§8.2 Power Iteration ⇒ "chap06_eigenvalue.key"

Numerical experiments

[Example 8.2.1 magicpower]

```
>> A = magic(5)/65
>> for j=1:8; x= A*x ; x= x/norm(x); end
```

For any x, x converges to $[1,1,1,1,1]^T$ WHY?

Analysis (of Power iteration)

Dominant Eigenvector

Suppose the eigenpairs of A are (λ_i, v_i) and

$$\underbrace{|\lambda_1|}_{} > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$$

↳ dominant eigenvalue

$\forall x$, we have $x = \sum c_j x_j$

Then
$$\begin{aligned} Ax &= c_1 Av_1 + c_2 Av_2 + \cdots + c_n Av_n \\ &= c_1\lambda_1 v_1 + c_2\lambda_2 v_2 + \cdots + c_n \lambda_n v_n \end{aligned}$$

$$\Rightarrow A^k x = c_1 \lambda_1^k v_1 + c_2\lambda_2^k v_2 + \cdots + c_n \lambda_n^k v_n$$

① $$= \boxed{\lambda_1^k \left[c_1 v_1 \right.} + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k v_n \Big]$$

②三角不等式

$$\Rightarrow \left\| \frac{A^k x}{\lambda_1^k} - c_1 v_1 \right\| \leq |c_2| \left|\frac{\lambda_2}{\lambda_1}\right|^k \|v_2\| + \cdots + |c_n| \left|\frac{\lambda_n}{\lambda_1}\right|^k \|v_n\|$$

$\to 0$ $\to 0$

$$\Rightarrow \boxed{A^k x \parallel v_1} !$$

(Note: $A(A(\cdots(Ax)))$ not $A^k x$)

Known already. But λ_1^k can make $\|A^k x\|$ very big or very small $\left(\because \left\| \frac{A^k x}{\lambda_1^k} - \underbrace{c_1 v_1}_{\text{fixed}} \right\| \to 0\right)$

Solution: <u>normalization</u>

Given $x^{(0)} \in \mathbb{C}^n$ and let $\boxed{y^{(0)} = \frac{x^{(0)}}{\|x^{(0)}\|}}$

for $i = 1, 2, \cdots$ until converges

$$x^{(k)} = A y^{(k-1)}$$

$$\boxed{y^{(k)} = \frac{x^{(k)}}{\|x^{(k)}\|}}$$

normalized eigenvector

$$\lambda^{(k)} = [y^{(k)}]^* A y^{(k)}$$

eigenvalue by Rayleigh quotient

Note: (1) As $y^{(k)}$ converges to the dominant eigenvector

$$A y^{(k)} \approx \lambda y^{(k)} \Rightarrow (y^{(k)})^* A y^{(k)} \approx \lambda \underbrace{(y^{(k)})^* y^{(k)}}_{=1}$$

(2) In short, $y^{(k)} = \beta^{(k)} A^k y^{(0)} = \prod_{i=1}^{k} \frac{1}{\|x^{(i)}\|} A^k y^{(0)}$

§8.4 Krylov Subspace

- Motivation:

  In Power Method, we have $u, Au, A^2u, \cdots$ but use only $\underbrace{\text{the latest vector } A^k u}$ 一根釣竿

- Idea:

  Use $\underbrace{\text{linear combinations of } \{u, Au, A^2u, \cdots, A^{m-1}u\}}$ (m-dim)

  — 一張低維度的網子

  — Needs only mtx-vec mult.

- Krylov matrix $K_m \subset \mathbb{C}^{n\times m}$

$$K_m = \begin{array}{cc} & m \\ n & \begin{bmatrix} u & Au & \cdots & A^{m-1}u \end{bmatrix} \end{array}$$

$A: n\times n$

$K_m: n\times m$

- Krylov Subspace $\mathcal{K}_m \in \mathbb{C}^n$

  Range of K_m

  $\parallel$

  Column space of K_m

$$\begin{array}{ccc} K_m & w & K_m w \\ (n\times m) \times (m\times 1) & \Rightarrow & (n\times 1) \end{array}$$

- Notes:
  - K_m and $\mathcal{K}_m$ depends on A & u, but we usually denote m (the dimension of $\mathcal{K}$ & rank of K_m) only. Rather than $K_{m,A,u}$ or $\mathcal{K}_{m,A,u}$

Lemma 8.4.1

dimension reduction

$A: n\times n$, $0 < \boxed{m < n}$

係數向量

If $x \in \mathcal{K}_m$, then (i) $\exists\, z \in \mathbb{C}^m$, s.t. $x = K_m z$

(ii) $x \in \mathcal{K}_{m+1}$

(iii) $Ax \in \mathcal{K}_{m+1}$

Pf: (i) If $x \in \mathcal{K}_m$, $\exists\, c_1, \cdots, c_m$ such that

$$x = c_1 u + c_2 Au + \cdots + c_m A^{m-1} u \quad \text{———} \quad (+)$$

Take $z = \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix}$

(ii) $x = c_1(u) + c_2(Au) + \cdots + c_m A^{m-1} u + 0 A^m u \in \mathcal{K}_{m+1}$

(iii) $A(+) \Rightarrow \underbrace{Ax}_{\substack{\cap \\ \mathcal{K}_{m+1}}} = \underline{0}\, u + \underline{c_1} Au + \cdots + \underline{c_m} A^m u$

Dimension Reduction by Krylov Subspace Approach

Replace the full n-dimensional space with a much lower-dimensional $\mathcal{K}_m$ for $m \ll n$.

Krylov subspace

$$Ax = b$$

$\Leftrightarrow$ $\min\limits_{x \in \mathbb{C}^n} \| Ax - b \|$ $\qquad x \in \mathbb{C}^n$

$\Rightarrow$ $\min\limits_{x \in \mathcal{K}_m} \| Ax - b \|$ $\qquad x \in \mathcal{K}_m,\ m \ll n$

放鬆成 x 在低維度空間 $(\mathcal{K}_m)$ 的最佳逼近解

$\Leftrightarrow$ $\min\limits_{z \in \mathbb{C}^m} \| A(K_m z) - b \|$ $\qquad z \in \mathbb{C}^m,\ m \ll n$

$\forall x \in \mathcal{K}_m,\ x = K_m z$ (Lemma 8.41)

$\rightsquigarrow$ Least square problem!

Note: In $K_m = [b, Ab, A^2 b, \cdots, A^{m-1} b]$, we

(Take b as the seed vector)

compute $A^k b = \underbrace{(A(A \cdots \underbrace{(A \underbrace{(A \cdot b)})} \cdots))}$

$\hookrightarrow$ All sparse mtx vector mult. $A \cdot u$

Not $\underbrace{(A \cdot A \cdots A)}_{\text{Dense! even } A \text{ is sparse}} \cdot b$.

Example 8.4.1 (p. 333, krylovunstable)

The Arnoldi Iteration

The Arnoldi iteration finds an orthonormal basis for a Krylov subspace

Issues:

$$K_m = [u \ \ Au \ \ A^2u \ \cdots \ A^{m-1}u]$$

$\longrightarrow$ approach to the dominant eigenvector

愈來愈"平行"於同一向量

逐漸失去線性獨立，愈來愈線性相依

$\Rightarrow$ Numerical cancellation!

K_m 的 condition number $\uparrow$ as $m \uparrow$

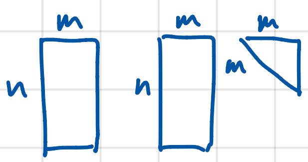
Ideas:

Find an orthonormal basis

Approach:

Conceptually, we can perform skinny QR on K_m to get

$$K_m = Q_m R_m = \underbrace{[q_1 \ q_2 \cdots q_m]}_{\text{orthonormal basis of } K_m} \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1m} \\ & R_{22} & \cdots & R_{2m} \\ & & \ddots & \vdots \\ & & & R_{mm} \end{bmatrix}$$

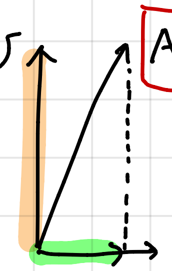


Example : To find $\mathcal{K}_3$

Given u

$$q_1 = \frac{u}{\|u\|}$$

$$v = \boxed{Aq_1} - (q_1^* \boxed{Aq_1}) \cdot q_1$$

① 

v, $\boxed{Aq_1} \in \mathcal{K}_2$, $q_1 \in \mathcal{K}_1$

② $Aq_1 = \times q_1 + \times q_2$

$$q_2 = \frac{v}{\|v\|}$$

$$v = \boxed{Aq_2} - (q_1^* \boxed{Aq_2})\, q_1 - (q_2^* \boxed{Aq_2}) \cdot q_2$$

$$(\because Aq_2 = \times q_1 + \times q_2 + \times q_3)$$

$$q_3 = \frac{v}{\|v\|}$$

$q_m \in \mathcal{K}_m$

$\Rightarrow A q_m \in \mathcal{K}_{m+1}$ (by lemma 8.4.1)

$\Rightarrow \exists\, H_{ij}$

s.t. $\boxed{Aq_m} = \underbrace{H_{1m}} q_1 + \underbrace{H_{2m}} q_2 + \cdots + \underbrace{H_{mm}} q_m + \boxed{\underbrace{H_{m+1,m}\, q_{m+1}}_{\text{unknown}}}$

新的向量，用來擴展高一維的 Krylov subspace ($\mathcal{K}_{m+1}$)

$$q_i^*(Aq_m) = H_{im}$$

$$(\because q_1 \perp q_2 \perp \cdots \perp q_m)$$

unknown

need $\|q_{m+1}\| = 1$

General case

To construct $K_m = [u, Au, A^2u, \cdots, A^m u]$

$= \text{span}\{q_1, q_2, \cdots, q_m\}$

(1) let $q_1 = \dfrac{u}{\|u\|}$

(2) For $m = 1, 2, \cdots$

(i) Let $H_{im} = q_i^* A q_m$ for $i = 1:m$

(ii) let $v = Aq_m - H_{1m}q_1 - \cdots - H_{mm}q_m$

(iii) Let $H_{m+1,m} = \|v\|$

(iv) let $q_{m+1} = \dfrac{v}{H_{m+1,m}}$

normalization

---

check orthonormal : $Q = [q_1\ q_2\ q_3]$

$\Rightarrow$ compute $\|Q^T Q - I_3\|$

check span : $K = [u, Au, A^2u]$

$\Rightarrow$ compute rank $([Q\ K])$

Check: $q_{m+1} \perp q_i$ for $i = 1:m$

$$q_i^* q_{m+1} = q_i^* \frac{1}{\|v\|}\left(Aq_m - \cancelto{0}{H_{1m}q_1} - \cdots - (q_i^* A q_m)q_i - \cdots - \cancelto{0}{H_{mm}q_m}\right)$$

$$= \frac{1}{\|v\|}\left[q_i^* A q_m - (q_i^* A q_m)\,\cancelto{=1}{q_i^* q_i}\right] = 0$$

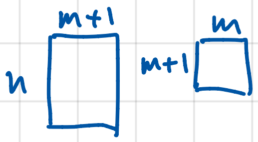
Fundamental Identity of Krylov Subspace Methods

$$A Q_m = [A q_1 \; A q_2 \; \cdots \; A q_m]$$

$$= [q_1 \; q_2 \cdots q_{m+1}] \begin{bmatrix} H_{11} & H_{12} & \cdots & H_{1m} \\ H_{21} & H_{22} & \cdots & H_{2m} \\ & H_{32} & \ddots & \vdots \\ & & \ddots & H_{mm} \\ & & & H_{m+1,m} \end{bmatrix}$$

H_m is an upper Hessenberg

$$= Q_{m+1} H_m$$



Exercise 8.4.7 on p.338

- How to solve the eigenvalue problem $Ax = \lambda x$ by using the Arnoldi iteration?
- That is, how to approximate the eigenvalue problem $Ax = \lambda x$ over $\mathcal{K}_m$?

§8.5 GMRES

The Arnoldi iteration can be used to solve $Ax=b$

$$Ax=b$$

$$\Rightarrow \min_{x\in\mathbb{C}^n} \|Ax-b\|$$

$$\approx \min_{x\in K_m} \|Ax-b\|$$

$$= \min_{z\in\mathbb{C}^m} \|A\boxed{K_m}z-b\|$$

ill-conditioned as $m\uparrow$

don't use $x=K_m z$

use $x=\underline{\underline{Q_m}}z$ (orthonormal basis)

$$\Rightarrow \min_{z\in\mathbb{C}^m} \|A\boxed{Q_m}z-b\|$$

by the key identity

$$\Rightarrow \min_{z\in\mathbb{C}^m} \|Q_{m+1}H_m z-b\|$$

$q_1=\frac{b}{\|b\|} \Rightarrow b=\|b\|\cdot q_1$

$=\|b\|\cdot[q_1\ q_2\cdots q_{m+1}]$

$=\|b\|\cdot Q_{m+1}\cdot e_1$

$$\Rightarrow \min_{z\in\mathbb{C}^m} \|Q_{m+1}\underbrace{(H_m z-\|b\|e_1)}_{\overset{\shortparallel}{w}}\|$$

least square problem $(n\times m)$

$\forall w\in\mathbb{C}^{m+1}$, we have

$\|Q_{m+1}w\|=(w^*Q_{m+1}^*Q_{m+1}w)^{1/2}$

$=(w^*w)^{1/2}=\|w\|$

$$\Rightarrow \min_{z\in\mathbb{C}^m} \|H_m z-\|b\|e_1\|$$

upper Hessenberg and smaller $(m+1\times m)$ size! (No n, where $m\ll n$)

$$\Rightarrow z_m=\arg\min \|H_m z-\|b\|e_1\|$$

$\Rightarrow x_m=Q_m z_m$ is the mth approximation solution of $Ax=b$

Restarting

Issue:

Entries in $H_m \uparrow$, dimension of Krylov subspace $\uparrow$

columns in $Q_m \uparrow$, computation & storage $\uparrow$

As $m \uparrow$

Solution:

Restarting.

Let $\underset{\substack{\uparrow \\ \text{True}}}{x} = \underset{\substack{\uparrow \\ \text{Approx.}}}{\hat{x}} + u.$ $\Rightarrow$ $A(\hat{x} + u) = b$ $\Rightarrow$ $Au = b - A\hat{x} = r$

$\Rightarrow$ Solve $Au = r$ to get a "correction"

However.

- Restarting preserves progress made in previous iteration,
- the Krylov space information is discarded and
- the residual minimization process starts again over lower-dimensional choices
- Which can retard or even stagnate the convergence
  减慢 停滞

Example 8.5.1 (p. 339, krylovstable)

§ 8.7 Matrix-free Iterations

Linear transformation

- $\begin{cases} f(x+y) = f(x) + f(y) \\ f(\alpha x) = \alpha f(x) \end{cases}$

- Every linear transformation between finite dimensional vector spaces can be represented as a matrix vector multiplication.

- By computing $f(x)$, we can do many things.

① 斜率 $\dfrac{f(x_2) - f(x_1)}{x_2 - x_1}$ in the secant method

root-finding $f(x) = 0$ (or find $f^{-1}(0)$)

② Matrix-vector mult. $f(x) = Ax$

GMRES, MINRES, CG for solving $Ax = b$

Krylov sub-space methods can be used to invert a linear transformation if one provides code for the transformation, even if its associated matrix is not known explicitly.

Blurring images (e.g. 柔焦 in portrait photos)

$$B_{ij} = \begin{cases} 1/2 & \text{if } i=j \\ 1/4 & \text{if } |i-j|=1 \\ 0 & \text{others} \end{cases} \qquad \underset{m\times m}{B} = \begin{bmatrix} 1/2 & 1/4 & & & \\ 1/4 & 1/2 & 1/4 & & \\ & \ddots & \ddots & \ddots & 1/4 \\ & & & 1/4 & 1/2 \end{bmatrix}$$

$$\underset{(m\times m)}{B} \cdot \underbrace{\underset{(m\times n)}{X}}_{\text{image}} = \underbrace{\begin{bmatrix} & & \\ & B & \\ & & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_m \end{bmatrix}}$$

→ To blur the image vertically

$\Rightarrow B^k X$ To increase the amount of blur

$\Rightarrow (\underset{(n\times n)}{C}\ \underset{(n\times m)}{X^T})^T = XC^T = XC$ To blur "horizontally"

B和C結構相同，大小不同

$\Rightarrow \text{blur}(X) = \underset{(m\times m)}{B^k}\ \underset{(m\times n)}{X}\ \underset{(n\times n)}{C^k}$

Example 8.7.1

1st row
$$\begin{pmatrix} \frac{1}{2} & \frac{1}{4} & & \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & \\ & \frac{1}{4} & \frac{1}{2} & \end{pmatrix} \begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

2nd row
$$\begin{pmatrix} \frac{1}{2} & \frac{1}{4} & & \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & \\ & \frac{1}{4} & \frac{1}{2} & \end{pmatrix} \begin{bmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

$= \frac{1}{4}x_1 + \frac{1}{2}x_2 + \frac{1}{4}x_3$

Deblurring

To invert blur(X). But **how**?

Idea:

(1) $\text{blur}(X) = B^k X C^k$ is a linear transformation

(2) Let $\text{vec}(X) = x$ and $\text{unvec}(x) = X$ $\qquad \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \underset{\text{unvec}}{\overset{\text{vec}}{\rightleftarrows}} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

$$\Rightarrow \underbrace{A}_{mn \times mn} \underbrace{\text{vec}(X)}_{mn \times 1} = \text{vec}(Z) \Rightarrow Ax = z$$

For an 1024×768 image, $mn = 786432$ and A has $786432^2 = 618{,}475{,}290{,}624$ entries!

$\Rightarrow$ Consume a lot of memory & almost impossible to invert

(3) $$\boxed{\underset{\text{original image}}{X} \underset{\text{unvec}}{\overset{\text{vec}}{\rightleftarrows}} x \underset{\text{Deblur, } A^{-1}z = x}{\overset{\text{Blur, } Ax = z}{\rightleftarrows}} z \underset{\text{vec}}{\overset{\text{unvec}}{\rightleftarrows}} \underset{\text{blurred image}}{Z}}$$

$$\forall u, \underbrace{A}_{(mn \times mn)} \underbrace{u}_{mn \times 1} \equiv \underbrace{\text{vec}\Big(\underbrace{\underbrace{B^k}_{m \times m} \underbrace{(\text{unvec}(\underbrace{u}_{mn\times 1}))}_{m \times n} \underbrace{C^k}_{n \times n}}_{m \times n} \Big)}_{mn \times 1}$$

Then we can solve $Ax = z$ by, e.g. GMRES

Note that there is no need to form B to do Bx

$$B \cdot x = \begin{bmatrix} & & & \\ & 1/4 & 1/2 & 1/4 \\ & & & \end{bmatrix} \begin{bmatrix} \vdots \\ x_{i-1} \\ x_i \\ x_{i+1} \\ \vdots \end{bmatrix} = \underbrace{\begin{bmatrix} \vdots \\ \frac{1}{4}x_{i-1} + \frac{1}{2}x_i + \frac{1}{4}x_{i+1} \\ \vdots \end{bmatrix}}_{\text{Computable without forming } B \text{ explicitly}}$$

Example 8.7.1 (p. 350, blurimage)

Example 8.7.2 (p. 351, deblurimage)