Math 170A HW #5

1) $w_1 = [1,1,1,1]^T \quad w_2 = [3,3,-1,-1]^T \quad w_3 = [6,0,2,0]^T \quad w_i \in \mathbb{R}^4$

Find an orthogonal basis for the subspace spanned by w_1, w_2, w_3.

Formulas

$$\begin{cases} w_1 = r_{11} q_1 \\ w_2 = r_{12} q_1 + r_{22} q_2 \\ w_3 = r_{13} q_1 + r_{23} q_2 + r_{33} q_3 \end{cases}$$

- $r_{ji} = \langle w_i, q_j \rangle \quad j < i$
- $r_{ii} = \| a_i - \sum_{j=1}^{j-1} r_{ji} q_j \|_2$
- $q_i = \dfrac{a_i - \sum_{j=1}^{j-1} r_{ji} q_j}{r_{ii}}$

orthog

$q_1 = \dfrac{w_1}{r_{11}} = \dfrac{w_1}{\|w_1\|_2} = \boxed{\dfrac{[1,1,1,1]^T}{2} = q_1}$

$r_{11} = \|w_1\|_2 = \boxed{2 = r_{11}}$

$w_2 = r_{12} q_1 + r_{22} q_2$

do $\langle w_2, q_1 \rangle = \langle r_{12} q_1 + r_{22} q_2, q_1 \rangle$

"projection" $= r_{12}\langle q_1, q_1\rangle + \cancel{r_{22}\langle q_2, q_1\rangle}$ 0

$= r_{12}(1)$

$q_2 = \dfrac{w_2 - r_{12} q_1}{r_{22}} = \dfrac{w_2 - r_{12} q_1}{\|w_2 - r_{12} q_1\|_2}$

$r_{12} = \langle w_2, q_1 \rangle$

$= \langle [3,3,-1,-1]^T, \frac{1}{2}[1,1,1,1]^T \rangle$

$[3,3,-1,-1] \cdot \frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} = \frac{3}{2} + \frac{3}{2} + \frac{-1}{2} + \frac{-1}{2}$

$= 3 - 1 = \boxed{2 = r_{12}}$

$q_2 = \dfrac{\begin{bmatrix}3\\3\\-1\\-1\end{bmatrix} - 2\frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}}{\text{norm}} = \dfrac{\begin{bmatrix}2\\2\\-2\\-2\end{bmatrix}}{4} = \boxed{\frac{1}{2}\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix} = q_2}$

$\boxed{r_{22} = 4}$

$w_3 = r_{13} q_1 + r_{23} q_2 + r_{33} q_3$

$q_3 = \dfrac{w_3 - r_{13} q_1 - r_{23} q_2}{r_{33}}$

$r_{13} = \langle w_3, q_1 \rangle$

$= [6,0,2,0] \cdot \frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$

$= 3 + 0 + 1 + 0 = 4$

$\boxed{r_{13} = 4}$

$r_{23} = \langle w_3, q_2 \rangle$

$= [6,0,2,0] \cdot \frac{1}{2}\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}$

$= 3 + 0 + (-1) + 0$

$= \boxed{2 = r_{23}}$

$q_3 = \dfrac{\begin{bmatrix}6\\0\\2\\0\end{bmatrix} - 4\begin{bmatrix}1/2\\1/2\\1/2\\1/2\end{bmatrix} - 2\begin{bmatrix}1/2\\1/2\\-1/2\\-1/2\end{bmatrix}}{r_{33}} = \begin{bmatrix}6\\0\\2\\0\end{bmatrix} - \begin{bmatrix}2\\2\\2\\2\end{bmatrix} - \begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}$

$q_3 = \dfrac{\begin{bmatrix}3\\-3\\1\\-1\end{bmatrix}}{r_{33}}$

$r_{33} = \text{norm} = \sqrt{9+9+1+1}$

$\boxed{r_{33} = \sqrt{20}}$

$\boxed{q_3 = \frac{1}{\sqrt{20}}\begin{bmatrix}3\\-3\\1\\-1\end{bmatrix}}$

Now orthogonal basis is $\{q_1, q_2, q_3\}$

1 correct

$$\boxed{Q = \begin{pmatrix} 1/2 & 1/2 & 3/\sqrt{20} \\ 1/2 & 1/2 & -3/\sqrt{20} \\ 1/2 & -1/2 & 1/\sqrt{20} \\ 1/2 & -1/2 & -1/\sqrt{20} \end{pmatrix}}$$

2. Let $A = \begin{bmatrix} 0 & 1 \\ \varepsilon & 0 \end{bmatrix}$ for $0 < \varepsilon < 1$ small and let $b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

a) calculate the soln to $Ax = b$

$\begin{bmatrix} 0 & 1 \\ \varepsilon & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow 0x_1 + x_2 = 1 \quad \boxed{x_2 = 1}$ $\quad \varepsilon x_1 + 0x_2 = 0$

$x_1 = \frac{0}{\varepsilon} = 0 \quad \boxed{x_1 = 0}$

$\boxed{x = \begin{bmatrix} 0 \\ 1 \end{bmatrix}}$ ✓

b) compute $K(A)$ wrt the ∞-norm.

recall $\|A\|_\infty = \max_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_i \sum_{j=1}^{n} |a_{ij}|$

sum over rows and get max

$\boxed{K(A) = \|A\| \cdot \|A^{-1}\|} \longrightarrow \|A\|_\infty = \begin{bmatrix} 0 & 1 \\ \varepsilon & 0 \end{bmatrix} \begin{matrix} = \Sigma 1 \\ = \Sigma \varepsilon \end{matrix} \quad \|A\|_\infty = 1$

$A^{-1} = \frac{1}{-\varepsilon}\begin{bmatrix} 0 & -1 \\ -\varepsilon & 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{\varepsilon} \\ 1 & 0 \end{bmatrix}; \quad \|A^{-1}\|_\infty = \begin{matrix} \Sigma \frac{1}{\varepsilon} \\ \Sigma 1 \end{matrix}$

$\|A^{-1}\|_\infty = \frac{1}{\varepsilon} \qquad \frac{1}{\varepsilon} > 1$ b/c $\varepsilon < 1$

$\therefore K(A) = \|A\|_\infty \cdot \|A^{-1}\|_\infty$

$= (1) \cdot \frac{1}{\varepsilon} = \boxed{\frac{1}{\varepsilon} = K(A)}$ ✓

c) Let $\delta b = \begin{bmatrix} 0 \\ \varepsilon \end{bmatrix}$. Solve $A\hat{x} = b + \delta b$

and let $\delta x = \hat{x} - x$ with x in part a.

$A(x + \delta x) = b + \delta b$

$Ax + A\delta x = b + \delta b$

$A\delta x = \delta b$

$\begin{bmatrix} 0 & 1 \\ \varepsilon & 0 \end{bmatrix} \delta x = \begin{bmatrix} 0 \\ \varepsilon \end{bmatrix}$

$0\delta x_1 + \delta x_2 = 0$

$\varepsilon \delta x_1 + 0\delta x_2 = \varepsilon$

$\delta x_1 = \frac{\varepsilon}{\varepsilon} = 1$

$\delta x = \begin{bmatrix} \delta x_1 \\ \delta x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$\hat{x} = x + \delta x$

$= \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$\boxed{\hat{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}}$

Problem is:
Adx = db
dx = A^-1 * db
forgot to use A^-1 instead of A we get dx = [1;0]

d) verify $\frac{\|\delta x\|}{\|x\|} \leq K(A) \frac{\|\delta b\|}{\|b\|}$ $\quad \|\cdot\| \Rightarrow \infty$-norm

$\|\delta x\|_\infty = 1$
$\|x\|_\infty = 1$
$\|\delta b\|_\infty = \varepsilon$
$\|b\|_\infty = 1$

$\frac{1}{1} \leq \frac{1}{\varepsilon} \cdot \frac{\varepsilon}{1} \Rightarrow 1 \leq 1$ ✓

luckily infinity norm of dx is still one in part d

e) When comparing δx and δb, explain why large $K(A)$ is problematic

If $K(A)$ is large, the relative error in b could be small or big we don't know. $K(A)$ large but rel. error $\left(\frac{\|b - \hat{b}\|}{\|b\|} = \frac{\|\delta b\|}{\|b\|}\right)$ small, $K(A) \cdot \frac{\|\delta b\|}{\|b\|}$ May be large...

If both $K(A)$, rel error are small, then $\frac{\|x - \hat{x}\|}{\|x\|} = \frac{\|\delta x\|}{\|x\|}$ also has to be small

$\boxed{\text{Essentially, large uncertainty if } K(A) \text{ big}}$ ✓.

relative error in x can be big even if relative error in b is small when k(A) is large

$(\sum |a_{ij}|^2)^{\frac{1}{2}}$

$\|A\|_F = \|v\|_2$ where $v = \begin{matrix} a_{11} \\ \vdots \\ a_{1n} \\ \vdots \end{matrix}$

$\|a_i\|_2 = \langle a_i, a_i \rangle = \sum_{i=1}^{n} |a_i|^2$

3.) Let A be $(n \times n)$ matrix and $\|A\|_F$ is Frobenius norm of A.

(a) Show that $\|A\|_F = \|UA\|_F$ for any orthogonal $(n \times n)$ matrix U

$\|A\|_F = \|UA\|_F$

$$\|A\|_F = \sqrt{\sum_{i=1}^{n} \|a_i\|_2} \underset{\substack{\text{multiply by} \\ \text{orthog matrix} \\ \text{still equivalent}}}{=} \sqrt{\sum_{i=1}^{n} \|Ua_i\|_2} = \|UA\|_F \quad \boxed{\text{QED}}$$

a_i is column of A

$\langle a_i, a_i \rangle = \sum_{i=1}^{n} |a_i|^2$

this takes all entries of A and squares and sums

(b) Prove $\|A\|_F = \|AV\|_F$ by transposing.

$$\|A\|_F = \|A^T\|_F = \|A^T V^T\|_F = \|(VA)^T\|_F = \|(VA)\|_F \quad \boxed{\text{QED}}$$

equivalent b/c Frob norm column stacks matrix & takes 2 norm sum order of elements doesn't matter.

multiplying orthog matrix doesn't change length

transpose property.

transpose doesn't change $\|\cdot\|_F$ value

(c) conclude $\|A\|_F = \sqrt{\sum_{i=1}^{n} \sigma_i^2}$

$A = U \Sigma V^T$

$$\Sigma = \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \sigma_2 & \\ 0 & & \ddots \; \sigma_n \end{bmatrix}_{n \times n}$$

$$\|A\|_F = \|U \Sigma V^T\|_F \underset{\substack{\uparrow \\ \text{from} \\ a), b)}}{=} \|\Sigma\|_F = \sqrt{\sum_{i=1}^{n} \sigma_i^2} \checkmark$$

(d) Find formula for $K_F(A)$ in terms of singular values of A.

$K_F(A) = \|A\|_F \cdot \|A^{-1}\|_F$

Frobenius

$\Rightarrow$ from (c)

$= \sqrt{\sum_{i=1}^{n} \sigma_i^2} \sqrt{\sum_{i=1}^{n} (1/\sigma_i)^2}$

$$\Sigma^{-1} = \begin{bmatrix} \frac{1}{\sigma_1} & & \\ & \frac{1}{\sigma_2} & \\ & & \ddots \end{bmatrix}$$

$$\boxed{K(A) = \sqrt{\sum_{i=1}^{n} \sigma_i^2} \cdot \sqrt{\sum_{i=1}^{n} (1/\sigma_i)^2}}$$

4) Let $A \in \mathbb{R}^{4\times 3}$, rank$(A) = 2$, consider full SVD of A, $A = U\Sigma V^T$

$U - (4\times4)$ $V - (3\times3)$ $\Sigma - (4\times3)$

$\uparrow$ Σ only has 2 nonzero entries rank$(A) = 2$

(a) Show $A = \sum_{i=1}^{2} \sigma_i u_i v_i^T$ where i cols of each matrix, σ_i singular values

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}_{4\times3}$$ because rank$(A) = 2$

$$U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ u_{21} & u_{22} & u_{23} & u_{24} \\ u_{31} & u_{32} & u_{33} & u_{34} \\ u_{41} & u_{42} & u_{43} & u_{44} \end{bmatrix} \quad V = \begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix} \quad V^T = \begin{bmatrix} v_{11} & v_{21} & v_{31} \\ v_{12} & v_{22} & v_{32} \\ v_{13} & v_{23} & v_{33} \end{bmatrix}$$

doing $\underset{(4\times3)(3\times3)}{\Sigma V^T} = \begin{bmatrix} \sigma_1 v_{11} & \sigma_1 v_{21} & \sigma_1 v_{31} \\ \sigma_2 v_{12} & \sigma_2 v_{22} & \sigma_2 v_{32} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}_{4\times3}$

doing $\underset{(4\times3)(4\times3)}{U(\Sigma V^T)} \Rightarrow \ldots \begin{pmatrix} u_{13} & u_{14} \\ \vdots & \vdots \\ u_{43} & u_{44} \end{pmatrix}$ get mapped to zero

$$A = \Sigma = \begin{pmatrix} u_{11}\sigma_1 v_{11} + u_{12}\sigma_2 v_{12} & u_{11}\sigma_1 v_{21} + u_{12}\sigma_2 v_{22} & u_{11}\sigma_1 v_{31} + u_{12}\sigma_2 v_{32} \\ u_{21}\sigma_1 v_{11} + u_{22}\sigma_2 v_{22} & \cdots & \vdots \\ \vdots & & \\ u_{41}\sigma_1 v_{11} + u_{42}\sigma_2 v_{12} & \cdots & u_{41}\sigma_1 v_{31} + u_{42}\sigma_2 v_{32} \end{pmatrix}$$

$$\underset{\substack{4\times1 \; 1\times3 \\ =4\times3}}{\sigma_1 u_1 v_1^T} = \begin{bmatrix} u_{11} \\ u_{21} \\ u_{31} \\ u_{41} \end{bmatrix} \begin{bmatrix} v_{11} & v_{21} & v_{31} \end{bmatrix} = \begin{bmatrix} u_{11}v_{11} & u_{11}v_{21} & u_{11}v_{31} \\ u_{21}v_{11} & u_{21}v_{21} & u_{21}v_{31} \\ \cancel{u_{31}v_{11}} & \cancel{u_{31}v_{21}} & \cancel{u_{31}v_{31}} \\ u_{41}v_{11} & u_{41}v_{21} & u_{41}v_{31} \end{bmatrix} \sigma_1$$

$$\sigma_2 u_2 v_2^T = \begin{bmatrix} u_{12} \\ u_{22} \\ u_{32} \\ u_{42} \end{bmatrix} \begin{bmatrix} v_{12} & v_{22} & v_{32} \end{bmatrix} = \begin{bmatrix} u_{12}v_{12} & u_{12}v_{22} & u_{12}v_{32} \\ u_{22}v_{12} & u_{22}v_{22} & u_{22}v_{32} \\ \cancel{u_{32}v_{12}} & \cancel{u_{32}v_{22}} & \cancel{u_{32}v_{32}} \\ u_{42}v_{12} & u_{42}v_{22} & v_{42}v_{32} \end{bmatrix} \sigma_2$$

summing$(\sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T)$ we get $\begin{pmatrix} u_{11}\sigma_1 v_{11} + u_{12}\sigma_2 v_{12} \ldots \\ \vdots \\ \ldots \quad u_{41}\sigma_1 v_{31} + v_{42}\sigma_2 v_{32} \end{pmatrix}$

which is what we get from doing $A = U\Sigma V^T$ thus $A = \sum_{i=1}^{2} \sigma_i u_i v_i^T$.

$\boxed{\text{QED}}$ ✓

✓ I only had to show LHS.
So expand A = UEV^T = for i=1:2

4.) show that the decomposition in (a) is rank-1 decomposition

X (b) i.e., show that for $u \in \mathbb{R}^4$, $v \in \mathbb{R}^3$, the matrix $uv^T \in \mathbb{R}^{4\times 3}$ has rank 1.

$$\sigma_1 \underset{(4\times1)}{u_1} \underset{(1\times3)}{v_1^T} = \sigma_1 \cdot \begin{bmatrix} u_{11} \\ u_{21} \\ u_{31} \\ u_{41} \end{bmatrix} \cdot \begin{bmatrix} v_{11} & v_{21} & v_{31} \end{bmatrix} = \begin{bmatrix} u_{11}v_{11} & u_{11}v_{21} & u_{11}v_{31} \\ u_{21}v_{11} & u_{21}v_{21} & u_{21}v_{31} \\ u_{31}v_{11} & u_{31}v_{21} & u_{31}v_{31} \\ u_{41}v_{11} & u_{41}v_{21} & u_{41}v_{31} \end{bmatrix} \sigma_1$$

Here we multiply two vectors u_1, v_1^T to form a matrix which has rank 1, and similarly for u_2, v_2^T.

From (a) we know $A = \sum_{i=1}^{2} \sigma_i u_i v_i^T$ which is the sum of rank one matrices with rank(A) = 2.

By the definition this decomposition is a rank-1 decomposition.

Show that uv^T has rank one. u = [u1, u2, u3, u4]^T; v = [v1,v2,v3,v4];
do expansion: same as above but ignore extra "1" in each element, I realize that u,v are vectors not matrices…

Reason: looking at expansion of uv^T we see that the second and third columns are multiples of the first column, which means uv^T has rank=1.

MATH170A HW5

Tyler Barbero
Instructor: Caroline Moosmueller

Due 16 November 2020

Problem 5

a) Explain line by line what the code does (you might need to google some of the commands).

4. reads image from file and stores data type (uint8) in array A.
5. converts truecolor image to grayscale by eliminating hue and saturation, retains luminance (intensity of light). I assume hue and saturation are 2/3 z-dimensions turning it into a 2d array of just luminance.
7. convert matrix A to double data type from uint8 unsigned integer 8bit.
11. returns size of B (x,y)
12. stores rank (scalar) of matrix b in var r.
13. does singular value decomp of matrix B and outputs u,v (orthog matrices, and s (singular value matrix).
17. stores numbers in a matrix "ranks" with r=rank(B)=480 as the last element.
18. stores length of matrix "ranks" in var "l"
20. starts for loop for i=1 to i=l, repeats code in loop l-times.
24. store the i-th element of ranks in var k.
26. matrix multiplication: U(all rows, 1 to k columns) * S(1 to k rows, 1 to k columns) * V(all rows, 1 to k columns) transposed. stores result in matrix approxB.
28. convert approxB data type from double to uint8 and store in approxA
32. designate figure 1.
33. create a 8 subplots in 2 rows and 4 columns (assigns a plot to one subplot in each iteration of loop).
34. for each subplot(2,4,i) plot the approxA matrix which changes based on i-th rank.
35. titles each subplot. sprintf will screen print the text and format a number into the first argument of sprintf which is a char array. Number is formatted into char array by %d.

b) Explain mathematically what the code does with the original image.

The image is a 480x640 uint8 matrix. The code will use more singular values of S as k = ranks(i) where with each iteration of the loop k (rank) will increase in value.

c) The approximation gets better as we increase k. Already for k= 100, the resulting approximation looks reasonable. What is the advantage to use/store the k= 100approximation instead of the original image? What is the disadvantage?

The advantage is that we don't need all of the of the data in A outside the values corresponding to the rank k=100. The disadvantage is that it will be an approximation and we won't have the full quality.

d) For a general image, by using the SVD, how can one determine a value for k that results in a reasonable approximation?

We can use the low-rank approximation. We approximate A with a rank-k matrix such that k ¡ r. We know that the singular values are descending meaning $\sigma_1 > \sigma_2 > ... > \sigma_k > ... > \sigma_r$. We can choose a σ_k that is very small and if we cut off $\sigma_{k+1}...\sigma_r$ it won't affect the result very much and we still get a good approximation.