Power Method(s)

Monday, December 2, 2024 12:41 PM

Class 39: December 2, 2024

Recall: Last time, we discussed the implementation and performance of the QR iteration. We first concluded that, in its basal form, t his algorithm is abysmal:

- Cost per iteration is one QR decomposition, which is O(N^3) for general matrices.
- Convergence is linear, and the rate seems to depend on the ratios between successive eigenvalues. If the eigenvalues cluster at all, there are blocks of the iterate Ak that converge to T painfully slowly.

We then discussed 3 ideas involved in the acceleration and practical implementation of QR iteration. These address both compl aints above. First, we address cost per iteration:

- We can use Householder reflectors to find a matrix A0 which is unitarily similar to A, but which is "upper Hessenberg" (upper triangular except for one extra diagonal under). This is a O(N^3) precomputation step.
- If we use A0 to do QR iteration, then all iterates Ak are also Hessenberg, and the cost of a QR goes down to O(N^2).
- If A is symmetric, A0 is then tridiagonal. The cost of a QR (and QR iteration step) goes down to O(N)!

Then, we address rate of convergence and further optimize by:
1. "Deflating" parts of our matrix which are done (where the matrix is already upper triangular / diagonal).
2. Using "shifts" to quickly hone into a specific eigenvalue, then deflate.

This gives us an algorithm which converges cubically to each eigenvalue, which means it takes O(N) iterations to converge. Th is version of QR is incredibly fast.

Today, we take a step back and talk about the mother of all eigensolvers: Power Method(s). These are aimed at finding one spe cific eigenvector (and eigenvalue) of interest. However, as we will see, QR iteration and other modern eigensolvers (Lanczos, Arnoldi) can be interpreted as "soupe d up" power methods!

POWER METHOD:

GOAL - Find (approximate) the "dominant" eigenvector u_1 (for $\lambda_1 = \max\{|\lambda_k|\}$).

- Assumption 1: $A = PDP^{-1}$, basis of eigenvectors u_k w/ eigenvalues λ_k.
- Assumption 2: $|\lambda_1| > |\lambda_2| > |\lambda_3| > \ldots > |\lambda_N| \geq 0$ (distinct & ordered).

Give a random vector $\vec{b}$ (or guess for u_1).

$\hookrightarrow \vec{b} = \underbrace{\alpha_1 \vec{u}_1} + \alpha_2 \vec{u}_2 + \ldots + \alpha_N \vec{u}_N \quad (\alpha_1 \neq 0)$

$$A\vec{b} = \alpha_1 \lambda_1 u_1 + \alpha_2 \lambda_2 u_2 + \ldots + \alpha_N \lambda_N u_N.$$

$$A^2 b = \alpha_1 \lambda_1^2 u_1 + \alpha_2 \lambda_2^2 u_2 + \ldots$$

$$\vdots$$

$$\boxed{A^k b = \alpha_1 \lambda_1^k u_1 + \alpha_2 \lambda_2^k u_2 + \ldots}$$

Given b, A, $z_0 = b/\|b\|_\infty$

Given b, A, $z_0 = b/\|b\|_\infty$

while ($n \le n_{max}$ AND)

$y_k = A z_{k-1}$

$z_k = y_k / \|y_k\|_\infty$.

Use z_k $\longrightarrow$ $\boxed{\lambda_k}$

$$\frac{A^k z_0}{\|A^k z_0\|} = \frac{\lambda_1^k \left(\alpha_1 u_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k u_2 + \dots + \alpha_N \left(\frac{\lambda_N}{\lambda_1}\right)^k u_N \right)}{|\lambda_1|^k \| \quad '' \quad \|_\infty}$$

$$z_k \xrightarrow[k \to \infty]{} \pm \frac{u_1}{\|u_1\|_\infty} = \hat{u}_1$$

$$e_k = \| z_k - \hat{u}_1 \|_\infty, \quad \frac{e_{k+1}}{e_k} \xrightarrow[k \to \infty]{} \left| \left(\frac{\lambda_2}{\lambda_1} \right) \right|$$

$A z_k \approx \lambda_k z_k$, random vec $\vec{c}$,

$$c^T A z_k \approx \lambda_k (c^T z_k) \to \boxed{\lambda_k = \frac{c^T A z_k}{c^T z_k}}$$

Better:

Rayleigh quotient

$$\boxed{R_A(z) = \frac{z^T A z}{z^T z}} \longleftarrow$$

$$\lambda_k = R_A(z_k), \quad \frac{|\lambda_k - \lambda_1|}{|\lambda_{k-1} - \lambda_1|} \sim \left(\frac{\lambda_2}{\lambda_1} \right) \text{ (General)}$$

Symmetric

$(\lambda_2/\lambda_1)^2$

@ "Shifted Power Method"

$\mu \in \mathbb{R}$, PM $\to (A - \mu I)$

$\mu \in \mathbb{R}$, PM $\to (A - \mu I)$

$\text{eig}(A-\mu I) \to \{\lambda_1 - \mu, \lambda_2 - \mu, \dots, \lambda_N - \mu\}$

Ex: $\{0.5, 0.6, \dots, 1\}$ $\mu = 1$

$\to \{-0.5, -0.4, \dots, 0\}$

Easy: add μ at the end

Changes rate of convergence.

Inverse Power Method:

A invertible $\to A^{-1}$ $\left\{\frac{1}{|\lambda_N|} > \frac{1}{|\lambda_{N-1}|} > \dots > \frac{1}{|\lambda_1|}\right\} > 0$

$\to u_N$ smallest eigenvector (value).

"$y_k = A^{-1} z_{k-1}$" $\to$ SOLVE

Inverse Shifted PM.

$(A - \mu I)^{-1} \to \left\{\frac{1}{\lambda_1 - \mu}, \frac{1}{\lambda_2 - \mu}, \dots, \frac{1}{\lambda_N - \mu}\right\}$

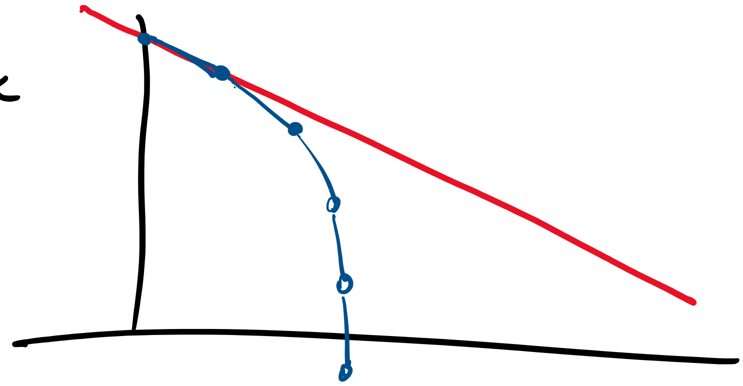
(*) Converge to (u_k, λ_k) for which μ is closest to λ_k.

Rate: $\dfrac{\frac{1}{|\lambda_j - \mu|}}{\frac{1}{|\lambda_k - \mu|}} = \dfrac{|\lambda_k - \mu|}{|\lambda_j - \mu|}$ $\begin{matrix} \to \text{closest} \\ \to \text{2nd closest} \end{matrix}$

Variable shift (Rayleigh shift) ISPM.

IDEA: μ_0 initial guess

Set $\mu_k = \lambda_k$



$\boxed{\text{SUPERLINEAR}}$!

A is non-symmetric

↳ quadratic

A is symmetric ⟶ cubic.

QR iteration ⟶ simultaneous orthogonal iter.