DS284: Numerical Linear Algebra − Assignment 3

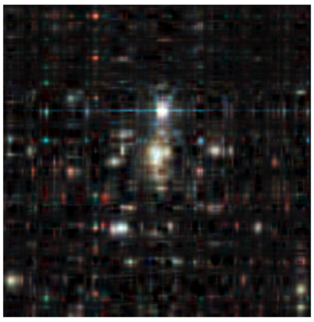
Aneesh Panchal
06-18-01-10-12-24-1-25223
Indian Institute of Science (IISc), Bangalore, IN

SEP 2024

Solution 2

Solution 2 (a)

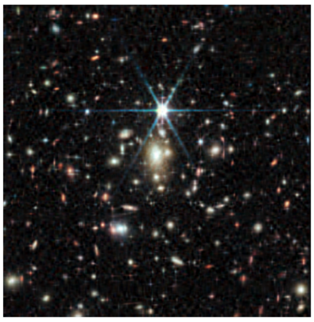
Reconstructed Image with $n(\sigma)$ =10



Reconstructed Image with $n(\sigma)$ =20

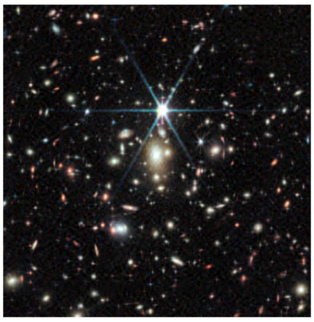
Reconstructed Image with $n(\sigma)$ =30

Reconstructed Image with $n(\sigma)$ =40



Reconstructed Image with $n(\sigma)$ =50

Reconstructed Image with $n(\sigma)$ =60



Reconstructed Image with $n(\sigma)$ =70

Reconstructed Image with $n(\sigma)$ =80

Reconstructed Image with $n(\sigma)$ =90

Reconstructed Image with $n(\sigma)$ =100

Original Image

According to my observations, minimum of $\mathbf{100}$ singular values are required to approximate the image, such that it looks indistinguishable from the original image (according to naked eye view).

Solution 2 (b)

Given size of image is $2000 \times 1968 \times 3$. From part (a), we assumed number of singular values required are $k = 100$.
Dim(Red colour matrix) = Dim(Green colour matrix) = Dim(Blue colour matrix) = 2000×1968.
Final Number of values to be Transmitted $=$ ((Entries in U) $+ k +$ (Entries in V^T)) $\times 3$

$$((2000 \times 100) + 100 + (1968 \times 100)) \times 3 = \mathbf{1190700}$$

But if we pass the full image then Number of values to be transmitted $= 1968 \times 2000 \times 3 = 11808000$. If we use only 100 singular values then only 1190700 are required to pass for reconstruction of image.

Solution 2 (c)

From part (a), we have considered the number of singular values to be $k = 100$.
For Red colour matrix,
$\|R - R_{100}\|_2 = 7.420379$, and $\sigma_{101} = 7.420378928710866$
$\|R - R_{100}\|_F = 66.487118$, and $\sqrt{\sigma_{101}^2 + \sigma_{102}^2 + \cdots + \sigma_{1968}^2} = 66.48711801948993$

For Blue colour matrix,
$\|B - B_{100}\|_2 = 7.021533$, and $\sigma_{101} = 7.021533289088522$
$\|B - B_{100}\|_F = 63.113877$, and $\sqrt{\sigma_{101}^2 + \sigma_{102}^2 + \cdots + \sigma_{1968}^2} = 63.113877142893934$

For Green colour matrix,
$\|G - G_{100}\|_2 = 6.794495$, and $\sigma_{101} = 6.794494660341246$
$\|G - G_{100}\|_F = 54.661232$, and $\sqrt{\sigma_{101}^2 + \sigma_{102}^2 + \cdots + \sigma_{1968}^2} = 54.66123216977876$

From above it is easy to see that the given theorem in the question hold for these errors.

Solution 3 (a) Given symmetric matrix $A \in \mathbb{R}^{m \times m}$

As A is symmetric, $\exists$ orthogonal matrix Q s.t. $A = Q \Lambda Q^T$

$$A = Q\,|\Lambda|\,\underbrace{\text{sign}(\Lambda)\, Q^T}_{\text{assume this as } V^T}$$

$$V^T V = \text{sign}(\Lambda)\, Q^T (\text{sign}(\Lambda)\, Q^T)^T$$

$$= \text{sign}(\Lambda)\, \underbrace{Q^T Q}_{I} (\text{sign}(\Lambda))^T$$

$$= \text{sign}(\Lambda)\, I\, (\text{sign}(\Lambda))^T$$

$$= \text{sign}(\Lambda)(\text{sign}(\Lambda))^T = I$$

Hence, we get $A = Q\,|\Lambda|\,V^T$

ie. singular values of A are abs. values of eig. values of A

Vector induced matrix norm $\|A\|_2$ is max. abs. eig. value of A

$\|A\|_2 = \|Q\,|\Lambda|\,V^T\|_2 \leq \|Q\|_2\, \||\Lambda|\|_2\, \|V^T\|_2 = \||\Lambda|\|_2$

as $|\Lambda|$ is diagonal matrix with entries as $|\lambda_1|, |\lambda_2|, \ldots$

$\|A\|_2 \leq \||\Lambda|\|_2 \leq |\lambda_{max}|$

Hence, proved.

Solution 3 (b) Cauchy Schwartz inequality $= |u^T v| \leq \|u\|_2 \|v\|_2$

$$|x^T A x| \leq \|x\|_2 \|Ax\|_2$$

$$\leq \|x\|_2 \|A\|_2 \|x\|_2$$

As given $x \in \mathbb{R}^m$ is a unit vector ie. $\|x\|_2 = 1$

$$|x^T A x| \leq \|A\|_2$$

Hence, proved.

Aneesh Panchal

06-18-01-10-12-24-1-25223

Solution 3 (c) $\tilde{A}\tilde{u} = \tilde{\lambda}\tilde{u}$

$(A+\delta A)(u+\delta u) = (\lambda+\delta\lambda)(u+\delta u)$

$\cancel{Au} + A\delta u + \delta A u + \cancel{\delta A \delta u} = \cancel{\lambda u} + \lambda \delta u + \delta\lambda u + \cancel{\delta\lambda\delta u}$

as given $Au = \lambda u$ & assume $\delta A \delta u, \delta\lambda\delta u \to 0$

$\delta\lambda u = (A-\lambda I)\delta u + \delta A u$

As A is symmetric $\Rightarrow$ $A-\lambda I$ is also symmetric

premultiply by u^T we get,

$\delta\lambda u^T u = ((A-\lambda I)u)^T \delta u + u^T \delta A u$

as $(A-\lambda I)u = 0$ we get,

$\delta\lambda u^T u = u^T \delta A u$

$|\delta\lambda| \, \|u^T u\| = \|u^T \delta A u\|$

$$|\delta\lambda| = \frac{\|u^T\delta A u\|}{\|u^T u\|} \leq \max_{\|u\|_2=1} \frac{\|u^T \delta A u\|_2}{\|u^T u\|_2} = \|\delta A\|_2$$

$|\delta\lambda| \leq \|\delta A\|_2$

Hence, proved.

Solution 3 (d) Relative condition no., $\hat{K}^R = \max\limits_{\|\delta A\|_2} \dfrac{\frac{|\delta\lambda|}{|\lambda|}}{\frac{\|\delta A\|_2}{\|A\|_2}}$

$$\hat{K}^R = \max_{\|\delta A\|_2} \frac{|\delta\lambda|}{|\lambda|} \frac{\|A\|_2}{\|\delta A\|_2} \leq \frac{\|A\|_2}{|\lambda|}$$ (as given in previous part $|\delta\lambda| \leq \|\delta A\|_2$)

Solution 3 (e) $\hat{K}^R \leq \dfrac{\|A\|_2}{|\lambda_{min}|}$ for given matrix $A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} = M$

$\|A\|_2 = |\lambda_{max}| = |a|$

$|\lambda_{min}| = |a|$

$\therefore$ we get $\hat{K}^R = 1$

which is required relative condition no.

Aneesh Panchal

06-18-01-10-12-24-1-25223

Solution 3 (f) Given algorithm S is backward stable

$$\tilde{f}(x) = f(\tilde{x}) \quad \text{for} \quad \frac{\|x-\tilde{x}\|}{\|x\|} = O(\varepsilon_M)$$

$$K(x) = \max_{\delta x} \frac{\|\delta f\|}{\|f\|}\,\frac{\|x\|}{\|\delta x\|} \Rightarrow \left(\frac{\|f(\tilde{x}) - f(x)\|}{\|f(x)\|}\right)\left(\frac{\|\tilde{x} - x\|}{\|x\|}\right)^{-1} \leq K(x)$$

relative forward error $= \dfrac{\|f(\tilde{x}) - f(x)\|}{\|f(x)\|} \leq K(x)\dfrac{\|\tilde{x} - x\|}{\|x\|} \approx O(K(x)\,\varepsilon_M)$

$\therefore$ for given algorithm S & i/p M,

relative forward error $\approx O(K(M)\,\varepsilon_M) = O(\varepsilon_M)$

ie. order of machine epsilon.

Solution 3 (g) Given all relative errors are machine epsilon ie. ε_M

a does not have floating pt. error.

Assume $\alpha = 1 \pm \varepsilon_M$

$$\left|\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right| = \begin{vmatrix} (a-\lambda)\alpha & \varepsilon_M \\ \varepsilon_M & (a-\lambda)\alpha \end{vmatrix}$$

$$= (a-\lambda)^2 \alpha^3 \ominus \varepsilon_M^2 \alpha$$

$$= (a-\lambda)^2 \alpha^4 - \varepsilon_M^2 \alpha^2$$

Now assume $\varepsilon_M^2 \to 0$ we get,

$p_M(\lambda) = \det(M - \lambda I) = (a-\lambda)^2 (1 \pm 4\varepsilon_M)$

for $p_M(\lambda) = 0$ we get,

$(1 \pm 4\varepsilon_M)\lambda^2 - \lambda\, 2a(1 \pm 4\varepsilon_M) + a^2(1 \pm 4\varepsilon_M) = 0$

as we know,

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \left[\frac{2a(1 \pm 4\varepsilon_M) \pm \sqrt{4a^2(1 \pm 4\varepsilon_M)^2\alpha \ominus 4a^2(1 \pm 4\varepsilon_M)^2\alpha^2}}{2(1 \pm 4\varepsilon_M)\,\alpha}\right]\alpha^2$$

$$= a\left[\frac{\cancel{2(1 \pm 4\varepsilon_M)}\left[1 \pm \sqrt{\alpha^2 - \alpha^3}\right]}{\cancel{2(1 \pm 4\varepsilon_M)}\,\cancel{\alpha}}\right]\alpha^2 \Rightarrow \text{now using assumption}$$

$\alpha = 1 \pm \varepsilon_M$ & put value we get

$$= a(1 \pm \varepsilon_M)(1 \pm \sqrt{\varepsilon_M})$$

$\therefore$ Floating pt. approximation of λ is $a(1 \pm \varepsilon_M)(1 \pm \sqrt{\varepsilon_M})$

As floating pt. approximation of λ is $a(1\pm\varepsilon_M)(1\pm\sqrt{\varepsilon_M})$

ie. $\lambda = a(1\pm\varepsilon_M\pm\sqrt{\varepsilon_M})$ (neglecting the term $\varepsilon_M^{3/2}\to 0$)

$\therefore$ Floating pt. approximation error,

$|\lambda - a| = \pm a(\varepsilon_M \pm \sqrt{\varepsilon_M})$

λ ↳ actual eig. value

Solution 3 (h) Forward Relative error $= \dfrac{|\lambda - a|}{a}$ ie. $\dfrac{|\tilde{f}(x) - f(x)|}{|f(x)|}$

$$= \frac{a(\pm\varepsilon_M\pm\sqrt{\varepsilon_M})}{a} = \pm\varepsilon_M\pm\sqrt{\varepsilon_M}$$

As we know $\varepsilon_M \ll 1$ & $\varepsilon_M > 0$

$\therefore \sqrt{\varepsilon_M} \gg \varepsilon_M$

Hence, we also know that $K(M) = 1$ (from part (e))

$\therefore$ forward relative error $= \pm\varepsilon_M\pm\sqrt{\varepsilon_M} > O(\varepsilon_M)$

So, it is clear that algorithm U is unstable.

Hence, proved.

Aneesh Panchal

06-18-01-10-12-24-1-25223