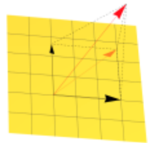
HW 7 Matrix norms and SVD answers

More projections!



① The following code generates 10 random vectors $\vec{v}_1, \ldots, \vec{v}_{10}$ and $\vec{u} \in \mathbb{R}^{10}$ (different in Matlab and Python)

seed random number generator

dimension

\# vectors

```
rng(1)
d = 10;
n = 10;
v = randn(d, n);
u = randn(d, 1);
```

```
np.random.seed(1)

d = 10
n = 10

v = randn(d, n)
u = randn(d, 1)
```

Using these same vectors, fill in the following table. Show your work.

k	$\lVert \vec{u} - \mathrm{Proj}_{\mathrm{span}(v_1,\ldots,v_k)}(\vec{u}) \rVert$
1	
2	
⋮	
n	

Python code:

```
import numpy as np
from scipy.linalg import norm, orth
from numpy.random import randn
np.random.seed(1)

d = 10
n = 10

v = randn(d, n)
u = randn(d, 1)

for k in range(n):
  o = orth(v[:,:k+1]) # orthonormal basis
                      # for first k columns
  projection = o.dot(o.T.dot(u))
  error = norm(u - projection)
  print(k+1, error)
```

```
1 2.1892117477365662
2 2.155577868662765
3 2.11195399881934
4 1.6323855457812806
5 1.6144862764208916
6 0.7269465908604331
7 0.7069842069043837
8 0.6990117688191947
9 0.52519313123074
10 3.4108880966772296e-15
```

```
# This code also works.  Instead of recomputing orth()
# at every step, it just takes the QR decomposition
# (similar to Gram-Schmidt) once at the beginning

o = np.linalg.qr(v)[0]
projection = np.zeros(d)
for k in range(n):
  projection += np.dot(o[:,k], u) * o[:,k]
  error = norm(u[:,0] - projection)
  print(k+1, error)
```

```
1 2.1892117477365662
2 2.155577868662765
3 2.11195399881934
4 1.6323855457812806
5 1.6144862764208914
6 0.7269465908604326
7 0.7069842069043831
8 0.699011768819194
9 0.5251931312307403
10 9.515144070671971e-16
```

```
# NOTE that this code gives the wrong answers.!
# The reason is that orth(v)[:,:k+1] is NOT
# guaranteed to span the first k columns of v

o = orth(v)
projection = np.zeros(d)
for k in range(n):
  projection += np.dot(o[:,k], u) * o[:,k]
  error = norm(u[:,0] - projection)
  print(k+1, error)
```

```
1 2.1892117477365662
2 2.155577868662765
3 2.11195399881934
4 1.6323855457812806
5 1.6144862764208916
6 0.7269465908604331
7 0.7069842069043837
8 0.6990117688191947
9 0.52519313123074
10 3.4108880966772296e-15
```

```python
# NOTE that this code gives the wrong answers.
# The reason is that orth(v)[:,:k+1] is NOT
# guaranteed to span the first k columns of v

o = orth(v)
projection = np.zeros(d)
for k in range(n):
  projection += np.dot(o[:,k], u) * o[:,k]
  error = norm(u[:,0] - projection)
  print(k+1, error)
```

```
1 2.2136769660885376
2 2.025283910507803
3 1.9537889045706607
4 1.9268327527440012
5 1.5309784000142703
6 1.5301008737781605
7 1.0758076394975862
8 0.7239283153907264
9 0.7091014734850234
10 3.751215075901527e-15
```

Matlab code:

```matlab
>> for k = 1:n
     o = orth(v(:,1:k));
     projection = o * (o' * u);
     error = norm(u - projection);
     fprintf('%d %f\n', k, error);
   end;
1 2.957450
2 2.769828
3 2.471525
4 2.378050
5 1.991503
6 1.621573
7 1.600526
8 1.463147
9 1.317561
10 0.000000
```

2

Let $f(x) = \cosh(x) = \dfrac{e^x + e^{-x}}{2}$.

For $k = 0, 1, \dots, 10$,
find the degree-k polynomial $p_k(x)$ that minimizes

$$\mathbb{E}_{X \sim N(0,1)}\left[(f(X) - p_k(X))^2\right].$$

For your answer, then fill in this table:

k	$p_k(x)$	$\sqrt{\mathbb{E}_X\left[(f(X)-p_k(X))^2\right]}$
0		
1		
$\vdots$		

1

⋮

10

Explain how you computed $p_k(x)$ and show at least some of your work. For example, if you use Wolfram Alpha, take at least one screenshot showing what the integrals look like.

Hints: ① Recall the Hermite polynomials

https://en.wikipedia.org/wiki/Hermite_polynomials#Definition

"probabilists' Hermite polynomials"

$$
\begin{aligned}
He_0(x) &= 1,\\
He_1(x) &= x,\\
He_2(x) &= x^2 - 1,\\
He_3(x) &= x^3 - 3x,\\
He_4(x) &= x^4 - 6x^2 + 3,\\
He_5(x) &= x^5 - 10x^3 + 15x,\\
He_6(x) &= x^6 - 15x^4 + 45x^2 - 15,\\
He_7(x) &= x^7 - 21x^5 + 105x^3 - 105x,\\
He_8(x) &= x^8 - 28x^6 + 210x^4 - 420x^2 + 105,\\
He_9(x) &= x^9 - 36x^7 + 378x^5 - 1260x^3 + 945x,\\
He_{10}(x) &= x^{10} - 45x^8 + 630x^6 - 3150x^4 + 4725x^2 - 945.
\end{aligned}
$$

↑orthogonal with respect to

$$\langle f,g\rangle = \int_{-\infty}^{\infty} f(x)g(x)e^{-x^2/2}\,dx$$

"physicists' Hermite polynomials"

$$
\begin{aligned}
H_0(x) &= 1,\\
H_1(x) &= 2x,\\
H_2(x) &= 4x^2 - 2,\\
H_3(x) &= 8x^3 - 12x,\\
H_4(x) &= 16x^4 - 48x^2 + 12,\\
H_5(x) &= 32x^5 - 160x^3 + 120x,\\
H_6(x) &= 64x^6 - 480x^4 + 720x^2 - 120,\\
H_7(x) &= 128x^7 - 1344x^5 + 3360x^3 - 1680x,\\
H_8(x) &= 256x^8 - 3584x^6 + 13440x^4 - 13440x^2 + 1680,\\
H_9(x) &= 512x^9 - 9216x^7 + 48384x^5 - 80640x^3 + 30240x,\\
H_{10}(x) &= 1024x^{10} - 23040x^8 + 161280x^6 - 403200x^4 + 302400x^2 - 30240.
\end{aligned}
$$

↑orthogonal with respect to

$$\langle f,g\rangle = \int_{-\infty}^{\infty} f(x)g(x)e^{-x^2}\,dx$$

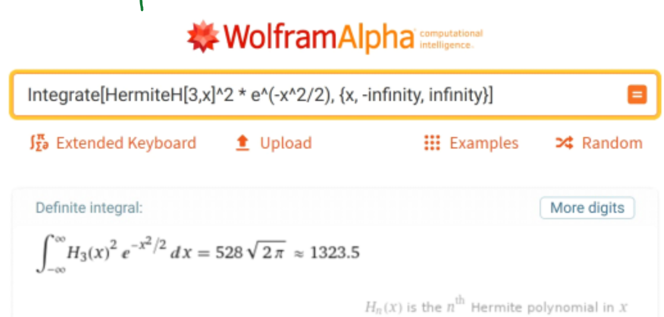
The Mathematica function HermiteH[j, x] returns the jth physicist poly.

② You can easily use Mathematica or Wolfram Alpha to compute any necessary integrals. https://www.wolframalpha.com/ For example,

Mathematica:

$$\int_{-\infty}^{\infty} \texttt{HermiteH[3, x]}^2\, e^{-x^2/2}\, dx$$

$528\sqrt{2\pi}$



③ I'll give you one for free.

$p_0(x) = \mathbb{E}[\cosh X] = \sqrt{e}$, and $\sqrt{\mathbb{E}_X[(f(X)-p_0(X))^2]} = \frac{e-1}{\sqrt{2}} \approx 1.215$

Answer:

More generally, consider the problem of finding

$$\min_{\substack{p \in \text{Span}\{1, x, \dots, x^k\} \\ \parallel \\ S_k}} \mathbb{E}_X\left[(f(X) - p(X))^2\right]$$

where X is drawn from a distribution with probability density $\sigma(x)$.

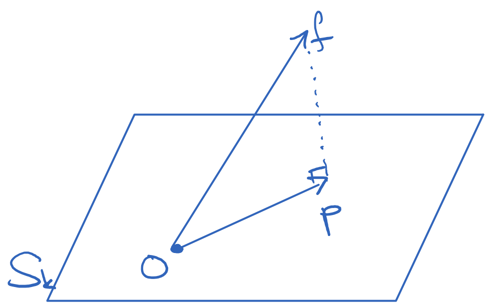
Then $\mathbb{E}_X\left[(f(X)-p(X))^2\right] = \int_{-\infty}^{\infty} dx\, \sigma(x) \left(f(x) - p(x)\right)^2$

$= \|f - p\|^2$

where $\|f-p\|^2 = \langle f-p, f-p \rangle$ for the inner product $\langle \cdot , \cdot \rangle$ defined by

$$\langle g, h \rangle := \int_{-\infty}^{\infty} dx\, \sigma(x)\, g(x)^* h(x)$$

For $p \in S_k = \text{Span}\{1, x, \dots, x^k\}$, the norm $\|f-p\|^2$ is minimized for for



$p = \text{Projection}_{S_k}(f)$

where the projection is with respect to the same inner product.

To compute the projection, we need an orthogonal basis for S_k. In general, one can be computed using Gram-Schmidt.

For this problem, however, we are given an orthogonal basis, the (probabilists') Hermite polynomials.

The density of X is given by $\sigma(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

Our goal is to minimize

$$\int_{-\infty}^{\infty} dx\, \sigma(x) \left(f(x) - p_k(x)\right)^2$$

Recall that the Hermite polynomials are orthogonal with respect to the inner product

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x) g(x) e^{-x^2/2} dx$$

Therefore the projection of f onto the span of $h_0, \dots, h_k$

is given by $P_k f := \sum_{j=0}^{k} \frac{\langle h_j, f\rangle}{\|h_j\|^2} h_j$

We can use Mathematica or Wolfram Alpha to compute each term.

We have $\|f\|^2 = \langle f, f\rangle = \int_{-\infty}^{\infty} \text{Cosh}[x]^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, dx$

$\frac{1}{2}(1 + e^2)$

```
Table[{k, HermiteH[k, x]}, {k, 0, 10}] // MatrixForm
```

$$\begin{pmatrix}
0 & 1 \\
1 & 2x \\
2 & -2+4x^2 \\
3 & -12x+8x^3 \\
4 & 12-48x^2+16x^4 \\
5 & 120x-160x^3+32x^5 \\
6 & -120+720x^2-480x^4+64x^6 \\
7 & -1680x+3360x^3-1344x^5+128x^7 \\
8 & 1680-13440x^2+13440x^4-3584x^6+256x^8 \\
9 & 30240x-80640x^3+48384x^5-9216x^7+512x^9 \\
10 & -30240+302400x^2-403200x^4+161280x^6-23040x^8+1024x^{10}
\end{pmatrix}$$

Observe that Mathematica's HermiteH[] polynomials are orthogonal with respect to the density e^{-x^2}. A change of variables makes them orthogonal with respect to $e^{-x^2/2}$.

$$\text{Table}\left[\int_{-\infty}^{\infty} \text{HermiteH}[j, x]\,\text{HermiteH}[k, x]\, e^{-x^2} \, dx, \{j, 0, 10\}, \{k, 0, 10\}\right] \text{ // MatrixForm}$$

$$\begin{pmatrix}
\sqrt{\pi} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 2\sqrt{\pi} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 8\sqrt{\pi} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 48\sqrt{\pi} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 384\sqrt{\pi} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 3840\sqrt{\pi} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 46080\sqrt{\pi} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 645120\sqrt{\pi} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10321920\sqrt{\pi} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 185794560\sqrt{\pi} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 37158912
\end{pmatrix}$$

$$\text{Table}\left[\int_{-\infty}^{\infty} \text{HermiteH}\left[j, \frac{x}{\sqrt{2}}\right] \text{HermiteH}\left[k, \frac{x}{\sqrt{2}}\right] \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, dx, \{j, 0, 10\}, \{k, 0, 10\}\right] \text{ // MatrixForm}$$

$$\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 48 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 384 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 3840 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 46080 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 645120 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10321920 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 185794560 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3715891200
\end{pmatrix}$$

Projecting a function f in the direction of H_k gives $(\langle H_k, f\rangle / \|H_k\|^2) * H_k$.

$$\text{coefficients} = \text{Table}\left[\frac{\int_{-\infty}^{\infty} \text{Cosh}[x]\,\text{HermiteH}\left[k, \frac{x}{\sqrt{2}}\right] \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx}{\int_{-\infty}^{\infty} \text{HermiteH}\left[k, \frac{x}{\sqrt{2}}\right]^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx}, \{k, 0, 10\}\right];$$

$$\text{Table}\left[\left\{k, \text{HermiteH}\left[k, \frac{x}{\sqrt{2}}\right], \text{coefficients}[[1+k]]\right\}, \{k, 0, 10\}\right] \text{ // MatrixForm}$$

$$\begin{pmatrix}
0 & 1 & \sqrt{e} \\
1 & \sqrt{2}\,x & 0 \\
2 & -2+2x^2 & \frac{\sqrt{e}}{4} \\
3 & -6\sqrt{2}\,x+2\sqrt{2}\,x^3 & 0 \\
4 & 12-24x^2+4x^4 & \frac{\sqrt{e}}{96} \\
5 & 60\sqrt{2}\,x-40\sqrt{2}\,x^3+4\sqrt{2}\,x^5 & 0 \\
6 & -120+360x^2-120x^4+8x^6 & \frac{\sqrt{e}}{5760} \\
7 & -840\sqrt{2}\,x+840\sqrt{2}\,x^3-168\sqrt{2}\,x^5+8\sqrt{2}\,x^7 & 0 \\
8 & 1680-6720x^2+3360x^4-448x^6+16x^8 & \frac{\sqrt{e}}{645120} \\
9 & 15120\sqrt{2}\,x-20160\sqrt{2}\,x^3+6048\sqrt{2}\,x^5-576\sqrt{2}\,x^7+16\sqrt{2}\,x^9 & 0 \\
10 & -30240+151200x^2-100800x^4+20160x^6-1440x^8+32x^{10} & \frac{\sqrt{e}}{116121600}
\end{pmatrix}$$

$$\text{projections} = \text{Table}\left[\sum_{j=0}^{k} \text{coefficients}[[1+j]]\,\text{HermiteH}\left[j, \frac{x}{\sqrt{2}}\right], \{k, 0, 10\}\right] \text{ // Simplify;}$$

$$\text{Table}\left[\left\{k, \text{projections}[[1+k]], \text{error} = \sqrt{\int_{-\infty}^{\infty} (\text{Cosh}[x] - \text{projections}[[1+k]])^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx} \text{ // Simplify}, \text{error // N}\right\}, \{k, 0, 10\}\right] \text{ // MatrixForm}$$

$$\begin{array}{cccc}
k & p_k(x) & \text{error} & \approx \\
0 & \sqrt{e} & \frac{-1+e}{\sqrt{2}} & 1.21501 \\
1 & \sqrt{e} & \frac{-1+e}{\sqrt{2}} & 1.21501 \\
2 & \frac{1}{2}\sqrt{e}\,(1+x^2) & \sqrt{\frac{1}{2}(1+(-3+e)\,e)} & 0.342207 \\
3 & \frac{1}{2}\sqrt{e}\,(1+x^2) & \sqrt{\frac{1}{2}(1+(-3+e)\,e)} & 0.342207 \\
4 & \frac{1}{24}\sqrt{e}\,(15+6x^2+x^4) & \frac{1}{2}\sqrt{2-\frac{37e}{6}+2e^2} & 0.0619965 \\
5 & \frac{1}{24}\sqrt{e}\,(15+6x^2+x^4) & \frac{1}{2}\sqrt{2-\frac{37e}{6}+2e^2} & 0.0619965 \\
6 & \frac{1}{720}\sqrt{e}\,(435+225x^2+15x^4+x^6) & \sqrt{\frac{1}{2}-\frac{1111e}{720}+\frac{e^2}{2}} & 0.00825666 \\
7 & \frac{1}{720}\sqrt{e}\,(435+225x^2+15x^4+x^6) & \sqrt{\frac{1}{2}-\frac{1111e}{720}+\frac{e^2}{2}} & 0.00825666 \\
8 & \frac{\sqrt{e}\,(24465+12180x^2+1050x^4+28x^6+x^8)}{40320} & \sqrt{\frac{1}{2}-\frac{6913e}{4480}+\frac{e^2}{2}} & 0.000868788 \\
9 & \frac{\sqrt{e}\,(24465+12180x^2+1050x^4+28x^6+x^8)}{40320} & \sqrt{\frac{1}{2}-\frac{6913e}{4480}+\frac{e^2}{2}} & 0.000868788 \\
10 & \frac{\sqrt{e}\,(2200905+1100925x^2+91350x^4+3150x^6+45x^8+x^{10})}{3628800} & \sqrt{\frac{1}{2}-\frac{799933e}{518400}+\frac{e^2}{2}} & 0.0000755394
\end{array}$$

Another way of computing the errors is to use Pythagoras's theorem; the error is basically the amount of $\|f\|^2$ unaccounted for by the projection. This gives the same answer as above.

$$\text{norm2} = \int_{-\infty}^{\infty} \text{Cosh}[x]^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, dx$$

```
norm2 // N
```

$\frac{1}{2}(1 + e^2)$

4.19453

```
"The errors";
```

Another way of computing the errors is to use Pythagoras's theorem; the error is basically the amount of $\|f\|^2$ unaccounted for by the projection. This gives the same answer as above.

$$\texttt{norm2} = \int_{-\infty}^{\infty} \texttt{Cosh[x]}^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx$$

```
norm2 // N
```

$$\frac{1}{2}\left(1+e^2\right)$$

```
4.19453
```

```
"The errors";
```

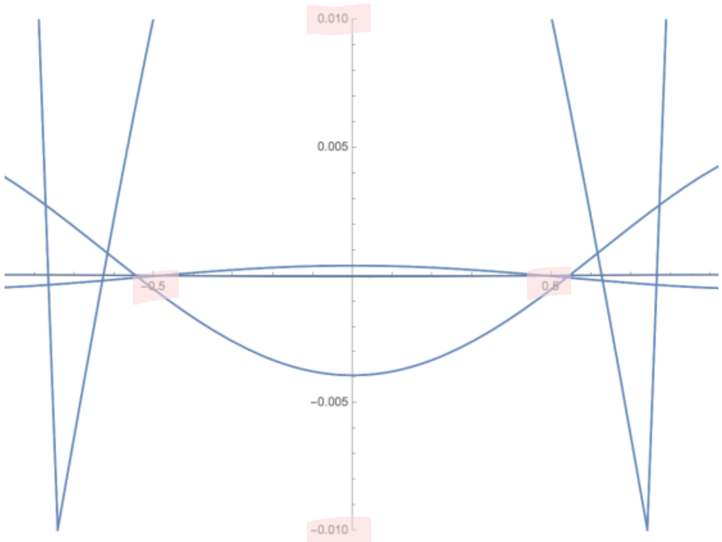
$$\texttt{Table}\left[\left\{k, \sqrt{\texttt{norm2} - \int_{-\infty}^{\infty} \texttt{projections[[1 + k]]}^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx} \texttt{ // N}\right\}, \{k, 0, 10\}\right] \texttt{ //}$$

```
MatrixForm
```

0	1.21501
1	1.21501
2	0.342207
3	0.342207
4	0.0619965
5	0.0619965
6	0.00825666
7	0.00825666
8	0.000868788
9	0.000868788
10	0.0000755394

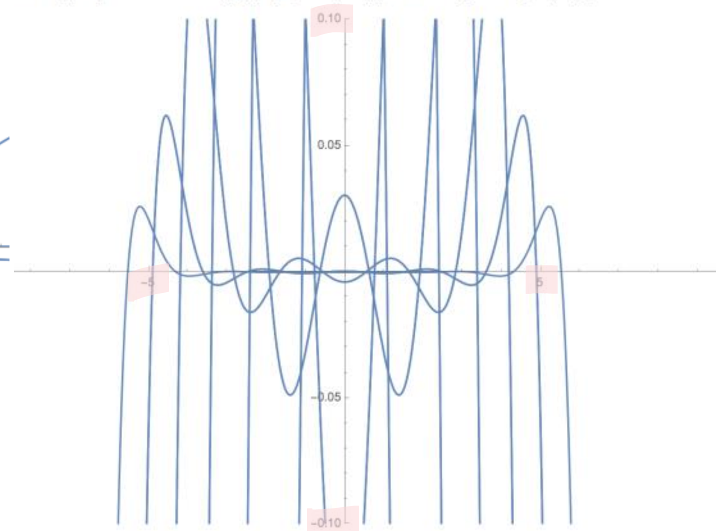
For higher-degree polynomials, the errors away from *f* get smaller and smaller:

```
Plot[projections - Cosh[x], {x, -1, 1}, PlotRange → .01 {-1, 1}]
```



But the errors still diverge for larger values of $|x|$. Fortunately, the density $e^{-x^2/2}$ makes these values less important.

```
Plot[projections - Cosh[x], {x, -10, 10}, PlotRange → .1 {-1, 1}]
```



Matrix norms

③ Compute the exact norms of these matrices:

$$A = \begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 4 & -2 & 4 \\ -2 & 1 & -2 \\ 4 & -2 & 4 \end{pmatrix}$$

$$C = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 0 & -3 \\ -3 & 0 & 2 \end{pmatrix} \qquad D = \frac{1}{\sqrt{3}}\begin{pmatrix} 3 & -1 \\ 0 & \sqrt{8} \end{pmatrix}$$

$$E = D^{-1} \qquad F = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

$$G = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Do the calculations by hand.

Use calculus. Do not use eigenvalues/eigenvectors.

Feel free to use Matlab/Mathematica to check your answers.

Answers:

A is rank one, so $\|A\| = \sqrt{\sum_{ij} |a_{ij}|^2} = \boxed{\sqrt{10}}$

B is also rank one, $B = \begin{pmatrix} 4 \\ -2 \\ 4 \end{pmatrix}\begin{pmatrix} 1 & -\frac{1}{2} & 1 \end{pmatrix}$, $\|B\| = \sqrt{81} = \boxed{9}$

$$\|C\| = \max\left\{ \underbrace{\|(1)\|}_{=1}, \left\|\begin{pmatrix} 2 & -3 \\ -3 & 2 \end{pmatrix}\right\| \right\} = |2| + |3| = \boxed{5}$$

In general, **Claim:** $\left\|\begin{pmatrix} a & b \\ b & a \end{pmatrix}\right\| = |a| + |b|$.

Proof:

$$\left\|\begin{pmatrix} a & b \\ b & a \end{pmatrix}\right\| \leq \left\|\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}\right\| + \left\|\begin{pmatrix} 0 & b \\ b & 0 \end{pmatrix}\right\| \quad (\triangle \text{ ineq.})$$
$$= |a| \cdot \left\|\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right\| + |b| \cdot \left\|\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right\|$$
$$= |a| + |b|. \qquad \square$$

$D = \frac{1}{\sqrt{3}}\begin{pmatrix} 3 & -1 \\ 0 & \sqrt{8} \end{pmatrix}$. Let's use calculus:

$$\mathcal{L} = \|D\vec{x}\|^2 + \lambda(\|\vec{x}\|^2 - 1)$$
$$= \frac{1}{3}\left[(3x_1 - x_2)^2 + 8x_2^2\right] + \lambda(x_1^2 + x_2^2 - 1)$$
$$= (3x_1^2 - 2x_1x_2 + 3x_2^2) + \lambda(x_1^2 + x_2^2 - 1)$$

We want $\frac{\partial \mathcal{L}}{\partial x_1} = 0$, $\frac{\partial \mathcal{L}}{\partial x_2} = 0$, $\frac{\partial \mathcal{L}}{\partial \lambda} = 0$.

Now $\frac{\partial \mathcal{L}}{\partial x_1} = (6 + 2\lambda)x_1 - 2x_2$, etc.

We don't need to do any work, though.

Since $\mathcal{L}$ is symmetrical between x_1 and x_2, the maximum will be with $x_1 = \pm x_2$.

$$\left\|D\tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\|^2 = \frac{1}{6}\cdot(4^2 + 8) = 4, \quad \left\|D\tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\|^2 = \frac{1}{6}(2^2 + 8) < 4$$
$$\Rightarrow \|D\| = \boxed{2}.$$

$$E = D^{-1} = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 1/\sqrt{8} \\ 0 & 3/\sqrt{8} \end{pmatrix}$$

The easiest way to get the norm is to use the SVD:

$$D = \sigma_1 \vec{v}_1 \vec{u}_1^T + \sigma_2 \vec{v}_2 \vec{u}_2^T \quad \Rightarrow E = \frac{1}{\sigma_1}\vec{u}_1\vec{v}_1^T + \frac{1}{\sigma_2}\vec{u}_2\vec{v}_2^T$$

Here $\sigma_1 = \|D\| \geq \sigma_2 > 0$, and $\|E\| = \frac{1}{\sigma_2}$.

We computed $\vec{u}_1 = \pm\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ in part (d). Since $\vec{u}_1 \perp \vec{u}_2$, Here $\sigma_1 = \|D\| \geq \sigma_2 > 0$, and $\|E\| = \frac{1}{\sigma_2}$.

We computed $\vec{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$ in part 2. Since $\vec{u}_2 \perp \vec{u}_1$, it must be that $\vec{u}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$ (or you can use $\frac{-1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$).

$$\Rightarrow \sigma_2 = \|D\vec{u}_2\| = \frac{1}{\sqrt{6}}\left\|\begin{pmatrix}3 & -1\\0 & \sqrt{8}\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}\right\| = \frac{1}{\sqrt{6}}\sqrt{2^2+\sqrt{8}^2} = \sqrt{2}.$$

$$\Rightarrow \boxed{\|E\| = \frac{1}{\sigma_2} = \frac{1}{\sqrt{2}}}$$

Or, we can use calculus again:

One way to compute $\|E\|$ is to optimize

$$\|E\| = \max_{\vec{x}\neq\vec{0}} \frac{\|E\vec{x}\|}{\|\vec{x}\|} = \sqrt{\max_{y\in\mathbb{R}} \frac{\left\|E\begin{pmatrix}1\\y\end{pmatrix}\right\|^2}{\left\|\begin{pmatrix}1\\y\end{pmatrix}\right\|^2}}$$

Then
$$\frac{\left\|E\begin{pmatrix}1\\y\end{pmatrix}\right\|^2}{\left\|\begin{pmatrix}1\\y\end{pmatrix}\right\|^2} = \frac{1}{1+y^2}\cdot\frac{1}{3}\left(\left(1+\frac{y}{\sqrt{8}}\right)^2+\frac{9}{8}y^2\right)$$
$$= \frac{1}{3(1+y^2)}\left(1+\frac{y}{\sqrt{2}}+\frac{5}{4}y^2\right)$$

Setting $\frac{d}{dy}(*) = 0$:
$$0 = (1+y^2)\left(\frac{1}{\sqrt{2}}+\frac{5}{2}y\right) - \left(1+\frac{y}{\sqrt{2}}+\frac{5}{4}y^2\right)2y$$
$$= \frac{1}{\sqrt{2}} + \frac{1}{2}y - \frac{1}{\sqrt{2}}y^2$$
$$= -\frac{1}{\sqrt{2}}(y-\sqrt{2})\left(y+\frac{1}{\sqrt{2}}\right)$$
$$\Rightarrow y \in \left\{-\frac{1}{\sqrt{2}}, \sqrt{2}\right\}$$

Since we are maximizing, we don't want cancellation, so $y>0$ is best $\Rightarrow y = \sqrt{2}$.

$$\frac{\left\|E\begin{pmatrix}1\\\sqrt{2}\end{pmatrix}\right\|^2}{\left\|\begin{pmatrix}1\\\sqrt{2}\end{pmatrix}\right\|^2} = \frac{1}{3}\cdot\frac{1}{3}\left(\left(\frac{3}{2}\right)^2+\left(\frac{3}{2}\right)^2\right) = \frac{1}{2} \quad \Rightarrow \boxed{\|E\| = \frac{1}{\sqrt{2}}}$$

Another way is to use calculus to compute $\|E\| = \max_{\vec{x}:\|\vec{x}\|=1} \|E\vec{x}\|$:

$$\mathcal{L} = \|E\vec{x}\|^2 + \lambda(\|\vec{x}\|^2-1)$$
$$= \underbrace{\frac{1}{3}\left[\left(x_1+\frac{1}{\sqrt{8}}x_2\right)^2+\frac{9}{8}x_2^2\right]}_{=\frac{1}{3}\left(x_1^2+\frac{1}{\sqrt{2}}x_1x_2+\frac{5}{4}x_2^2\right)} + \lambda(x_1^2+x_2^2-1)$$

$$\frac{\partial\mathcal{L}}{\partial x_1} = \left(\frac{2}{3}+2\lambda\right)x_1 + \frac{1}{3\sqrt{2}}x_2 = 0 \quad \Rightarrow x_2 = -3\sqrt{2}\left(\frac{2}{3}+2\lambda\right)x_1$$

$$\frac{\partial\mathcal{L}}{\partial x_2} = \frac{1}{3\sqrt{2}}x_1 + \left(\frac{5}{6}+2\lambda\right)x_2 = 0$$

$$\frac{1}{3\sqrt{2}} - 3\sqrt{2}\left(\frac{2}{3}+2\lambda\right)\left(\frac{5}{6}+2\lambda\right) = 0$$
$$-1 + 18\left(\frac{5}{9} + 2\underbrace{\left(\frac{2}{3}+\frac{5}{6}\right)}_{3}\lambda + 4\lambda^2\right) = 0$$
$$= 9(8\lambda^2+6\lambda+1)$$

$$-1+18\Big(\tfrac{5}{9}+\underbrace{2(\tfrac{2}{3}+\tfrac{5}{6})}_{3}\lambda+4\lambda^2\Big)=0$$

$$=9(8\lambda^2+6\lambda+1)$$

$$=9(4\lambda+1)(2\lambda+1)$$

$$\Rightarrow \lambda\in\{-\tfrac{1}{2},-\tfrac{1}{4}\}$$

Now substitute back into $x_2=-3\sqrt{2}\left(\frac{2}{3}+2\lambda\right)x_1$

For $\lambda=-\frac{1}{4}$, $x_2=-\frac{1}{\sqrt{2}}x_1$.

For $\lambda=-\frac{1}{2}$, $x_2=\sqrt{2}x_1$.

Since we are trying to <u>maximize</u> $\|E\vec{x}\|$, we don't want cancellation, so should use $\lambda=-\frac{1}{2}$.

$$\frac{\left\|E\begin{pmatrix}1\\ \sqrt{2}\end{pmatrix}\right\|^2}{\left\|\begin{pmatrix}1\\ \sqrt{2}\end{pmatrix}\right\|^2}=\frac{1}{3}\cdot\frac{1}{3}\left(\left(\tfrac{3}{2}\right)^2+\left(\tfrac{3}{2}\right)^2\right)=\frac{1}{2}\Rightarrow\boxed{\|E\|=\frac{1}{\sqrt{2}}}$$

Next:

$$F=\begin{pmatrix}1&1&-1\\0&1&1\\1&0&1\end{pmatrix}\qquad \begin{pmatrix}1\\0\\1\end{pmatrix}-\frac{1}{2}\begin{pmatrix}1\\1\\0\end{pmatrix}=\begin{pmatrix}1/2\\-1/2\\1\end{pmatrix}\propto\begin{pmatrix}1\\-1\\2\end{pmatrix}$$

Let

$$\mathcal{U}=\begin{pmatrix}-\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}&0\\ \frac{1}{\sqrt{6}}&\frac{-1}{\sqrt{6}}&\frac{2}{\sqrt{6}}\end{pmatrix}.$$

This is unitary, so $\|\mathcal{U}F\|=\|F\|$.

$$\|\mathcal{U}F\|=\left\|\begin{pmatrix}0&0&\sqrt{3}\\ 1/\sqrt{2}&\sqrt{2}&0\\ \sqrt{\frac{3}{2}}&0&0\end{pmatrix}\right\|=\max\left\{\sqrt{3},\left\|\begin{pmatrix}1/\sqrt{2}&\sqrt{2}\\ \sqrt{3/2}&0\end{pmatrix}\right\|\right\}$$

$$\left\|\begin{pmatrix}1/\sqrt{2}&\sqrt{2}\\ \sqrt{3/2}&0\end{pmatrix}\right\|^2=\frac{1}{2}\left\|\begin{pmatrix}1&2\\ \sqrt{3}&0\end{pmatrix}\right\|^2$$

$$\max_y\frac{\left\|\begin{pmatrix}1&2\\ \sqrt{3}&0\end{pmatrix}\begin{pmatrix}1\\ y\end{pmatrix}\right\|^2}{1+y^2}=\max_y\frac{(1+2y)^2+3}{1+y^2}=\max_y\frac{4(y^2+y+1)}{y^2+1}$$

Setting $\frac{d}{dy}(\sim)=0$ gives $(y^2+1)(2y+1)-(y^2+y+1)2y=0$

$$=(\cancel{2y^3}+y^2+\cancel{2y}+1)-(\cancel{2y^3}+2y^2+\cancel{2y})$$

$$=-y^2+1$$

$$\Rightarrow y\in\{1,-1\}$$

$$\Rightarrow\left\|\begin{pmatrix}1/\sqrt{2}&\sqrt{2}\\ \sqrt{3/2}&0\end{pmatrix}\right\|=\frac{1}{\sqrt{2}}\sqrt{\frac{4(y^2+y+1)}{y^2+1}}\Bigg|_{y=1}=\sqrt{3}.$$

$$\Rightarrow\boxed{\|F\|=\sqrt{3}.}$$

Next:

$$G=\begin{pmatrix}1&-1&0\\0&1&1\end{pmatrix}$$

$$\|G\|^2=\max_{\vec{x}\neq\vec{0}}\frac{\|G\vec{x}\|^2}{\|\vec{x}\|^2}=\max_{\vec{x}\neq\vec{0}}\frac{(x_1-x_2)^2+(x_2+x_3)^2}{x_1^2+x_2^2+x_3^2}$$

By symmetry, the maximum will be achieved with $x_1=-x_3$.
We can also scale so $x_2=1$: $\vec{x}=\begin{pmatrix}-y\\1\\y\end{pmatrix}$. And $y>0$ to avoid cancellation.

$$\Rightarrow\|G\|^2=\max\ 2(1+y)^2$$

We can also scale so $x_2 = 1$: $\vec{x} = \binom{-y}{y}$. And $y > 0$ to avoid cancellation.

$$\Rightarrow \|G\|^2 = \max_{y \in \mathbb{R}} \frac{2(1+y)^2}{1+2y^2}$$

$$0 = \frac{d}{dy}(\sim) \propto (1+2y^2)\cdot 2(1+y) - (1+y)^2 \cdot 4y$$

$$\propto (\cancel{2y^3} + 2y^2 + y + 1) - (\cancel{2y^3} + 4y^2 + 2y)$$

$$= -2y^2 - y + 1$$

$$= (1+y)(1-2y)$$

$$\Rightarrow y \in \{-1, 1/2\} \qquad \Rightarrow y = 1/2$$

$$\Rightarrow \|G\| = \sqrt{\frac{2(1+1/2)^2}{1+2(1/2)^2}} = \boxed{\sqrt{3}}$$

Another way to solve this is to use $\|G\| = \|G^T\|$.

$$\|G^T\|^2 = \max_{y \in \mathbb{R}} \frac{\|G^T\binom{1}{y}\|^2}{1+y^2} \left\{ = \frac{1+(1-y)^2+y^2}{1+y^2} = 1 + \frac{(1-y)^2}{1+y^2} \right.$$

Setting $0 = \frac{d}{dy}(\sim) \propto -(1+y^2)(1-y) - (1-y)^2 y$

$$= (\cancel{y^3} - y^2 + \cancel{y} - 1) - (\cancel{y^3} - 2y^2 + \cancel{y})$$

$$= y^2 - 1$$

$$\Rightarrow y \in \{\pm 1\}.$$

$$\Rightarrow \|G^T\| = \sqrt{1 + \frac{4}{2}} = \boxed{\sqrt{3}}$$

4

a) Give an example of a 2×2 matrix A

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

such that the matrix $B = \begin{pmatrix} 0 & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ has norm $\|B\| > \|A\|$. (Feel free to use Matlab.)

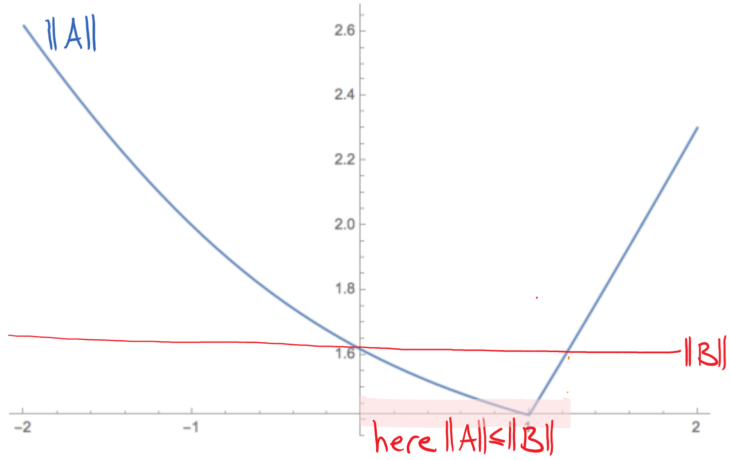
a) I ran the following in Matlab a few times until it found an example:

```
>> A = randn(2,2); norm(A)

ans =

    1.0227

>> B = A; B(1,1) = 0; norm(B)

ans =

    1.0236 >||A||

>> A

A =

    0.2821    0.2182
   -0.0465    0.9990
```

Playing around some more, here's one construction I found that works:

$$A = \begin{pmatrix} a & 1 \\ -1 & 1 \end{pmatrix} \text{ for } 0 < a < \sqrt{5} - 1.$$

```
Plot[Norm[( a  1 ; -1  1 )], {a, -2, 2}]
```



b) If the entries a_{ij} are all ≥ 0, prove that $\|B\| \leq \|A\|$.

Proof:

$$\|A\| = \sqrt{\max_{\vec{x} \neq \vec{0}} \frac{\|A\vec{x}\|^2}{\|\vec{x}\|^2}}$$

The numerator is $(a_{11}x_1 + a_{12}x_2)^2 + (a_{21}x_1 + a_{22}x_2)^2$.
Since all $a_{ij} \geq 0$ the maximum will be at a point with $x_1 \geq 0$ and $x_2 \geq 0$.
Similarly for B. But if $x_1, x_2 \geq 0$ then

$$(a_{12}x_2)^2 + (a_{21}x_1 + a_{22}x_2)^2 \leq (a_{11}x_1 + a_{12}x_2)^2 + (a_{21}x_1 + a_{22}x_2)^2,$$

so $\|B\| \leq \|A\|$.

⑤ Solve parts a,b,c,d in order.
Let $n = 50$ and let A be the $n \times n$ matrix given by

$$A_{ij} = \begin{cases} 0 & \text{if } i = j = 1 \\ r^{i+j-2} & \text{otherwise,} \end{cases}$$

where $r = 1 - \frac{1}{1000}$.

Here's some Matlab code to generate A:

```
n = 50;
r = 1 - 1/1000;
v = r .^ (0:n-1)';
A = v * v';
A(1,1) = 0;
```

```
A = v * v';
A(1,1) = 0;
```

Let $S = \sum_{i=0}^{n-1} r^{2i} = \frac{1-r^{2n}}{1-r^2} \approx 47.6277$.

a) Use the triangle inequality to argue that
$$S-1 \le \|A\| \le S+1$$
Can you give a simple argument that $\|A\| \le S$?

a) Let $\vec{v} = \begin{pmatrix} 1 \\ r \\ r^2 \\ \vdots \\ r^{n-1} \end{pmatrix}$, $B = \vec{v}\vec{v}^T$ and $C = \vec{e}_1\vec{e}_1^T$.

These are both rank-one matrices, with $\|B\| = S$, $\|C\| = 1$.
$$A = B - C$$
$$\Rightarrow \|B\| - \|C\| \le \|A\| \le \|B\| + \|C\|. \checkmark$$

Note that
$$\|A\| = \max_{\substack{\vec{x} \in \mathbb{R}^n \\ \vec{x} \ne \vec{0}}} \frac{\|A\vec{x}\|}{\|\vec{x}\|}.$$

The optimal $\vec{x}$ will have all coefficients $x_1, \ldots, x_n > 0$, to avoid cancellation. Then
$$\|A\vec{x}\|^2 \le \|B\vec{x}\|^2,$$
so $\|A\| \le \|B\|$. $\checkmark$

b) Implement the following pseudocode in Matlab, and use it to lower-bound $\|A\|$:

repeat 10^6 times:
  $\vec{v} \leftarrow$ random vector
  lower bound $\leftarrow \frac{\|A\vec{v}\|}{\|\vec{v}\|}$
output best bound found

b) This is what I got:

```
bestsofar = 0;
for i = 1:10^6
  v = randn(n,1);
  ratio = norm(A*v) / norm(v);
  bestsofar = max(bestsofar, ratio);
end
bestsofar

bestsofar =
```

$\|A\| \ge$ 29.6134   It's not a very good bound!

c) Prove that $\|A\| \ge 47.6$

Do this by finding a vector $\vec{v}$ with $\frac{\|A\vec{v}\|}{\|\vec{v}\|} \ge 47.6$

c) A is a perturbation of the matrix $B = \vec{v}\vec{v}^T$ from part ⓐ. That vector $\vec{v} = \begin{pmatrix} 1 \\ r \\ r^2 \\ \vdots \\ r^{n-1} \end{pmatrix}$ maximized $\frac{\|B\vec{v}\|}{\|\vec{v}\|} = S$.

from part ⓐ. That vector $\vec{v} = \begin{pmatrix} 1 \\ r \\ r^2 \\ \vdots \\ r^{n-1} \end{pmatrix}$ maximized $\frac{\|B\vec{v}\|}{\|\vec{v}\|} = S$.

So it makes sense to try the same vector with A:

```
>> v = r .^ (0:n-1)';
>> norm(A*v)/norm(v)

ans =

   47.6070
```

We have therefore proved that

$$47.607 \leq \|A\| \leq \|B\| = S \leq 47.628.$$

The norm is actually

```
>> norm(A)

ans =

   47.6072
```

It is easy to compute exactly because $\mathrm{rank}(A) = 2$.

In the orthonormal basis for $R(A)$ $\vec{e}_1, \frac{1}{\sqrt{T}} \sum_{j=2}^{n} r^{j-1} \vec{e}_j$

$$T = \sum_{j=2}^{n} r^{2j-2} = r^2 \frac{1 - r^{2(n-1)}}{1 - r^2}$$

A is given by $\begin{pmatrix} 0 & \sqrt{T} \\ \sqrt{T} & T \end{pmatrix}$

for which the norm is $\frac{1}{\sqrt{2}} \sqrt{T} \sqrt{2 + T + \sqrt{T(4+T)}}$.

⑥ Using the definition $\|A\| = \max_{\vec{x} \neq 0} \frac{\|A\vec{x}\|}{\|\vec{x}\|}$,

prove that for any invertible matrix A,

$$\|A\| = \frac{1}{\min_{y : \|y\| = 1} \|A^{-1} y\|}.$$

Answer:

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \quad \text{(definition)}$$

$$= \max_{y \neq 0} \frac{\|y\|}{\|A^{-1}y\|} \quad \text{(since } A \text{ is invertible, } A^{-1}y \text{ varies over all } x \neq 0\text{)}$$

$$= \max_{y : \|y\| = 1} \frac{1}{\|A^{-1}y\|} \quad \text{(by linearity, it is enough to vary over } y \text{ with } \|y\| = 1\text{)}$$

$$= \frac{1}{\min_{y : \|y\| = 1} \|A^{-1}y\|} \quad \text{(moving the max into the denominator)}$$

✓

⑦ Recall the $n \times n$ matrix

$$A_n \equiv \begin{pmatrix} -2 & 1 & & & 1 \\ 1 & -2 & 1 & 0 & \\ & 1 & -2 & 1 & \\ & & 1 & -2 & 1 \end{pmatrix}.$$

$$A_n \equiv \begin{pmatrix} -2 & 1 & & & & 1 \\ 1 & -2 & 1 & & 0 & \\ & 1 & -2 & 1 & & \\ & & 1 & -2 & 1 & \\ & 0 & & \ddots & & \\ & & & 1 & -2 & 1 \\ 1 & & & & 1 & -2 \end{pmatrix}.$$

(This matrix arose as a discretization of the second derivative, on an interval with periodic boundary conditions.) In this problem, you will solve for $\|A_n\|$, at least when n is even.

ⓐ Using Matlab or similar software, determine numerically the norms of

$$A_4 = \begin{pmatrix} -2 & 1 & 0 & 1 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 1 & 0 & 1 & -2 \end{pmatrix}$$

$$A_5 = \begin{pmatrix} -2 & 1 & 0 & 0 & 1 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 1 & 0 & 0 & 1 & -2 \end{pmatrix}$$

$$A_6 = \begin{pmatrix} -2 & 1 & 0 & 0 & 0 & 1 \\ 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 1 & 0 & 0 & 0 & 1 & -2 \end{pmatrix}$$

(You do not have to prove that your answers are correct.)

Answer (in Matlab/Octave):

```
A4 = [-2 1 0 1; 1 -2 1 0; 0 1 -2 1; 1 0 1 -2]
A5 = [-2 1 0 0 1; 1 -2 1 0 0; 0 1 -2 1 0; 0 0 1 -2 1; 1 0 0 1 -2]
A6 = [-2 1 0 0 0 1; 1 -2 1 0 0 0; 0 1 -2 1 0 0; 0 0 1 -2 1 0; 0 0 0 1 -2 1; 1 0 0 0 1 -2]

norm(A4) → 4
norm(A5) → 3.6180
norm(A6) → 4
```

this does the same for n up to 20

```
for n = 3:20
  A = zeros(n,n);
  for i = 1:n
    A(i,i) = -2;
    A(i, 1+mod(i,n)) = 1;
    A(i, 1+mod(i-2,n)) = 1;
  end;
  disp([n,norm(A)]);
end;
```

these are equivalent ways of building up the same matrix A_n

```
for n = 3:20
  diagonal = -2 * eye(n);
  abovediagonal = diag(ones(n-1,1), 1) + diag(1, 1-n);
```

```
for n = 3:20
  A = zeros(n,n);
  for i = 1:n
    A(i,i) = -2;
    A(i, 1+mod(i,n)) = 1;
    A(i, 1+mod(i-2,n)) = 1;
  end;
  disp([n,norm(A)]);
end;
```

these are equivalent ways of building up the same matrix A_n

```
for n = 3:20
  diagonal = -2 * eye(n);
  abovediagonal = diag(ones(n-1,1), 1) + diag(1, 1-n);
  belowdiagonal = abovediagonal';
  A = diagonal + abovediagonal + belowdiagonal;
  disp([n, norm(A)]);
end;
```

Answer: In Mathematica:

$$A_4 = \begin{pmatrix} -2 & 1 & 0 & 1 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 1 & 0 & 1 & -2 \end{pmatrix};$$

$$A_5 = \begin{pmatrix} -2 & 1 & 0 & 0 & 1 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 1 & 0 & 0 & 1 & -2 \end{pmatrix};$$

$$A_6 = \begin{pmatrix} -2 & 1 & 0 & 0 & 0 & 1 \\ 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 1 & 0 & 0 & 0 & 1 & -2 \end{pmatrix};$$

```
Norm[A4]
Norm[A5]
Norm[A6]
```

4

$$\sqrt{\frac{5}{2}\left(3+\sqrt{5}\right)}$$

4

```
Table[
  A = ConstantArray[0, {n, n}];
  For[i = 1, i ≤ n, i++,
   A[[i, i]] = -2;
   A[[i, 1 + Mod[i, n]]] = 1;
   A[[i, 1 + Mod[i - 2, n]]] = 1;
  ];
  {n, Norm[A // N]}
  ,
  {n, 3, 20}
 ] // MatrixForm
```

atrixForm=

n	$\|A_n\|$
3	3.
4	4.
5	3.61803
6	4.
7	3.80194
8	4.
9	3.87939
10	4.
11	3.91899
12	4.
13	3.94188
14	4.
15	3.9563
16	4.
17	3.96595
18	4.
19	3.97272
20	4.

Assuming that n is even:

b) Permute the rows and columns of A_n like this

$$\begin{array}{c|c} & \text{odd columns } 1,3,5,7,\dots \qquad \text{even columns } 2,4,6,8,\dots \\ \begin{array}{c}\text{odd rows} \\[2em] \text{even rows}\end{array} & \left(\begin{array}{c|c} & \\[2em] \hline & \\[2em] \end{array}\right) \end{array}$$

This doesn't change the norm!

Find vector $\vec{x}$ so that $\|A\vec{x}\| = 4$

This doesn't change the norm!

Find a vector $\vec{v}$ so that $\frac{\|A_n\vec{v}\|}{\|\vec{v}\|} = 4.$

Conclude that $\|A_n\| \geq 4.$

b)

$$\text{permuted matrix} = \left(\begin{array}{ccc|ccc} -2 & & 0 & 1 & & 0 \\ & \ddots & & 1 & \ddots & \\ 0 & & -2 & 0 & 1 & 1 \\ \hline 1 & & 0 & -2 & & 0 \\ 1 & \ddots & & & \ddots & \\ 0 & 1 & 1 & 0 & & -2 \end{array}\right)$$

By symmetry, $\vec{v}$ should be constant on each half, so either

$$\vec{u} = (1, 1, 1, \ldots, 1) \quad \text{or} \quad \vec{v} = (\overset{\text{odd}}{1}, \overset{\text{even}}{-1}, \overset{\text{odd}}{1}, \overset{\text{even}}{-1}, \ldots, 1, -1)$$

should be good.

The latter vector is better, with $\|\vec{v}\| = \sqrt{n}$ and

$$A_n\vec{v} = \begin{pmatrix} -4 \\ -4 \\ \vdots \\ -4 \end{pmatrix} \quad \Longrightarrow \frac{\|A_n\vec{v}\|}{\|\vec{v}\|} = 4. \checkmark$$

(d) What is the norm of a permutation matrix, e.g.,

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} ? \quad \text{Why?}$$

Permutation matrices are isometries; they don't change any lengths. Therefore they'll always have norm 1.

(e) Now prove an upper bound on $\|A_n\|$ that matches the lower bound you gave in part (c).

This finishes the calculation of $\|A_n\|$ (for n even).

(Hint: "Break A_n into pieces." That is, write A_n as the sum of three or four permutation matrices, possibly with weights, and then use the triangle inequality $\|B+C\| \leq \|B\| + \|C\|$.)

Answer:

Use the same decomposition we used in part (a):

$$A_n = -2I + \begin{pmatrix} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & 0 & 1 & & \\ & & & 0 & 1 & \\ & & & & \ddots & \\ 1 & & & & & 0 \end{pmatrix} + \begin{pmatrix} 0 & & & & & 1 \\ 1 & 0 & & & & \\ & 1 & 0 & & & \\ & & 1 & \ddots & & \\ & & & & 0 & \\ & & & & 1 & 0 \end{pmatrix}$$

The identity and the latter two matrices are all permutations, so have norm 1.

$\Rightarrow A_n \leq 2\cdot 1 + 1 + 1 = 4$ ✓

$\Rightarrow$ If n is even, $\|A_n\| = 4$.

(f) [Optional] What about $\|A_n\|$ for n odd?

This is slightly trickier. Observe, though, that A_n is symmetrical under cyclic shifts of its row and columns (because of the banded structure).

One can argue that this implies A_n is diagonal in the Fourier basis.

Experimenting with Mathematica...

```
An[n_] := Module[{A},
   A = ConstantArray[0, {n, n}];
   For[i = 1, i ≤ n, i++,
    A[[i, i]] = -2;
    A[[i, 1 + Mod[i, n]]] = 1;
    A[[i, 1 + Mod[i - 2, n]]] = 1;
   ];
   A
  ];
```

function to create A_n

```
Fn[n_] := 1/Sqrt[n] Table[e^(2 π i/n j k), {j, 0, n - 1}, {k, 0, n - 1}];
```

function to create Fourier basis-change matrix

```
Adjoint[M_] := Conjugate[Transpose[M]];

n = 7;
Fn[n].An[n].Adjoint[Fn[n]] // N // Chop // MatrixForm
```

$\leftarrow$ change basis with FAF^{-1}

atrixForm=

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.75302 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2.44504 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3.80194 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -3.80194 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2.44504 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.75302 \end{pmatrix}$$

yes, it is diagonal, and these are the largest magnitude values

$\Rightarrow \|A_7\| = 3.80194$

It appears that to maximize the stretch, we should set $\vec{u}$ equal to one of the middle two Fourier vectors, e.g., if $n = 2k+1$,

$$\vec{u} = \frac{1}{\sqrt{n}}\left(1, \omega^k, \omega^{2k}, \omega^{3k}, \ldots, \omega^{(n-1)k}\right)$$

where $\omega = \exp(2\pi i/n)$.

Then
$$\|A_n \vec{u}\|^2 = n \cdot |e_1^T A_n \vec{u}|^2 \quad \text{by symmetry}$$
$$= |-2 + \omega^k + \omega^{-k}|^2$$
$$= \left(2 + 2\cos\frac{\pi}{n}\right)^2$$

$$\Rightarrow \|A_n\| = 2\left(1 + \cos\frac{\pi}{n}\right)$$
$$= 4 - \frac{\pi^2}{n^2} + O\left(\frac{1}{n^4}\right).$$

$\Rightarrow \|A_n\| = 2(1+\cos\frac{\pi}{n})$
$= 4 - \frac{\pi^2}{n^2} + O(\frac{1}{n^4})$.

(To formally prove this would take a little more work.)

SVD

(8) Compute the singular-value decompositions of the following matrices. Do not use a computer! Don't use eigenvalues either.

(a) $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$

(b) $B = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 4 & 0 & 0 \end{pmatrix}$

(c) $C = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 4 \end{pmatrix}$

(d) $D = \begin{pmatrix} 0 & 0 & 5 \\ \sqrt{\frac{2}{3}} & -\sqrt{\frac{1}{3}} & 0 \\ \sqrt{\frac{1}{3}} & \sqrt{\frac{2}{3}} & 0 \end{pmatrix}$

(e) $E = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & i & 0 \\ 0 & i & 1 & 0 \\ 2 & 0 & 0 & 1 \end{pmatrix}$

What are the norms of the matrices?

Answer:

(a) A is a permutation matrix, so its norm is 1, all its singular values are 1, and you can read off its left and right singular vectors:

$$Ae_1 = e_3, \quad Ae_2 = e_1, \quad Ae_3 = e_2$$

$\Rightarrow$ right singular vectors are e_1, e_2, e_3 and the corresponding left singular vectors are, respectively, e_3, e_1, e_2.

In matrix form, the SVD of A is

$$A = \begin{pmatrix} | & | & | \\ e_3 & e_1 & e_2 \\ | & | & | \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} — e_1 — \\ — e_2 — \\ — e_3 — \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad ✓$$

Note: The singular vectors are not unique. If
$\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$
is any orthonormal basis for $\mathbb{R}^3$, then they can be taken as right singular vectors, with
$A\vec{v}_1, A\vec{v}_2, A\vec{v}_3$
the corresponding left singular vectors (also orthonormal).

(b) $B = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 4 & 0 & 0 \end{pmatrix} = 4 e_3 e_1^T + 3 e_2 e_3^T + 2 e_1 e_2^T$

$\|B\| = 4$

(c) $C = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 4 \end{pmatrix} = 2 e_1 e_2^T + 5 \begin{pmatrix} 0 \\ 3/5 \\ 4/5 \end{pmatrix} e_3^T + 0 \cdot \begin{pmatrix} 0 \\ 4/5 \\ -3/5 \end{pmatrix} e_1^T$

$\|C\| = 5$

(d) $D = \begin{pmatrix} 0 & 0 & 5 \\ \sqrt{\frac{2}{3}} & -\sqrt{\frac{1}{3}} & 0 \\ \sqrt{\frac{1}{3}} & \sqrt{\frac{2}{3}} & 0 \end{pmatrix}$

Observe that $\frac{1}{\sqrt{3}}\begin{pmatrix} \sqrt{2} & -1 \\ 1 & \sqrt{2} \end{pmatrix}$ is unitary, so an SVD for it is
$$1 \cdot \begin{pmatrix} \sqrt{2/3} \\ \sqrt{1/3} \end{pmatrix} e_1^T + 1 \cdot \begin{pmatrix} -\sqrt{\frac{1}{3}} \\ \sqrt{\frac{2}{3}} \end{pmatrix} e_2^T$$

Therefore an SVD for D is
$$D = 5 e_1 e_3^T + 1 \cdot \begin{pmatrix} 0 \\ \sqrt{2/3} \\ \sqrt{1/3} \end{pmatrix} e_1^T + 1 \cdot \begin{pmatrix} 0 \\ -1/\sqrt{3} \\ \sqrt{2/3} \end{pmatrix} e_2^T$$

$\|D\| = 5$

(e) $E = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & i & 0 \\ 0 & i & 1 & 0 \\ 2 & 0 & 0 & 1 \end{pmatrix}$

Sitting inside E are the noninteracting block submatrices
$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = 3 \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}}(1 \;\; 1) - 1 \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}}(1 \;\; -1)$$

and $\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$ ← This is $\sqrt{2}$ times a unitary matrix, so an SVD is $\sqrt{2}\,\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix} e_1^T + \sqrt{2}\,\frac{1}{\sqrt{2}}\begin{pmatrix} i \\ 1 \end{pmatrix} e_2^T$

Therefore an SVD for E is
$$E = 3 \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \frac{1}{\sqrt{2}}(1 \; 0 \; 0 \; 1) - 1 \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}}(1 \; 0 \; 0 \; -1)$$
$$+ \sqrt{2} \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ i \\ 0 \end{pmatrix}(0 \; 1 \; 0 \; 0) + \sqrt{2}\,\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ i \\ 1 \\ 0 \end{pmatrix}(0 \; 0 \; 1 \; 0)$$

$\|E\| = 3$.

$\|E\| = 3.$

⑨ Compute SVDs and norms for the matrices

ⓕ $F = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$

ⓖ $G = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 1 \end{pmatrix}$

It is okay to use a computer. However, please give **exact** answers (not just numerical).

Answer:

Matlab will give numerical answers, from which you can easily figure out the exact numbers (e.g., by squaring each number).

```
octave:3> [U,D,V] = svd([1 1 -1; 0 1 1; 1 0 1])
U =

  -8.1650e-01  -5.7735e-01   1.1102e-16
  -4.0825e-01   5.7735e-01  -7.0711e-01
  -4.0825e-01   5.7735e-01   7.0711e-01

D =

Diagonal Matrix

   1.7321        0        0
        0   1.7321        0
        0        0   1.0000

V =

  -0.70711   0.00000   0.70711
  -0.70711  -0.00000  -0.70711
  -0.00000   1.00000  -0.00000
```

```
octave:4> [U,D,V] = svd([1 -2 0; 0 1 1])
U =

  -0.89443   0.44721
   0.44721   0.89443

D =

Diagonal Matrix

   2.4495        0        0
        0   1.0000        0

V =

  -3.6515e-01   4.4721e-01  -8.1650e-01
   9.1287e-01  -1.1102e-16  -4.0825e-01
   1.8257e-01   8.9443e-01   4.0825e-01
```

But Mathematica will give the exact answers directly:

A

MatrixForm /@ SingularValueDecomposition$\left[\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}\right]$ // FullSimplify

$= F$

MatrixForm /@ SingularValueDecomposition[{{1, 1, -1}, {0, 1, 1}, {1, 0, 1}}] // FullSimplify

MatrixForm /@ SingularValueDecomposition[{{1, -2, 0}, {0, 1, 1}}] // FullSimplify

$= G$

$$\left\{\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\right\}$$

left singular vectors

right singular vectors

$$\left\{\begin{pmatrix} -\frac{1}{\sqrt{3}} & \sqrt{\frac{2}{3}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{pmatrix}\right\}$$

$\Rightarrow \|F\| = \sqrt{3}$

(largest singular value)

...values

$$\left\{ \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} 0 & \sqrt{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{pmatrix} \right\}$$

$\Rightarrow \|F\| = \sqrt{3}$

(largest singular value)

singular values

$$\left\{ \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}, \begin{pmatrix} \sqrt{6} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} -\sqrt{\frac{2}{15}} & \frac{1}{\sqrt{5}} & -\sqrt{\frac{2}{3}} \\ \sqrt{\frac{5}{6}} & 0 & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \end{pmatrix} \right\}$$

$\Rightarrow \|G\| = \sqrt{6}$

WolframAlpha computational intelligence. https://www.wolframalpha.com/

SingularValueDecomposition[{{1,-2,0},{0,1,1}}]

Extended Keyboard Upload Examples Random

Input:

singular value decomposition $\begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 1 \end{pmatrix}$

Result: Approximate forms

$M = U.\Sigma.V^{\dagger}$

where

$$M = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

$$U = \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \sqrt{6} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$V = \begin{pmatrix} -\sqrt{\frac{2}{15}} & \frac{1}{\sqrt{5}} & -\sqrt{\frac{2}{3}} \\ \sqrt{\frac{5}{6}} & 0 & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$

$m^{\dagger}$ gives the conjugate transpose of m