Homework #15

Not collected: this is practice for the final exam

1. For each matrix A of HW14#1:

$$\textbf{a)} \begin{pmatrix} 8 & 4 \\ 1 & 13 \end{pmatrix} \qquad \textbf{b)} \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \qquad \textbf{c)} \begin{pmatrix} -3 & 11 \\ 10 & -2 \\ 1 & 5 \\ -4 & 6 \end{pmatrix}$$

$$\textbf{d)} \begin{pmatrix} 9 & 7 & 10 & 8 \\ -13 & 1 & 5 & -6 \end{pmatrix} \qquad \textbf{e)} \begin{pmatrix} 3 & 7 & 1 & 5 \\ 3 & 1 & 7 & 5 \\ 6 & 2 & 2 & -2 \end{pmatrix}$$

find the singular value decomposition in the matrix form

$$A = U\Sigma V^T.$$

2. For each matrix A of Problem 1, write down orthonormal bases for all four fundamental subspaces. (This can be read off from your answers to Problem 1.)

3. Let S be a symmetric matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Let $S = QDQ^T$ be an orthogonal diagonalization of S, where D has diagonal entries $\lambda_1, \ldots, \lambda_n$.

Show that $S = QDQ^T$ is a singular value decomposition if and only if S is positive-semidefinite. [See HW14#6.]

4. Let A be a square, invertible matrix with singular values $\sigma_1, \ldots, \sigma_n$.

a) Show that A^{-1} has the same singular vectors as A^T, with singular values $\sigma_n^{-1} \geq \cdots \geq \sigma_1^{-1}$.

[**Hint:** Invert $A = U\Sigma V^T$.]

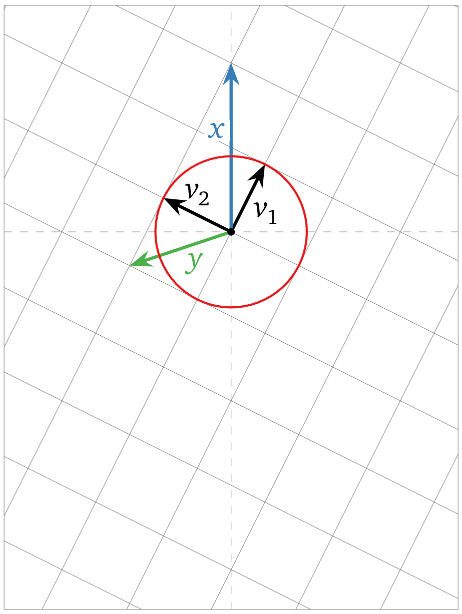
b) Let λ be an eigenvalue of A. Use HW14#9(c) and **a)** to show that $\sigma_n \leq |\lambda|$.

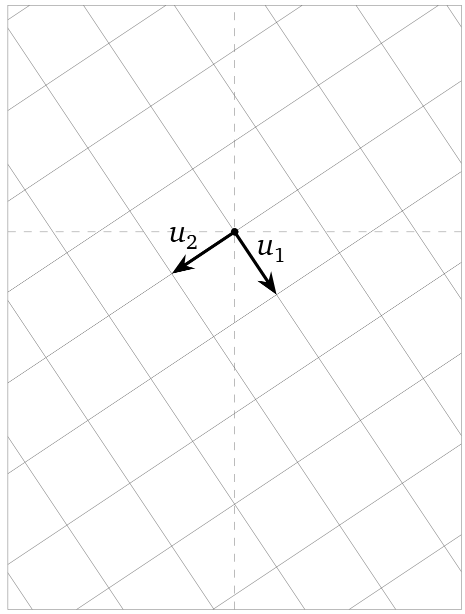
It follows that the absolute values of all eigenvalues of A are contained in the interval $[\sigma_n, \sigma_1]$. Compare HW14#5.

5. A certain 2×2 matrix A has singular values $\sigma_1 = 2$ and $\sigma_2 = 1.5$. The right-singular vectors v_1, v_2 and the left-singular vectors u_1, u_2 are shown in the pictures below.

a) Draw Ax and Ay in the picture on the right.

b) Draw $\{Ax : \|x\| = 1\}$ (what you get by multiplying all vectors on the unit circle by A) in the picture on the right.

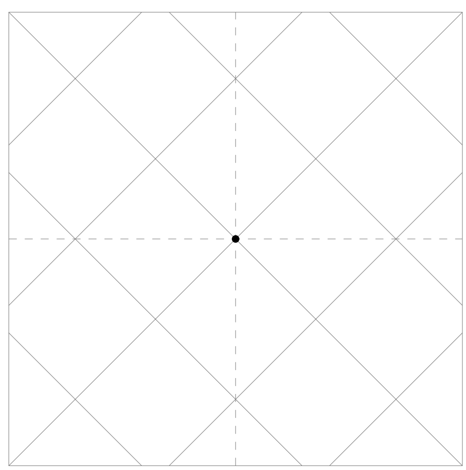




6. Consider the following 3×2 matrix A and its SVD:

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix}^T .$$

Draw $\{Ax : \|x\| = 1\}$ (what you get by multiplying all vectors on the unit sphere by A) in the picture on the right.



7. Compute the pseudoinverse of each matrix of Problem 1.

8. Consider the matrix

$$A = \begin{pmatrix} 3 & 7 & 1 & 5 \\ 3 & 1 & 7 & 5 \\ 6 & 2 & 2 & -2 \end{pmatrix}$$

of Problem 7(e). Find the matrix P for projection onto the row space of A in two ways:

a) Multiply out $P = A^+A$.

b) In Problem 2 you found $\mathrm{Nul}(A) = \mathrm{Span}\{v\}$ for $v = (1, -1, -1, 1)$. Compute $P_\perp = vv^T/v \cdot v$ and $P = I_4 - P_\perp$.

Your answers to **a)** and **b)** should be the same, of course!

9. Let A be an $m \times n$ matrix.

a) If A has full column rank, show that $A^+A = I_n$.

b) If A has full row rank, show that $AA^+ = I_m$.

In particular, a matrix with full column rank admits a *left inverse*, and a matrix with full row rank admits a *right inverse*. Compare HW5#15.

10. Let A be a matrix and let A^+ be its pseudoinverse. Match the subspaces on the left to the subspaces on the right:

$\mathrm{Col}(A)$	$\mathrm{Col}(A^+)$
$\mathrm{Nul}(A)$	$\mathrm{Nul}(A^+)$
$\mathrm{Row}(A)$	$\mathrm{Row}(A^+)$
$\mathrm{Nul}(A^T)$	$\mathrm{Nul}((A^+)^T)$

What is the rank of A^+?

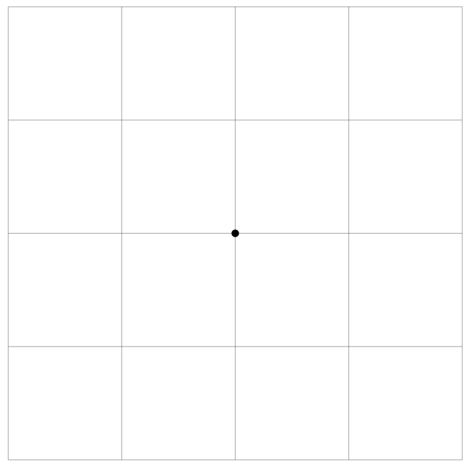
11. What is the pseudoinverse of the $m \times n$ zero matrix?

12. Consider the matrix $A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$ of Problem 7(b).

a) Find all least-squares solutions of $Ax = \binom{3}{1}$ in parametric vector form.

b) Find the shortest least-squares solution $\widehat{x} = A^{+}\binom{3}{1}$.

c) Draw your answers to **a)** and **b)** on the grid below.



13. Consider the following matrix holding 5 samples of 2 measurements each:

$$A_0 = \begin{pmatrix} 22 & -12 & 24 & -29 & 20 \\ 1 & -11 & 37 & -17 & -35 \end{pmatrix}.$$

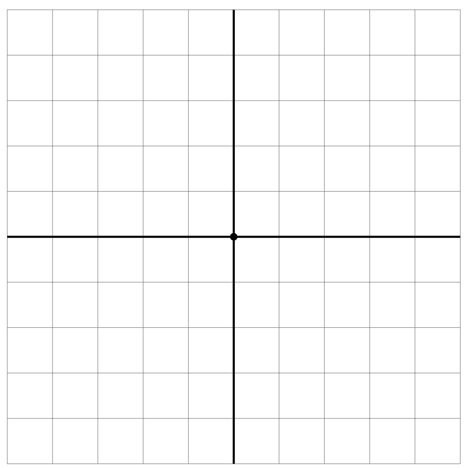
a) Subtract the means of the rows of A_0 to obtain the centered matrix A.

b) Compute the covariance matrix $S = \frac{1}{5-1}AA^T$. What is the total variance? What is the covariance of the first row with the second?

c) Find the eigenvalues λ_1, λ_2 and unit eigenvectors v_1, v_2 of S. What line best approximates the columns of A? What line best approximates the columns of A_0?

d) Find the orthogonal projections of the columns of A onto this line by computing the first summand of the SVD of A (in vector form). (Don't forget to rescale by $\sqrt{5-1}$.)

e) Draw the columns of A, the first best-fit line you found in **c)**, and the orthogonal projections you found in **d)** on the grid below. (Grid marks are 10 units apart.)



14. An online movie-streaming service collects star ratings from its viewers and uses these to predict what movies you will like based on your previous ratings. The following are the ratings that ten (fictitious) people gave to three (fictitious) movies, on a scale of 0–10:

	Abe	Amy	Ann	Ben	Bob	Eve	Dan	Don	Ian	Meg
Prognosis Negative	7.8	6.1	2.4	9.8	10	3.0	6.3	3.6	6.7	6.3
Ponce De Leon	6.0	7.9	6.4	8.1	7.1	6.4	7.3	7.9	6.2	8.1
Lenore's Promise	5.8	8.2	8.2	6.8	6.2	8.7	7.3	9.2	6.8	8.2

Using SymPy (in the Sage cell on the webpage) or your favorite linear algebra calculator, put the data into a matrix:

```
A0 = Matrix([[7.8, 6.1, 2.4, 9.8,  10, 3.0, 6.3, 3.6, 6.7, 6.3],
             [6.0, 7.9, 6.4, 8.1, 7.1, 6.4, 7.3, 7.9, 6.2, 8.1],
             [5.8, 8.2, 8.2, 6.8, 6.2, 8.7, 7.3, 9.2, 6.8, 8.2]])
```

Find the row averages and subtract them:

```
 # Multiplying by (1,1,...,1) sums the rows
averages = A0 * Matrix.ones(10,1)/10
A = A0 - averages * Matrix.ones(1, 10)
```

Now compute the covariance matrix:

```
S = A*A.transpose() / (10-1)
from sympy import pprint
pprint(S)
```

In this problem, please write your answers to two decimal places.

a) What is the variance in the number of stars given each of the three movies? What is the total variance? (Use `S.trace()`)

b) Which entry of S tells you that people who liked *Prognosis Negative* generally did not like *Lenore's Promise*?

Let us compute the eigenvalues of S in order, and the corresponding unit eigenvectors:

```
[(sigma3sq, v3), (sigma2sq, v2), (sigma1sq, v1)] \
    = sorted([(x[0], x[2][0]) for x in S.eigenvects()])
  # Verify the sum is equal to the total variance
print(sigma1sq + sigma2sq + sigma3sq, S.trace())
  # Compute unit eigenvectors
pprint([v1.normalized(), v2.normalized(), v3.normalized()])
```

c) Which is the direction with the most variance? What is the variance in that direction?

d) Explain how these calculations tell you that 68% of the ratings are at a distance of $\sigma_3 \approx 0.18$ stars from the plane $\mathrm{Span}\{v_1, v_2\}$.

e) Use the fact that $\{v_1, v_2, v_3\}$ is orthonormal to find an implicit equation for $\mathrm{Span}\{v_1, v_2\}$ of the form $x_3 = a_1x_1 + a_2x_2$.

f) Suppose that Joe gave *Prognosis Negative* a rating of 8.5 and *Ponce De Leon* a rating of 6.2. How would you expect Joe to rate *Lenore's Promise*?

Remark: According to a New York Times Magazine article, this really is the idea behind Netflix's algorithm—which earned its creator a \$1 000 000 prize.

15. Let A be a matrix with singular value decomposition

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T.$$

Recall from HW14#9 that $\|Ax\|/\|x\|$ is maximized at $x = v_1$ with maximum value σ_1.

a) Show that the maximum value of $\|Ax\|/\|x\|$ subject to the condition $x \cdot v_1 = 0$ is equal to σ_2, and is achieved at $x = v_2$.
[**Hint:** If $x \cdot v_1 = 0$ then $Ax = A'x$ for $A' = \sigma_2 u_2 v_2^T + \sigma_3 u_3 v_3^T + \cdots + \sigma_r u_r v_r^T$.]

b) More generally, show that the maximum value of $\|Ax\|/\|x\|$ subject to the conditions $x \cdot v_1 = 0,\ x \cdot v_2 = 0,\ \ldots,\ x \cdot v_j = 0$ is equal to σ_{j+1}, and is achieved at $x = v_{j+1}$.

c) If A has full column rank, show that the *minimum* value of $\|Ax\|/\|x\|$ is equal to σ_r, and is achieved at $x = v_r$.

In the language of principal component analysis, this says that v_2 is the direction of *second-largest variance*, etc.

16. Decide if each statement is true or false, and explain why.

a) If A is a matrix of rank r, then A is a linear combination of r rank-1 matrices.

b) If A is a matrix of rank 1, then A^+ is a scalar multiple of A^T.

c) If $A = U\Sigma V^T$ is the SVD of A, then the SVD of A^+ is $A^+ = V\Sigma^+ U^T$.