Power Iteration

Dynamical System –

$$\vec{x}_{n+1} = A\vec{x}_n$$

example:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}, \quad \text{eigenvectors } \vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$\vec{x}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\vec{x}_1 = A\vec{x}_0 = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

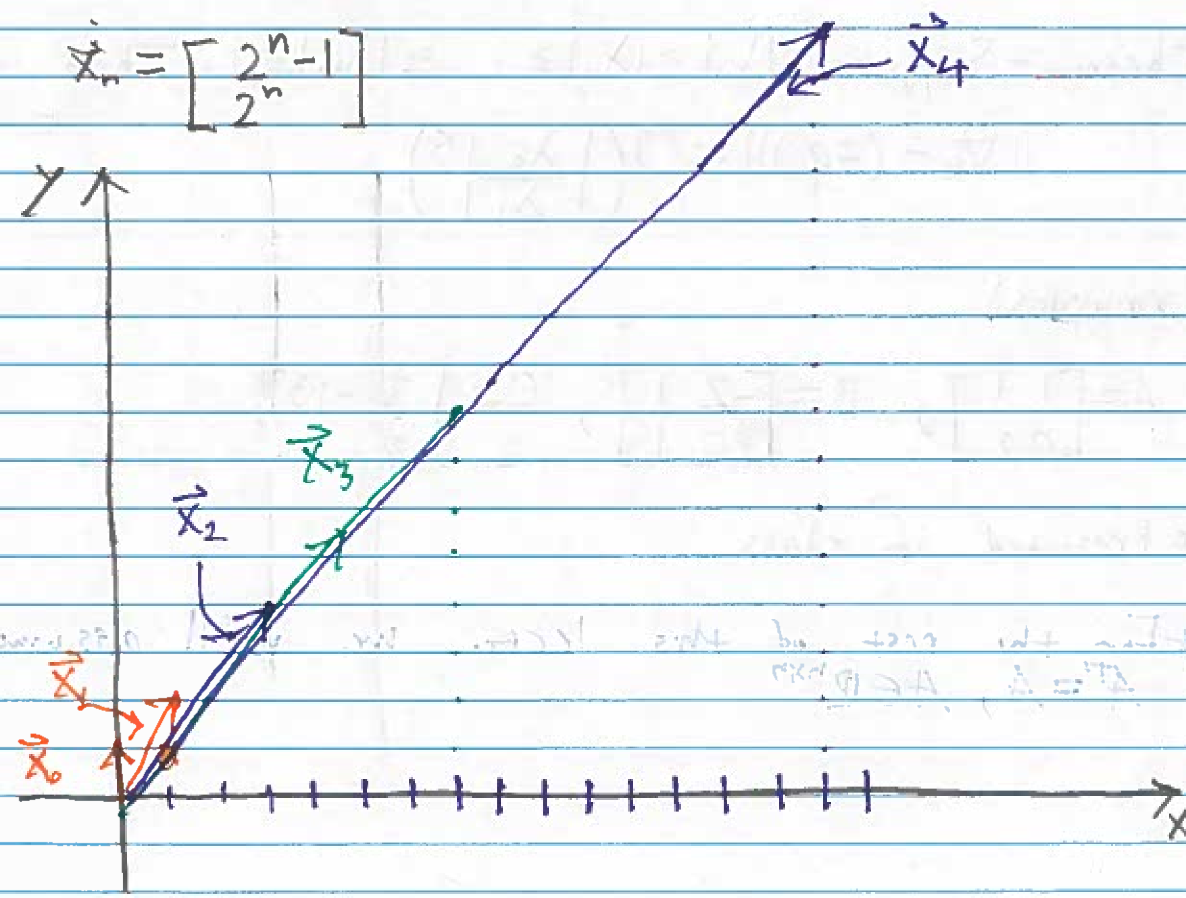
$$\vec{x}_2 = A\vec{x}_1 = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

$$\vec{x}_3 = A\vec{x}_2 = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \end{bmatrix}$$

$$\vec{x}_4 = \begin{bmatrix} 15 \\ 16 \end{bmatrix}$$

$$\vdots$$

$$\vec{x}_n = \begin{bmatrix} 2^n - 1 \\ 2^n \end{bmatrix}$$



The iterates $\vec{v}_n = \frac{x_n}{\|x_n\|}$ converge to $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$!!!

Why does this work??

Let A be a nonsingular matrix with n-eigenvalues $\lambda_1, \ldots, \lambda_n$ and n eigenvectors $\vec{q}_1, \ldots, \vec{q}_n$.

Let $\vec{v}_0 \in \mathbb{R}^n$ satisfy $\|\vec{v}_0\| = 1$. Write $\vec{v}_0$ as a linear combination:

$$\vec{v}_0 = a_1 \vec{q}_1 + \ldots + a_n \vec{q}_n$$

$$\Rightarrow \vec{v}_1 = \frac{A\vec{v}_0}{\|A\vec{v}_0\|}$$

$$= c_1 \cdot (a_1 \lambda_1 \vec{q}_1 + \ldots + a_n \lambda_n \vec{q}_n)$$

$$= c_1 \cdot \lambda_1 \left(a_1 \vec{q}_1 + a_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k + \ldots + a_n \left(\frac{\lambda_n}{\lambda_1} \right) \vec{q}_n \right)$$

$$\vdots$$

$$\vec{v}_k = c_k \lambda_1^k \left(a_1 \vec{q}_1 + a_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k + \ldots + a_n \left(\frac{\lambda_n}{\lambda_1} \right)^k q_n \right)$$

Theorem - Suppose $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_m| \geq 0$. Then

$$\| \vec{v}_k - (\pm q_1) \| = O\left(\left| \frac{\lambda_2}{\lambda_1} \right|^k \right)$$

examples:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} -2 & 1 \\ 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & -13 \\ 5 & 1 \end{bmatrix}$$

* Presented in class.

* For the rest of this lecture we will assume $A^T = A$, $A \in \mathbb{R}^{n \times n}$.

Inverse Iteration

Let $A \in \mathbb{R}^{n \times n}$ with eigenvalues $\lambda_1, \dots, \lambda_n$. Let $\nu \in \mathbb{R}$ with $\nu_i \neq \lambda_i$

Proposition: The eigenvectors of $(A-\nu I)^{-1}$ are the same as the eigenvectors of A, and the eigenvalues are $\{(\lambda_i - \nu)\}^{-1}$.

~~proof~~:

$$(A-\nu I)\vec{v}_i = (\lambda_i - \nu)\vec{v}_i$$

$$\longrightarrow \frac{1}{\lambda_i - \nu}\vec{v}_i = (A-\nu I)^{-1}\vec{v}_i$$

Idea: Pick ν to bias other eigenvalues!

Algorithm - Inverse Iteration

1. $\|v^{(0)}\| = 1$
2. while error < tol
   - Solve $(A-\nu I)w = v^{k-1}$ for w
   - $v^{(k)} = w/\|w\|$
3. loop

Rayleigh Quotient

$\vec{v} \in \mathbb{R}^n$

$$r(\vec{v}) = \frac{\vec{v}^T A \vec{v}}{\vec{v}^T \vec{v}}$$

If $\vec{v}$ is an eigenvector:

$$r(\vec{v}) = \lambda$$

$\uparrow$

Nice property #1

Idea:

Given a vector what is the "best" eigenvalue

$$A\vec{x} "=" \lambda\vec{x}$$

minimize $\|A\vec{x}-\lambda\vec{x}\|$, with λ as the variable.

Least squares:

$$\underbrace{\vec{x}}_{\text{matrix}}\lambda = \underbrace{A\vec{x}}_{\vec{b}}$$

Solve for me

Least squares:

minimize $\|r\| = \|b - A\vec{x}\|$

$r \perp \text{range}(A)$

$\Rightarrow r \in \text{Null}(A^*)$

$A^* r = 0$

$A^*(b - A\vec{x}) = 0$

$A^*A\vec{x} = A^*\vec{b}$ $\rightarrow$ Normal equations

$$\vec{x}\lambda = A\vec{x}$$

Apply least squares:

$$\vec{x}^*\vec{x}\lambda = x^*A\vec{x}$$

$$\Rightarrow \boxed{\lambda = \frac{\vec{x}^*A\vec{x}}{\vec{x}^*\vec{x}}}$$

Proposition: $\nabla r(\vec{x}) = \frac{2}{\vec{x}^*\vec{x}}\left(A\vec{x} - r(\vec{x})\vec{x}\right)$

proof:

$$\frac{\partial r}{\partial x_j} = \frac{\frac{\partial}{\partial x_j}(\vec{x}^* A\vec{x})}{\vec{x}^*\vec{x}} - \frac{\vec{x}^* A\vec{x}\ \frac{\partial}{\partial x_j}(\vec{x}^*\vec{x})}{(\vec{x}^*\vec{x})^2}$$

$$\frac{\partial}{\partial x_j}\vec{x}^* A\vec{x} = \sum_{k=1}^{n}\sum_{i=1}^{n}\frac{\partial}{\partial x_j} x_k A_{ki} x_i$$

$$= \sum_{k=1}^{n}\sum_{i=1}^{n}\left(\delta_{kj} A_{ki} x_i + x_k A_{ki}\delta_{ij}\right)$$

$$= \sum_{i=1}^{n}\sum_{k=1}^{n}\delta_{kj} A_{ki} x_i + \sum_{k=1}^{n}\sum_{i=1}^{n} x_k A_{ki}\delta_{ij}$$

$$= \sum_{i=1}^{n} A_{ji}\vec{x}_i + \sum_{k=1}^{n}(x_k A_{kj}) = 2\,(Ax)_j$$

Follow from $A_{kj} = A_{jk}$

$$\frac{\partial}{\partial x_j}\vec{x}^*\vec{x} = \sum_{i=1}^{n}\frac{\partial}{\partial x_j} x_i \cdot x_i = 2x_j$$

$$\Rightarrow \frac{\partial r}{\partial x_j} = \frac{2(Ax)_j}{\vec{x}^*\vec{x}} - \frac{(\vec{x}^* A\vec{x})\cdot 2x_j}{(\vec{x}^*\vec{x})^2}$$

$$\Rightarrow \nabla r(\vec{x}) = \frac{2}{\vec{x}^*\vec{x}}\left(A\vec{x} - r(x)\vec{x}\right)$$

■

Consequence:

$$r(x) - r(q_i) = O(\|\vec{x} - \vec{q}_i\|^2)$$

eigenvector quadratic accuracy!!

Theorem - Suppose $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_m| \geq 0$. Then, the iterates of the eigenvalues computed by power iteration satisfy

$$|\lambda^{(k)} - \lambda_1| = \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right)$$

as $k \to \infty$. Moreover, $\lim_{k\to\infty} \lambda^{(k)} = \lambda_1$.

Rayleigh Quotient Iteration

$\|v^{(0)}\| = 1$

$\lambda^{(0)} = (\vec{v}^{(0)})^T A \vec{v}^{(0)}$

while error > tol

  Solve $(A - \lambda^{(k-1)} I)\vec{w}^{(k-1)} = v^{(k-1)}$

  $v^{(k)} = \dfrac{\vec{w}^{(k-1)}}{\|\vec{w}^{(k-1)}\|}$

  $\lambda^{(k)} = (\vec{v}^{(k)})^T A \vec{v}^{(k)}$

Theorem - Rayleigh Quotient

$$\|\vec{v}^{(k+1)} - (\pm q_J)\| = \mathcal{O}(\|v^{(k)} - (\pm q_J)\|^3)$$

and

$$|\lambda^{(k+1)} - \lambda_J| = \mathcal{O}(|\lambda^{(k)} - \lambda_J|^3)$$

as $k \to \infty$.

proof:

1. If $\|\vec{v}^{(k)} - q_J\| \leq \varepsilon$ for sufficiently small ε then $|\lambda^{k} - \lambda_J| = \mathcal{O}(\varepsilon^2)$.

2. By inverse iteration

$$\|\vec{v}^{(k+1)} - \vec{q}_J\| = \mathcal{O}\Bigg(\underbrace{\frac{|\lambda^{(k)} - \lambda_J|}{|\lambda^{(k)} - \lambda_K|}}_{\mathcal{O}(\varepsilon^2)}\Bigg) \cdot \underbrace{\|v^{(k)} - q_J\|}_{\mathcal{O}(\varepsilon)}$$

$$= \mathcal{O}(\varepsilon^3) \;!!$$