

**INTRODUCTION
TO
LINEAR
ALGEBRA
Fifth Edition**

MANUAL FOR INSTRUCTORS

Gilbert Strang
Massachusetts Institute of Technology

math.mit.edu/linearalgebra

web.mit.edu/18.06

video lectures: ocw.mit.edu

math.mit.edu/~gs

www.wellesleycambridge.com

email: linearalgebrabook@gmail.com

Wellesley - Cambridge Press

Box 812060

Wellesley, Massachusetts 02482

Problem Set 7.1, page 370

- 1 $A = \mathbf{u}\mathbf{v}^T$ has rank 1 with $\mathbf{u}^T = \mathbf{v}^T = [1 \ 2 \ 3 \ 4]$. Those vectors have $\|\mathbf{u}\|^2 = \|\mathbf{v}\|^2 = 30$ so the SVD has a division by $\sqrt{30}$ to reach \mathbf{u}_1 and \mathbf{v}_1 . Multiply by $\sigma_1 = 30$ to recover A .

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T = 30 \frac{\mathbf{u}}{\sqrt{30}} \frac{\mathbf{v}^T}{\sqrt{30}} = U \Sigma V^T \quad (\text{1 column in } U \text{ and } V).$$

B has rank $r = 2$. The first two columns of B are independent (the pivot columns). Column 3 is a combination $2(\text{col } 2) - (\text{col } 1)$. Column 4 is $3(\text{col } 2) - 2(\text{col } 1)$:

$$B = \begin{bmatrix} 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \\ 5 & 6 & 7 & 8 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 4 \\ 4 & 5 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \end{bmatrix} \begin{matrix} (\text{col } 1)(\text{row } 1)^T \\ + \\ (\text{col } 2)(\text{row } 2)^T \end{matrix}$$

Those pivot columns come from the first half of the book: *not orthogonal!* They don't give the \mathbf{u} 's and \mathbf{v} 's of the SVD. For that we need eigenvalues and eigenvectors of $B^T B$ and BB^T .

- 2 All the singular values of I are $\sigma = 1$. We cannot leave out any of the terms $\mathbf{u}_i \cdot \mathbf{v}_i^T$ without making an error of size 1. And the matrix $A = I$ starts with size 1! None of the SVD pieces can be left out.

Notice that the SVD is $I = (U)(I)(U^T)$ so that $U = V$. The natural choice for the SVD is just $U \Sigma V^T = III$. But we could actually choose any orthogonal matrix U . (The eigenvectors of I are very far from unique—many choices! Any orthogonal matrix U holds orthonormal eigenvectors of I .)

One possible rank 5 flag with a 3 by 3 cross of zeros is $A = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & \frac{1}{2} & 1 & 1 & 1 \\ 1 & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

$$\begin{aligned}
 \mathbf{3} \quad & \begin{bmatrix} 1 & 2 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 1 & 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 2 & 2 & 2 & 2 \end{bmatrix} \\
 & \begin{bmatrix} 1 & 2 & 2 \\ 1 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \text{pivot} \\ \text{columns} \end{bmatrix} \begin{bmatrix} \text{rows} \\ \text{of } R \end{bmatrix}
 \end{aligned}$$

$$\mathbf{4} \quad BB^T = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 3 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 9 & 13 \\ 13 & 19 \end{bmatrix}. \text{ Trace } \mathbf{28}, \text{ Determinant } \mathbf{2}.$$

$$B^T B = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \\ 1 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 5 & 5 \\ 5 & 13 & 13 \\ 5 & 13 & 13 \end{bmatrix}. \text{ Trace } \mathbf{28}, \text{ Determinant } \mathbf{0}.$$

With a small singular value $\sigma_2 \approx \frac{1}{\sqrt{14}}$, B is compressible. But we don't just keep the first row and column of B . The good row \mathbf{v}_1 and column \mathbf{u}_1 are eigenvectors of $B^T B$ and BB^T .

$$\mathbf{5} \quad \text{My hand calculation produced } A^T A = \begin{bmatrix} 7 & 10 & 7 \\ 10 & 16 & 10 \\ 7 & 10 & 7 \end{bmatrix} \text{ and } \det(A^T A - \lambda I) = -\lambda^3 + 30\lambda^2 - 24\lambda.$$

This gives $\lambda = 0$ as one eigenvalue of $A^T A$ (correct). The others are :

$$\lambda^2 - 30\lambda + 24 = 0 \text{ gives } \lambda = 15 \pm \sqrt{15^2 - 24} \approx 15 \pm 14 = 29 \text{ and } 1.$$

So $\sigma_1 \approx \sqrt{29}$ and $\sigma_2 = 1$. The `svd(A)` command in MATLAB will give accurate σ 's and U and V .

6 The matrix A has trace 4 and determinant 0. So its eigenvalues are 4 and 0—*not used in the SVD!* The matrix $A^T A$ has trace 25 and determinant 0, so $\lambda_1 = 25 = \sigma_1^2$ gives $\sigma_1 = 5$.

The eigenvectors $\mathbf{v}_1, \mathbf{v}_2$ of $A^T A$ (a symmetric matrix !) are orthogonal :

$$\begin{bmatrix} 20 & 10 \\ 10 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \mathbf{25} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 20 & 10 \\ 10 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \mathbf{0} \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

Similarly AA^T has orthogonal eigenvectors $\mathbf{u}_1, \mathbf{u}_2$:

$$\begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \mathbf{25} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \mathbf{0} \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

7 Multiply both sides of $A = U\Sigma V^T$ by the matrix V to get $AV = U\Sigma$. Column by column this says that $Av_i = \sigma_i u_i$. Notice that Σ goes on the **right side** of U when we want to multiply every *column* of U by its singular value σ_i .

8 The text found $\lambda_1 = \sigma_1^2 = \frac{1}{2}(3 + \sqrt{5})$ and then $\sigma_1 = \frac{1}{2}(1 + \sqrt{5})$. Then $\sigma_1 + 1$ equals σ_1^2 .

Also $\lambda_2 = \sigma_2^2 = \frac{1}{2}(3 - \sqrt{5})$ and $\sigma_2 = \frac{1}{2}(\sqrt{5} - 1)$ and $\sigma_1 - \sigma_2 = \frac{1}{2} + \frac{1}{2} = 1$.

(**Why don't we choose** $\sigma_2 = \frac{1}{2}(1 - \sqrt{5})$?).

9 The 20 by 40 random matrices are $A = \mathbf{rand}(20, 40)$ and $B = \mathbf{randn}(20, 40)$. With those random choices the 20 rows are independent with probability 1. Notice for these *continuous probabilities*, this does not mean that the rows are **always** independent! A random determinant might be 0 even when the probability of nonzero is 1.

MATLAB again gives the singular values of a random A and B .

By averaging 100 samples you would begin to see the expected distribution of σ 's, which is highly important in "random matrix theory".

Problem Set 7.2, page 379

1 $A = \begin{bmatrix} 0 & 4 \\ 0 & 0 \end{bmatrix}$ has eigenvalues 0 and 0; $A^T A = \begin{bmatrix} 0 & 0 \\ 0 & 16 \end{bmatrix}$ has eigenvalues $\lambda = 16$ and 0. Then $\sigma_1(A) = \sqrt{16} = 4$. The eigenvectors of $A^T A$ and AA^T are the columns of $V = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

$$\text{Then } U\Sigma V^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ 0 & 0 \end{bmatrix} = A.$$

$$A = \begin{bmatrix} 0 & 4 \\ 1 & 0 \end{bmatrix} \text{ gives } A^T A = \begin{bmatrix} 1 & 0 \\ 0 & 16 \end{bmatrix} \text{ with } \lambda_1 = 16 \text{ and } \lambda_2 = 1. \text{ Same } U \text{ and } V.$$

$$\text{Then } U\Sigma V^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ 1 & 0 \end{bmatrix} = A.$$

$$\mathbf{2} \quad A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} \text{ leads to } A^T A = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix} \text{ with eigenvectors in } V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

$$\sigma_1^2 = 8 \quad \mathbf{u}_1 = \frac{A\mathbf{v}_1}{\sigma_1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 4 \\ 0 \end{bmatrix} \frac{1}{\sigma_1} \text{ has unit vector } \mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \sigma_1 = 2\sqrt{2}$$

$$\sigma_2^2 = 2 \quad \mathbf{u}_2 = \frac{A\mathbf{v}_2}{\sigma_2} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 2 \end{bmatrix} \frac{1}{\sigma_2} \text{ has unit vector } \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ and } \sigma_2 = \sqrt{2}$$

$$\text{The full SVD is } A = U\Sigma V^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2\sqrt{2} & \\ & \sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} / \sqrt{2}.$$

$$\mathbf{3} \quad \text{Problem 7.2.2 happens to have } AA^T = \text{diagonal matrix } \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix}. \text{ So its eigenvectors}$$

$(1, 0)$ and $(0, 1)$ go in $U = I$. Its eigenvalues are $\sigma_1^2 = 8$ and $\sigma_2^2 = 2$. The rows of A are orthogonal but not *orthonormal*. So $A^T A$ is not diagonal and V is not I .

$$\mathbf{4} \quad AA^T = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \text{ has } \sigma_1^2 = 3 \text{ with } \mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \text{ and } \sigma_2^2 = 1 \text{ with } \mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}.$$

$$A^T A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \text{ has } \sigma_1^2 = 3 \text{ with } \mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \sigma_2^2 = 1 \text{ with } \mathbf{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$$

$$\text{and } \mathbf{v}_3 = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}. \text{ Then } \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = U\Sigma.$$

- 5 (a) $A = \begin{bmatrix} 1 & 1 \\ 3 & 3 \end{bmatrix}$ has $\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ in its row space and $\mathbf{u}_1 = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ in its column space. Those are unit vectors.

Since $A^T A = \begin{bmatrix} 10 & 10 \\ 10 & 10 \end{bmatrix}$ has $\lambda_1 = 20$ and $\lambda_2 = 0$, A itself has $\sigma_1 = \sqrt{20}$ and has no σ_2 . (Remember that the r singular values have to be strictly positive!)

- (b) If we want square matrices U and V , choose \mathbf{u}_2 and \mathbf{v}_2 orthogonal to \mathbf{u}_1 and \mathbf{v}_1 :

$$U = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & 3 \\ 3 & -1 \end{bmatrix} \quad \text{and} \quad V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

- 6 If $A = U\Sigma V^T$ then $A^T = V\Sigma^T U^T$ and $A^T A = V\Sigma^T \Sigma V^T$. This is a diagonalization $V\Lambda V^T$ with $\Lambda = \Sigma^T \Sigma$ (so each $\sigma_i^2 = \lambda_i$). Similarly $AA^T = U\Sigma \Sigma^T U^T$ is a diagonalization of AA^T . We see that the eigenvalues in $\Sigma \Sigma^T$ are the same $\sigma_i^2 = \lambda_i$.

- 7 This small question is a key to everything. It is based on the associative law $(AA^T)A = A(A^T A)$. Here we are applying both sides to an eigenvector \mathbf{v} of $A^T A$:

$$(AA^T)A\mathbf{v} = A(A^T A)\mathbf{v} = A\lambda\mathbf{v} = \lambda A\mathbf{v}.$$

So $A\mathbf{v}$ is an eigenvector of AA^T with the same eigenvalue λ .

- 8 $A = U\Sigma V^T = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}^T = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & 3 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} \sqrt{50} & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$

- 9 This $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ is a 2 by 2 matrix of rank 1. Its row space has basis \mathbf{v}_1 , its nullspace has basis \mathbf{v}_2 , its column space has basis \mathbf{u}_1 , its left nullspace has basis \mathbf{u}_2 :

$$\begin{aligned} \text{Row space} & \quad \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix} & \quad \text{Nullspace} & \quad \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ -1 \end{bmatrix} \\ \text{Column space} & \quad \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ 3 \end{bmatrix}, & \quad \mathbf{N}(A^T) & \quad \frac{1}{\sqrt{10}} \begin{bmatrix} 3 \\ -1 \end{bmatrix}. \end{aligned}$$

- 10** If A has rank 1 then so does $A^T A$. The only nonzero eigenvalue of $A^T A$ is its trace, which is the sum of all a_{ij}^2 . (Each diagonal entry of $A^T A$ is the sum of a_{ij}^2 down one column, so the trace is the sum down all columns.) Then $\sigma_1 =$ square root of this sum, and $\sigma_1^2 =$ this sum of all a_{ij}^2 .
- 11** $A^T A = AA^T = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ has eigenvalues $\sigma_1^2 = \frac{3 + \sqrt{5}}{2}$, $\sigma_2^2 = \frac{3 - \sqrt{5}}{2}$. But A is indefinite
 $\sigma_1 = (1 + \sqrt{5})/2 = \lambda_1(A)$, $\sigma_2 = (\sqrt{5} - 1)/2 = -\lambda_2(A)$; $\mathbf{u}_1 = \mathbf{v}_1$ but $\mathbf{u}_2 = -\mathbf{v}_2$.
- 12** A proof that *eigshow* finds the SVD. When $\mathbf{V}_1 = (1, 0)$, $\mathbf{V}_2 = (0, 1)$ the demo finds $A\mathbf{V}_1$ and $A\mathbf{V}_2$ at some angle θ . A 90° turn by the mouse to \mathbf{V}_2 , $-\mathbf{V}_1$ finds $A\mathbf{V}_2$ and $-A\mathbf{V}_1$ at the angle $\pi - \theta$. Somewhere between, the constantly orthogonal \mathbf{v}_1 and \mathbf{v}_2 must produce $A\mathbf{v}_1$ and $A\mathbf{v}_2$ at angle $\pi/2$. Those orthogonal directions give \mathbf{u}_1 and \mathbf{u}_2 .
- 13** The number $\sigma_{\max}(A^{-1})\sigma_{\max}(A)$ is the same as $\sigma_{\max}(A)/\sigma_{\min}(A)$. This is certainly ≥ 1 . It equals 1 if all σ 's are equal, and $A = U\Sigma V^T$ is a multiple of an orthogonal matrix. The ratio $\sigma_{\max}/\sigma_{\min}$ is the important **condition number** of A studied in Section 9.2.
- 14** $A = UV^T$ since all $\sigma_j = 1$, which means that $\Sigma = I$.
- 15** A rank-1 matrix with $A\mathbf{v} = 12\mathbf{u}$ would have \mathbf{u} in its column space, so $A = \mathbf{u}\mathbf{w}^T$ for some vector \mathbf{w} . I intended (but didn't say) that \mathbf{w} is a multiple of the unit vector $\mathbf{v} = \frac{1}{2}(1, 1, 1, 1)$ in the problem. Then $A = 12\mathbf{u}\mathbf{v}^T$ to get $A\mathbf{v} = 12\mathbf{u}$ when $\mathbf{v}^T\mathbf{v} = 1$.
- 16** If A has orthogonal columns $\mathbf{w}_1, \dots, \mathbf{w}_n$ of lengths $\sigma_1, \dots, \sigma_n$, then $A^T A$ will be diagonal with entries $\sigma_1^2, \dots, \sigma_n^2$. So the σ 's are definitely the singular values of A (as expected). The eigenvalues of that diagonal matrix $A^T A$ are the columns of I , so $V = I$ in the SVD. Then the \mathbf{u}_i are $A\mathbf{v}_i/\sigma_i$ which is the unit vector \mathbf{w}_i/σ_i .
- The SVD of this A with orthogonal columns is $A = U\Sigma V^T = (A\Sigma^{-1})(\Sigma)(I)$.
- 17** Since $A^T = A$ we have $\sigma_1^2 = \lambda_1^2$ and $\sigma_2^2 = \lambda_2^2$. But λ_2 is negative, so $\sigma_1 = 3$ and $\sigma_2 = 2$. The unit eigenvectors of A are the same $\mathbf{u}_1 = \mathbf{v}_1$ as for $A^T A = AA^T$ and $\mathbf{u}_2 = -\mathbf{v}_2$ (notice the sign change because $\sigma_2 = -\lambda_2$, as in Problem 11).
- 18** Suppose the SVD of R is $R = U\Sigma V^T$. Then multiply by Q to get $A = QR$. So the SVD of this A is $(QU)\Sigma V^T$. (Orthogonal Q times orthogonal $U =$ orthogonal QU .)

19 The smallest change in A is to set its smallest singular value σ_2 to zero.

20 $A^T A = \begin{bmatrix} 1 & 100 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 100 & 1 \end{bmatrix} = \begin{bmatrix} 10001 & 100 \\ 100 & 1 \end{bmatrix}$ has eigenvalues $\lambda(A^T A) = \sigma^2(A)$.

$$\lambda^2 - 10002\lambda + 1 = 0 \text{ gives } \lambda = 5001 \pm \sqrt{(5001)^2 - 1} \approx 5001 \pm \left(5001 - \frac{1}{10002}\right).$$

So $\lambda \approx 10002$ and $1/10002$ and $\sigma \approx 100.01$ and $1/100.01$. Check $\sigma_1\sigma_2 \approx 1 = \det A$.

21 The singular values of $A + I$ are not $\sigma_j + 1$. They come from eigenvalues of $(A + I)^T(A + I)$. Test the diagonal matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$.

22 Since Q_1 and U are orthogonal, so is Q_1U . (check: $(Q_1U)^T(Q_1U) = U^TQ_1^TQ_1U = U^TU = I$.) So the SVD of the matrix $Q_1AQ_2^T$ is just $Q_1U\Sigma V^TQ_2^T = (Q_1U)\Sigma(Q_2V)^T$ and Σ is the same as for A . The matrices A and $Q_1AQ_2^T$ and Σ are all “isometric” = sharing the same Σ .

23 The singular values of Q are the eigenvalues of $Q^TQ = I$ (therefore all 1's).

24 (a) From $\mathbf{x}^T S \mathbf{x} = 3x_1^2 + 2x_1x_2 + 3x_2^2$ you can see that $S = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. Its eigenvalues are 4 and 2. The maximum of $\mathbf{x}^T S \mathbf{x} / \mathbf{x}^T \mathbf{x}$ is 4.

(b) The 1 by 2 matrix $A = \begin{bmatrix} 1 & 4 \end{bmatrix}$ leads to $\frac{\|A\mathbf{x}\|^2}{\|\mathbf{x}\|^2} = \frac{(x_1 + 4x_2)^2}{x_1^2 + x_2^2}$. The maximum value is $\sigma_1^2(A)$. For this matrix $A = \begin{bmatrix} 1 & 4 \end{bmatrix}$ that singular value squared is $\sigma_1^2 = 17$.

This is because $AA^T = \begin{bmatrix} 17 \end{bmatrix}$ and also $A^T A = \begin{bmatrix} 1 & 4 \\ 4 & 16 \end{bmatrix}$ has $\lambda = 17$ and 0.

25 The minimum value of $\frac{\mathbf{x}^T S \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$ is the *smallest* eigenvalue of S . The eigenvector is the minimizing \mathbf{x} . That eigenvector gives $\mathbf{x}^T S \mathbf{x} = \mathbf{x}^T \lambda_{\min} \mathbf{x}$.

Since $\frac{\|A\mathbf{x}\|^2}{\|\mathbf{x}\|^2} = \frac{\mathbf{x}^T A^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$ we see that the minimizing \mathbf{x} is an **eigenvector of $A^T A$** (and not usually an eigenvector of A).

26 From $AV = U\Sigma$ we know that $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ = first column of V goes to $2 \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}$ = first column of $U\Sigma$. Similarly the second column $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ goes to $\begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix}$. The two outputs are orthogonal and they are the axes of an ellipse. With $\theta = 30^\circ$ those axes are $\begin{bmatrix} \sqrt{3} \\ 1 \end{bmatrix}$ going out from $(0, 0)$ at 30° and $\frac{1}{2} \begin{bmatrix} -1 \\ \sqrt{3} \end{bmatrix}$ going out at 120° . Comparing to the picture in Section 7.4, the first step would be a reflection (not a rotation), then a stretch by factors 2 and 1, then a 30° rotation.

27 Start from $A = U\Sigma V^T$. The columns of U are a basis for the column space of A , and so are the columns of C , so $U = CF$ for some invertible r by r matrix F .

Similarly the columns of V are a basis for the row space of A and so are the columns of B , so $V = BG$ for some invertible r by r matrix G .

Then $A = U\Sigma V^T = C(F\Sigma G^T)B^T = CMB^T$ and $M = F\Sigma G^T$ is r by r and invertible.

Problem Set 7.3, page 391

1 The row averages of A_0 are 3 and 0. Therefore

$$A = \begin{bmatrix} 2 & 1 & 0 & -1 & -2 \\ -1 & 1 & 0 & 1 & -1 \end{bmatrix} \quad \text{and} \quad S = \frac{AA^T}{4} = \frac{1}{4} \begin{bmatrix} 10 & 0 \\ 0 & 4 \end{bmatrix}$$

The eigenvalues of S are $\lambda_1 = \frac{10}{4}$ and $\lambda_2 = \frac{4}{4} = 1$. The top eigenvector of S is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

I think this means that a **vertical line** is closer to the five points $(2, -1), \dots, (-2, -1)$ in the columns of A than any other line through the origin $(0, 0)$.

2 Now the row averages of A_0 are $\frac{1}{2}$ and 2. Therefore

$$A = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -1 & 0 & 1 & 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad S = \frac{AA^T}{5} = \frac{1}{5} \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & 4 \end{bmatrix}.$$

Again the rows of A are accidentally orthogonal (because of the special patterns of those rows). This time the top eigenvector of S is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. So a **horizontal line** is closer to the six points $(\frac{1}{2}, -1), \dots, (-\frac{1}{2}, -1)$ from the columns of A than any other line through the center point $(0, 0)$.

3 $A_0 = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 2 & 2 \end{bmatrix}$ has row averages 2 and so $A = \begin{bmatrix} -1 & 0 & 1 \\ 2 & -1 & -1 \end{bmatrix}$. Then $S = \frac{1}{2}AA^T = \frac{1}{2} \begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix}$.

Then $\text{trace}(S) = \frac{1}{2}(8)$ and $\det(S) = (\frac{1}{2})^2(3)$. The eigenvalues $\lambda(S)$ are $\frac{1}{2}$ times the roots of $\lambda^2 - 8\lambda + 3 = 0$. Those roots are $4 \pm \sqrt{16-3}$. Then the σ 's are $\sqrt{\lambda_1}$ and $\sqrt{\lambda_2}$.

4 This matrix A with orthogonal rows has $S = \frac{AA^T}{n-1} = \frac{1}{3} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 4 \end{bmatrix}$.

With λ 's in descending order $\lambda_1 > \lambda_2 > \lambda_3$, the eigenvectors are $(0, 1, 0)$ and $(0, 0, 1)$ and $(1, 0, 0)$. The first eigenvector shows the \mathbf{u}_1 direction. Combined with the second eigenvector \mathbf{u}_2 , the best plane is the yz plane.

These problems are examples where the sample **correlation matrix** (rescaling S so all its diagonal entries are 1) would be the identity matrix. If we think the original scaling is not meaningful and the rows should have the same length, then there is no reason to choose $\mathbf{u}_1 = (0, 1, 0)$ from the 8 in row 2.

5 The correlation matrix DSD which has 1's on the diagonal is

$$DSD = \begin{bmatrix} \frac{1}{2} & & \\ & \frac{1}{2} & \\ & & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 & 0 \\ 2 & 4 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & & \\ & \frac{1}{2} & \\ & & 1 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 1 \end{bmatrix}.$$

6 Working with letters instead of numbers, the correlation matrix $C = DSD$ is

$$\begin{bmatrix} 1 & c_{12} & c_{13} \\ c_{12} & 1 & c_{23} \\ c_{13} & c_{23} & 1 \end{bmatrix} \text{ with } c_{12} = \frac{S_{12}}{\sigma_1\sigma_2} \text{ and } c_{13} = \frac{S_{13}}{\sigma_1\sigma_3} \text{ and } c_{23} = \frac{S_{23}}{\sigma_2\sigma_3}.$$

$$\text{Then } D = \begin{bmatrix} 1/\sigma_1 & & \\ & 1/\sigma_2 & \\ & & 1/\sigma_3 \end{bmatrix} \text{ gives } DSD = C.$$

7 From each row of A_0 , subtract the average of that row (the average grade for that course) from the 10 grades in that row. This produces the centered matrix A . Then the sample covariance matrix is $S = \frac{1}{9}AA^T$. The leading eigenvector of the 5 by 5 matrix S tells the weights on the 5 courses to produce the “*eigencourse*”. This is the course whose grades have the most information (the greatest variance).

If a course gives everyone an A , the variance is zero and that course is not included in the eigencourse. We are looking for most information not best grade.

Problem Set 7.4, page 398

1 $A^T A = \begin{bmatrix} 10 & 20 \\ 20 & 40 \end{bmatrix}$ has $\lambda = 50$ and 0 , $\mathbf{v}_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ -1 \end{bmatrix}$; $\sigma_1 = \sqrt{50}$.

2 Orthonormal bases: \mathbf{v}_1 for row space, \mathbf{v}_2 for nullspace, \mathbf{u}_1 for column space, \mathbf{u}_2 for $N(A^T)$. All matrices with those four subspaces are multiples cA , since the subspaces are just lines. Normally many more matrices share the same 4 subspaces. (For example, all n by n invertible matrices share \mathbf{R}^n as their column space.)

3 $A = QS = \frac{1}{\sqrt{50}} \begin{bmatrix} 7 & -1 \\ 1 & 7 \end{bmatrix} \frac{1}{\sqrt{50}} \begin{bmatrix} 10 & 20 \\ 20 & 40 \end{bmatrix}$. S is semidefinite because A is singular.

4 $A^+ = V \begin{bmatrix} 1/\sqrt{50} & 0 \\ 0 & 0 \end{bmatrix} U^T = \frac{1}{50} \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$; $A^+A = \begin{bmatrix} .2 & .4 \\ .4 & .8 \end{bmatrix}$, $AA^+ = \begin{bmatrix} .1 & .3 \\ .3 & .9 \end{bmatrix}$.

5 $A^T A = \begin{bmatrix} 10 & 8 \\ 8 & 10 \end{bmatrix}$ has $\lambda = 18$ and 2 , $\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\sigma_1 = \sqrt{18}$ and $\sigma_2 = \sqrt{2}$.

6 $AA^T = \begin{bmatrix} 18 & 0 \\ 0 & 2 \end{bmatrix}$ has $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. The same $\sqrt{18}$ and $\sqrt{2}$ go into Σ .

7 $\begin{bmatrix} \sigma_1 \mathbf{u}_1 & \sigma_2 \mathbf{u}_2 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \end{bmatrix} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T$. In general this is $\sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T$.

8 $A = U\Sigma V^T$ splits into QK (polar): $Q = UV^T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ and $K = V\Sigma V^T = \begin{bmatrix} \sqrt{18} & 0 \\ 0 & \sqrt{2} \end{bmatrix}$.

9 A^+ is A^{-1} because A is invertible. Pseudoinverse equals inverse when A^{-1} exists!

10 $A^T A = \begin{bmatrix} 9 & 12 & 0 \\ 12 & 16 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ has $\lambda = 25, 0, 0$ and $\mathbf{v}_1 = \begin{bmatrix} .6 \\ .8 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} .8 \\ -.6 \\ 0 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Here $A = [3 \ 4 \ 0]$ has rank 1 and $AA^T = [25]$ and $\sigma_1 = 5$ is the only singular value in $\Sigma = [5 \ 0 \ 0]$.

11 $A = [1] [5 \ 0 \ 0] V^T$ and $A^+ = V \begin{bmatrix} .2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} .12 \\ .16 \\ 0 \end{bmatrix}$; $A^+ A = \begin{bmatrix} .36 & .48 & 0 \\ .48 & .64 & 0 \\ 0 & 0 & 0 \end{bmatrix}$; $AA^+ = [1]$

12 The zero matrix has no pivots or singular values. Then $\Sigma =$ same 2 by 3 zero matrix and the pseudoinverse is the 3 by 2 zero matrix.

13 If $\det A = 0$ then $\text{rank}(A) < n$; thus $\text{rank}(A^+) < n$ and $\det A^+ = 0$.

14 This problem explains why the matrix A transforms the circle of unit vectors $\|\mathbf{x}\| = 1$ into an **ellipse** of vectors $\mathbf{y} = A\mathbf{x}$. The reason is that $\mathbf{x} = A^{-1}\mathbf{y}$ and the vectors with $\|A^{-1}\mathbf{y}\| = 1$ do lie on an ellipse:

$$\|A^{-1}\mathbf{y}\|^2 = 1 \quad \text{is} \quad \mathbf{y}^T (A^{-1})^T A^{-1} \mathbf{y} = 1 \quad \text{or} \quad \mathbf{y}^T (AA^T)^{-1} \mathbf{y} = 1.$$

That matrix $(AA^T)^{-1}$ is *symmetric positive definite* (A is assumed invertible).

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{gives} \quad AA^T = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \quad \text{and} \quad (AA^T)^{-1} = \frac{1}{9} \begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix}.$$

So the ellipse $\|A^{-1}\mathbf{y}\|^2 = 1$ of outputs $\mathbf{y} = A\mathbf{x}$ has equation $5y_1^2 - 8y_1y_2 + 5y_2^2 = 9$.

The singular values of this positive definite A are its eigenvalues 3 and 1.

The ellipse $\|A^{-1}\mathbf{y}\| = 1$ has semi-axes of lengths $1/3$ and $1/1$.

- 15** (a) $A^T A$ is singular (b) This \mathbf{x}^+ in the row space does give $A^T A\mathbf{x}^+ = A^T \mathbf{b}$ (c) If $(1, -1)$ in the nullspace of A is added to \mathbf{x}^+ , we get another solution to $A^T A\hat{\mathbf{x}} = A^T \mathbf{b}$. But this $\hat{\mathbf{x}}$ is longer than \mathbf{x}^+ because the added part is orthogonal to \mathbf{x}^+ in the row space and $\|\hat{\mathbf{x}}\|^2 = \|\mathbf{x}^+\|^2 + \|\text{added part from nullspace}\|^2$.

- 16** \mathbf{x}^+ in the row space of A is perpendicular to $\hat{\mathbf{x}} - \mathbf{x}^+$ in the nullspace of $A^T A =$ nullspace of A . The right triangle has $c^2 = a^2 + b^2$.

- 17** $AA^+ \mathbf{p} = \mathbf{p}$, $AA^+ \mathbf{e} = \mathbf{0}$, $A^+ A\mathbf{x}_r = \mathbf{x}_r$, $A^+ A\mathbf{x}_n = \mathbf{0}$.

- 18** $A^+ = V\Sigma^+U^T$ is $\frac{1}{5} [.6 \ .8] = [.12 \ .16]$ and $A^+ A = [1]$ and $AA^+ = \begin{bmatrix} .36 & .48 \\ .48 & .64 \end{bmatrix} =$ projection.

- 19** L is determined by ℓ_{21} . Each eigenvector in X is determined by one number. The counts are $1 + 3$ for LU , $1 + 2 + 1$ for LDU , $1 + 3$ for QR (notice 1 rotation angle), $1 + 2 + 1$ for $U\Sigma V^T$, $2 + 2 + 0$ for $X\Lambda X^{-1}$.

- 20** LDL^T and $Q\Lambda Q^T$ are determined by $1 + 2 + 0$ numbers because A is symmetric.

Note Problem 20 should have referred to Problem 19 not 18.

- 21** Check the formula for $A^+ A$ using A^+ and A :

$$A^+ A = \left(\sum_1^r \frac{\mathbf{v}_i \mathbf{u}_i^T}{\sigma_i} \right) \left(\sum_1^r \sigma_j \mathbf{u}_j \mathbf{v}_j^T \right) = \sum_1^r \mathbf{v}_i \mathbf{u}_i^T \mathbf{u}_i \mathbf{v}_i^T \text{ because } \mathbf{u}_i^T \mathbf{u}_j = 0 \text{ when } i \neq j$$

Then every $\mathbf{u}_i^T \mathbf{u}_i = 1$ (unit vector) so $A^+ A = \sum_1^r \mathbf{v}_i \mathbf{v}_i^T$ is correct.

$$\text{Similarly } AA^+ = \left(\sum_1^r \sigma_j \mathbf{u}_j \mathbf{v}_j^T \right) \left(\sum_1^r \frac{\mathbf{v}_i \mathbf{u}_i^T}{\sigma_i} \right) = \sum_1^r \mathbf{u}_i \mathbf{u}_i^T.$$

22 $M = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} A\mathbf{v} \\ A^T\mathbf{u} \end{bmatrix} = \sigma \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}$. Thus $\begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}$ is an eigenvector.

The singular values of A are *eigenvalues* of this block matrix.