Computing eigenvalues and eigenvectors

Power iteration

Power iteration

Let $\mathbf{A} \in \mathbb{R}^{m \times m}$ has a full set of eigenvalues

$$\lambda_k, \quad k = 1, \ldots, m$$

Suppose that

$$|\lambda_1| > |\lambda_2| \geqslant |\lambda_3| \geqslant \cdots \geqslant |\lambda_m|$$

NB: $\lambda_1 \in \mathbb{R}$, otherwise λ_1^* is also an eigenvalue.

Power iteration

Only compute λ_1.

Iteration:

1. Take arbitrary $\mathbf{x}_0 \in \mathbb{R}^m$.
2. Iterate:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\,\mathbf{x}_k \\ \\ \lambda^{(k)} = \dfrac{\langle \mathbf{x}_k | \mathbf{x}_{k+1} \rangle}{\langle \mathbf{x}_k | \mathbf{x}_k \rangle} \quad \left(\equiv \dfrac{\langle \mathbf{x}_k | \mathbf{A} | \mathbf{x}_k | \rangle}{\langle \mathbf{x}_k | \mathbf{x}_k \rangle} \right), \end{cases} \qquad k = 0, 1, \cdots$$

Then, $\lambda^{(k)} \to \lambda_1, \quad k \to \infty.$

Power iteration: why does it work

Consider

$$\mathbf{x}_0 = c_1\,\mathbf{e}_1 + c_2\,\mathbf{e}_2 + \cdots + c_m\,\mathbf{e}_m$$

where $\{\mathbf{e}_k\}$ are normalized eigenvectors of $\mathbf{A}$:

$$\mathbf{A}\,\mathbf{e}_k = \lambda_k\,\mathbf{e}_k \qquad k = 1, \ldots, m$$

Then,

$$\mathbf{x}_k = \mathbf{A}^k\,\mathbf{x}_0 = \lambda_1^k\,c_1\,\mathbf{e}_1 + \lambda_2^k\,c_2\,\mathbf{e}_2 + \cdots + \lambda_m^k\,c_m\,\mathbf{e}_m$$

$$= \lambda_1^k \left[c_1\,\mathbf{e}_1 + \underbrace{\left(\frac{\lambda_2}{\lambda_1}\right)^k c_2\,\mathbf{e}_2 + \cdots + \left(\frac{\lambda_m}{\lambda_1}\right)^k c_m\,\mathbf{e}_m}_{\to\ 0 \text{ as } k\ \to\ \infty} \right]$$

Power iteration

Questions:

1. What if $c_1 = 0$ (i.e. $\mathbf{x}_0 \perp \mathbf{e}_1$)?

   numerical errors will generate $\neq 0$ components in the direction of $\mathbf{e}_1$

2. Numerical stability?

3. Rate of convergence?

   IOW, *a priori* error estimate?

4. *A posteriori* error estimate?

Stability

If $|\lambda_1| > 1$, $\lambda_1^k \to \infty$ overflow
$|\lambda_1| < 1$, $\lambda_1^k \to 0$ underflow

$\Rightarrow$ Need to keep $\mathbf{x}_k$ normalized: $\|\mathbf{x}_k\| = 1\,, \quad k = 0, 1, \ldots$

Modified iteration:

$$\begin{cases} \mathbf{y}_{k+1} = \mathbf{A}\,\mathbf{x}_k \\ \mathbf{x}_{k+1} = \dfrac{\mathbf{y}_{k+1}}{\|\mathbf{y}_{k+1}\|} \leftarrow \text{normalize} \\ \lambda^{(k)} = \langle \mathbf{x}_k | \mathbf{y}_{k+1} \rangle \end{cases}$$

$$k = 0, 1, \cdots$$

A priori error estimate

$$\frac{|\lambda^{(k)} - \lambda_1|}{|\lambda_1|} \leqslant \text{const} \times \left(\frac{\lambda_2}{\lambda_1}\right)^k$$

АКД §8.2

Convergence

$\lambda^{(k)}$ converges to λ_1.

Moreover, $\mathbf{x}_k$ converges to the eigenvector $\mathbf{e}_1$.

Convergence: $\sin \varphi_k \to 0\,, \qquad k \to \infty$

$$\cos \varphi_k = \frac{\langle \mathbf{x}_k | \mathbf{e}_1 \rangle}{\|\mathbf{x}_k\| \cdot \|\mathbf{e}_1\|}$$

A posteriori error estimate

There is no general solution.

For symmetric matrices:

Lemma: Let $\lambda_* \in \mathbb{R}$ and $\mathbf{x}_* \in \mathbb{R}^m$. Let $\mathbf{A} \in \mathbb{R}^{m \times m}$ is symmetric.
Then, $\exists$ an eigenvalue λ of $\mathbf{A}$, such that

$$|\lambda - \lambda_*| \leqslant \frac{\|\mathbf{A}\mathbf{x}_* - \lambda_* \mathbf{x}_*\|}{\|\mathbf{x}_*\|} .$$

A posteriori error estimate

Therefore, an *a posteriori* error bound for the power iteration for symmetric matrices:

$$|\lambda_1 - \lambda^{(k)}| \leqslant \frac{\|\mathbf{y}_k - \lambda^{(k)}\mathbf{x}_{k-1}\|}{\|\mathbf{x}_{k-1}\|}$$

i.e. stop the iteration when the r.h.s is smaller then the predefined tolerance.

The rate of convergence; shifts

Convergence is slow is the spectral gap is small, $|\lambda_2 - \lambda_1| \ll \lambda_1$, since then $\left|\frac{\lambda_2}{\lambda_1}\right| \approx 1$.

Shifts

Consider

$$\mathbf{B} = \mathbf{A} - \sigma\widehat{\mathbf{1}}$$

with some $\sigma \in \mathbb{R}$. eigvals(B) = eigvals$(A) - \sigma$

Use the power iteration for $\mathbf{B}$. Convergence is controlled by $\lambda_k - \sigma$. If $\lambda_1 - \sigma$ is the max eigenvalue of $\mathbf{B}$, then select σ such that

$$\max_{k=2,\cdots,m}\left|\frac{\lambda_k - \sigma}{\lambda_1 - \sigma}\right| \Rightarrow \min$$

Krylov subspaces

Krylov subspaces

$$\mathcal{K}_r(\mathbf{A}, \mathbf{x}_0) = \text{span}\left(\mathbf{x}_0, \mathbf{A}\mathbf{x}_0, \mathbf{A}^2\mathbf{x}_0, \cdots, \mathbf{A}^{r-1}\mathbf{x}_0\right)$$

Power iteration generates a sequence of *Krylov subspaces*

$$\mathcal{K}_1 \subset \mathcal{K}_2 \subset \cdots \subset \mathbb{R}^m$$

Note that also

$$\mathbf{A}\mathcal{K}_r \subset \mathcal{K}_{r+1}$$

Krylov subspace iterations

Power iteration generated vectors,

$$\mathbf{A}^s \mathbf{x}_0$$

are progressively closer to being linear dependent with $s = 1, 2, \cdots$.

Krylov subspace methods: form a basis of $\mathcal{K}_r$ by reorthogonalization.

Inverse iteration

Inverse iteration

Suppose we know $\lambda_* \approx \lambda_s$, a good approximation to λ_s. How to compute $\mathbf{e}_s$, its eigenvector?

Idea 1

By definition, $\mathbf{A}\mathbf{e}_s = \lambda_s \mathbf{e}_s$, so try looking for a non-trivial solution of

$$(\mathbf{A} - \lambda_* \widehat{\mathbf{1}})\, \mathbf{x} = 0$$

Since $\lambda_* \neq \lambda_s$, $\mathbf{A} - \lambda_* \widehat{\mathbf{1}}$ is non-degenerate $\Rightarrow$ there is only the trivial solution $\mathbf{x} = 0$.

Inverse iteration

The algorithm:

$$(\mathbf{A} - \lambda_* \widehat{\mathbf{1}})\, \mathbf{y}_{k+1} = \mathbf{x}_k \tag{1}$$
$$\mathbf{x}_{k+1} = \frac{\mathbf{y}_{k+1}}{\|\mathbf{y}_{k+1}\|}$$

for $k = 0, 1, \ldots$ and some initial condition $\mathbf{x}_0$.

Never invert a matrix explicitly!

Inverse iteration

Formally:

$$\mathbf{y}_{k+1} = (\mathbf{A} - \lambda_* \widehat{\mathbf{1}})^{-1} \, \mathbf{x}_k$$

$\Rightarrow$ inverse iteration is nothing but a power iteration with the matrix $(\mathbf{A} - \lambda_* \widehat{\mathbf{1}})^{-1}$

How and why does it work?

Inverse iteration

Consider symmetric $\mathbf{A}$. Let $\{\mathbf{e}_\alpha\}, \quad \alpha = 1, \ldots, m$ is an orthonormal eigensystem of $\mathbf{A}$.

$$\mathbf{x}_0 = \sum_{\alpha=1}^{m} c_\alpha \mathbf{e}_\alpha; \quad \mathbf{y}_1 = \sum_{\alpha=1}^{m} b_\alpha \mathbf{e}_\alpha$$

The iteration (1) becomes:

$$\sum_{\alpha=1}^{m} (\lambda_\alpha - \lambda_*) b_\alpha \mathbf{e}_\alpha = \sum_{\alpha=1}^{m} c_\alpha \mathbf{e}_\alpha$$

Due to orthonormality of $\mathbf{e}_\alpha \Rightarrow$

$$b_\alpha = \frac{c_\alpha}{\lambda_\alpha - \lambda_*}$$

Inverse iteration

Therefore,

$$\mathbf{y}_1 = \sum_{\alpha=1}^{m} \frac{c_\alpha}{\lambda_\alpha - \lambda_*}\mathbf{e}_\alpha = \frac{1}{\lambda_s - \lambda_*}\left[c_s\mathbf{e}_s + \underbrace{\sum_{\alpha\neq s} \frac{\lambda_s - \lambda_*}{\lambda_\alpha - \lambda_*}\mathbf{e}_\alpha}\right]$$

"small" if $\quad |\lambda_s - \lambda_*| \ll |\lambda_\alpha - \lambda_*|, \quad \alpha \neq s$

Each iteration multiplies $\mathbf{e}_{\alpha\neq s}$ by

$$\frac{\lambda_s - \lambda_*}{\lambda_\alpha - \lambda_*}\ .$$

Modified inverse iteration

Inverse iteration requires a good approximation to the eigenvalue.

Try using the Rayleigh ratio

$$\lambda^{(k)} = \langle \mathbf{x}_k | \mathbf{A}\mathbf{x}_k \rangle \, ,$$

$$\left(\mathbf{A} - \lambda^{(k)} \widehat{\mathbf{1}} \right) \mathbf{y}_{k+1} = \mathbf{x}_k \, ,$$

$$\mathbf{x}_{k+1} = \frac{\mathbf{y}_{k+1}}{\|\mathbf{y}_{k+1}\|}$$

for $k = 0, 1, \ldots$ and $\|\mathbf{x}_0\| = 1$

Modified inverse iteration

Modified inverse iteration converges fast if the angle $\widehat{(\mathbf{x}_0, \mathbf{e}_s)}$ is small.

How to get $\mathbf{x}_0 \approx \mathbf{e}_s$? E.g. a couple of power iterations.