \dots + c_n \lambda_n^2 v_n$$

$$\vdots$$

$$A^k \vec{x}_0 = c_1 \lambda_1^k \vec{v}_1 + c_2 \lambda_2^k \vec{v}_2 + \dots + c_n \lambda_n^k \vec{v}_n$$

$$\vec{x}_k = \frac{A^k \vec{x}_0}{\| A^k \vec{x}_0 \|} = \frac{\lambda_1^k \left[c_1 \vec{v}_1 + c_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k \vec{v}_2 + \dots + c_n \left(\frac{\lambda_n}{\lambda_1}\right)^k \vec{v}_n \right]}{\left\| \lambda_1^k \left[c_1 \vec{v}_1 + c_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k \vec{v}_2 + \dots + c_n \left(\frac{\lambda_n}{\lambda_1}\right)^k \vec{v}_n \right] \right\|}$$

as ~~k → ∞~~

$$\rightarrow \frac{\cancel{\lambda_1^k c_1 \vec{v}_1}}{\cancel{\| \lambda_1^k c_1 \vec{v}_1 \|}}$$

$$= \frac{\lambda_1^k}{|\lambda_1^k|} \frac{\left[c_1 \vec{v}_1 + c_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k \vec{v}_2 + \dots + c_n \left(\frac{\lambda_n}{\lambda_1}\right)^k \vec{v}_n \right]}{\| \quad \text{"} \quad \|}$$

$$\vec{x}_k = \frac{A^k \vec{x}_0}{\|A^k \vec{x}_0\|} = \pm \frac{\left[c_1 \vec{v}_1 + c_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k \vec{v}_2 + \dots + c_n \left(\frac{\lambda_n}{\lambda_1}\right)^k \vec{v}_n \right]}{\| \quad \|}$$

$\left(\frac{\lambda_i}{\lambda_1}\right)^k \to 0$ Since $|\lambda_1| > |\lambda_i| \quad \forall i > 1$

$$\vec{x}_k \to \pm \frac{c_1 \vec{v}_1}{\|c_1 \vec{v}_1\|} = \pm \frac{\vec{v}_1}{\|\vec{v}_1\|} \quad \text{as } k \xrightarrow{c_1 \neq 0} \infty$$

(III) Th^{m} Determining an eigenvalue from an eigenvector.

$$A\vec{v}_1 = \lambda_1 \vec{v}_1$$

$$\vec{v}_1^T A \vec{v}_1 = \vec{v}_1^T \lambda_1 \vec{v}_1 = \lambda_1 \vec{v}_1^T \vec{v}_1$$

$$\lambda_1 = \frac{\vec{v}_1^T A \vec{v}_1}{\vec{v}_1^T \vec{v}_1} \quad \text{Rayleigh quotient.}$$

If $\vec{x}_k \approx \vec{v}_1$, then

$$\lambda_1 \approx \lambda_1^{(k)} = \frac{\vec{x}_k^T A \vec{x}_k}{x_k^T x_k}$$

Note. 1. The power method does not work for $c_1 = 0$.

2. The rate of convergence $\{\vec{x}_k\}$ is faster when $\left|\frac{\lambda_2}{\lambda_1}\right|$ is smaller.

ex. Do three iterations by the power method with $\vec{x}_0 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ to approximate the dominant eigenvalue of $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ with its eigenvector.

$$A\vec{x}_0 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 3 \end{bmatrix} \Rightarrow \vec{x}_1 = \frac{A\vec{x}_0}{\|A\vec{x}_0\|_\infty} = \begin{bmatrix} 8 \\ 7 \end{bmatrix} = \frac{1}{8}\begin{bmatrix} 1 \\ 7/8 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 0.87 \end{bmatrix}$$

$$A\vec{x}_1 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0.87 \end{bmatrix} \Rightarrow \vec{x}_2 = \frac{1}{2.87}\begin{bmatrix} 2.74 \\ 2.87 \end{bmatrix} = \begin{bmatrix} 0.95 \\ 1 \end{bmatrix}$$

$$A\vec{x}_2 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 0.95 \\ 1 \end{bmatrix} = \begin{bmatrix} 2.95 \\ 2.9 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.98 \end{bmatrix}$$

$\cancel{A\vec{x}_3} = [$

$$\lambda_1^{(1)} = \frac{x_1^T A x_1}{x_1^T x_1} \approx \frac{5.23}{1.75} = 2.98.$$

$$\lambda_1^{(2)} = \frac{x_2^T A x_2}{x_2^T x_2} \approx \frac{5.70}{1.90} = 3.00$$

$$\lambda_1^{(3)} = \frac{x_3^T A x_3}{x_3^T x_3} \approx \frac{5.88}{1.96} = 3.00$$

The actual $\lambda_1 = 3$ and $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.