Lec 22 Power Method

We will continue our study of linear systems, but shift our attention to solving eigenvalue problems

$$Ax = \lambda x$$

MOTIVATION Eigenvalue problems are ubiquitous:

- Differential Equations:

  (I) Schrödinger Equation: $H|\psi\rangle = E|\psi\rangle$

  (II) undamped vibration: $m\frac{d^2}{dt^2}x = -kx$.

- Statistics/Machine Learning: PCA (principal component analysis). Here, the eigenvectors on a multidimensional data set with the largest eigenvalues, are those "directions" in the data set that "explain it" best.

- Search algorithms: Google's Page Rank. The basic idea is that a collection of webpages can be represented as a matrix, called the adjacency matrix

It's principal eigenvector (the one with the largest eigenvalue) then represents a ranking of the web-pages.

POWER METHOD

This is an algorithm for determining the largest eigenvalue of a matrix, and its corresponding eigenvector. It can be modified to extract other evectors/evalues also.

Suppose A, an $n \times n$ matrix, satisfies:

$$A v_i = \lambda_i v_i$$

where

$$\underbrace{|\lambda_1|}_{\text{dominant/principal evalue.}} > |\lambda_2| \geq |\lambda_3| \cdots \geq |\lambda_n|$$

and $v_1, \ldots, v_n$ are linearly independent and of unit length. Then for any $x_0 \in \mathbb{R}^n$, we have:

$$x_0 = \sum_{i=1}^{n} c_i v_i \quad \left(\begin{array}{l}\text{since } v_1 \ldots v_n \text{ is} \\ \text{a basis}\end{array}\right)$$

Let's further assume that $c_1 \neq 0$, ie. x_0 has a component along the principal eigenvector.

Now:

$$A^k x_0 = \sum_i c_i A^k v_i$$

$$= \sum_i c_i \lambda_i^k v_i$$

$$= c_1 \lambda_1^k \left[v_1 + \frac{c_2}{c_1}\left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \cdots + \frac{c_n}{c_1}\left(\frac{\lambda_n}{\lambda_1}\right)^k v_n \right]$$

$$\xrightarrow[k\to\infty]{} c_1 \lambda_1^k v_1$$

Since $(\lambda_i/\lambda_1)^k \to 0$ for $i = 2 \ldots n$.

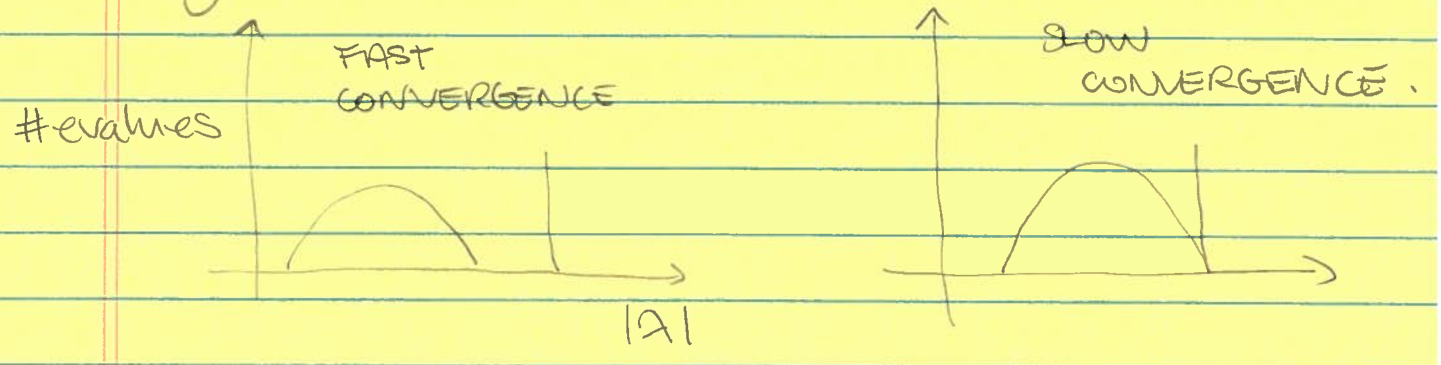
Thus we have a method for extracting v_1 in an iterative manner. There's one problem: $A^k x_0$ grows in magnitude if $|\lambda_1| > 1$. We can correct for this by re-scaling at each step:

Define: $x_k = \dfrac{A^k x_0}{\|A^k x_0\|}$

$$\to \frac{c_1 \lambda_1^k v_1}{c_1 \lambda_1^k} = v_1$$

$\|v_1\| = 1$.

CONVERGENCE Convergence is geometric with the largest multiplicative factor being λ_2/λ_1. This means that the method converges slowly if there is an evalue close (in magnitude) to the dominant eigenvalue.



How do we find the principal eigenvalue?

Define:

$$\mu_k = x_k^T A x_k \rightarrow v_1^T A v_1 = \lambda_1$$

$\|v_1\| = 1$

EXAMPLE

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

We can compute by standard means (eg using the characteristic polynomial) that the eigenvalues λ_i are:

$$\lambda_i = \quad 3 \qquad 2 \qquad 1$$

and the associated (normalized) evectors are:

$$v_i = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

Now, lets recover λ_1, v_1 using the power method:

Let's choose as a starting vector:

$$x_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Then:

$$A x_0 = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

$$\Rightarrow \quad \|A x_0\| = \sqrt{5}.$$

$$\Rightarrow \quad x_1 = \frac{A x_0}{\|A x_0\|} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

Next:

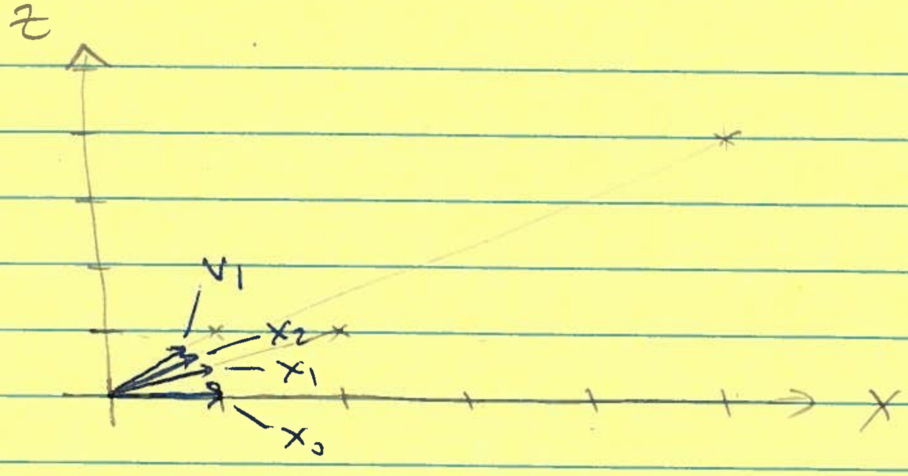
$$A^2 x_0 = A \cdot A x_0$$

$$= \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ 4 \end{bmatrix}$$

$$\Rightarrow \quad \|A^2 x_0\| = \sqrt{41}$$

$$\Rightarrow \quad x_2 = \frac{A^2 x_0}{\|A^2 x_0\|} = \frac{1}{\sqrt{41}} \begin{bmatrix} 5 \\ 0 \\ 4 \end{bmatrix}$$

In fact, this sequence $\{x_k\}$ is approaching v_1.



What about λ_1? can we recover that using the power method?

$$\mu_0 = x_0^T A x_0 = [1\ 0\ 0]\begin{bmatrix}2\\0\\1\end{bmatrix} = 2$$

$$\mu_1 = x_1^T A x_1 = \frac{1}{5}[2\ 0\ 1]\begin{bmatrix}2&0&1\\0&2&0\\1&0&2\end{bmatrix}\begin{bmatrix}2\\0\\1\end{bmatrix}$$

$$= \frac{1}{5}[2\ 0\ 1]\begin{bmatrix}5\\0\\4\end{bmatrix} = 2\tfrac{4}{5}.$$

$$\mu_2 = x_2^T A x_2 = \frac{1}{41}[5\ 0\ 4]\begin{bmatrix}2&0&1\\0&2&0\\1&0&2\end{bmatrix}\begin{bmatrix}5\\0\\4\end{bmatrix}$$

$$= \frac{1}{41}[5\ 0\ 4]\begin{bmatrix}14\\0\\13\end{bmatrix} = 2\tfrac{40}{41}$$

The sequence $\{\mu_k\}$ appears to be converging towards 3 which we know is the dominant eigenvalue