

Notes

Ryan Heslin

August 26, 2021

1.

1 and 2.

2.

The singular values of an orthogonal matrix are I_n because $Q^T Q = I$, which has those singular values.

3.

See above.

4.

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$
$$A^T A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

They come out ugly: $\sqrt{\frac{3 \pm \sqrt{5}}{2}}$

5.

$$A = \begin{bmatrix} p & -q \\ q & p \end{bmatrix}$$

The singular values are $\sqrt{p^2 + q^2}$. In the special case of a rotation, those are both $\sqrt{\sin^2 \theta + \cos^2 \theta} = 1$, since rotations are orthogonal; otherwise the scaling factor applied in the rotation-scaling.

6.

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

$\sigma_1 = 5$ since we know one eigenvalue is 0. The lone singular vector is \$ \$.

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 5/\sqrt{5} \\ 10/\sqrt{5} \end{bmatrix} \|Av_1\| = \sqrt{25/5 + 100/5} = \sqrt{25} = 5$$

7.

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Neat trick with the sign reversing

$$\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} / 2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

8.

How rotations work - scale basis if need be, then transform columns.

$$A = \begin{bmatrix} p & -q \\ q & p \end{bmatrix}$$

$$A = \frac{1}{\sqrt{p^2 + q^2}} \begin{bmatrix} p & -q \\ q & p \end{bmatrix} \begin{bmatrix} \sqrt{p^2 + q^2} & 0 \\ 0 & \sqrt{p^2 + q^2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

9.

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

$$A = \begin{bmatrix} 1/\sqrt{5} & 0 \\ 2/\sqrt{5} & 0 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 2/\sqrt{5} \\ 0 & 0 \end{bmatrix}$$

11.

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}$$

$$V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$AV = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 1 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 1/\sqrt{6} & -1/\sqrt{2} \\ 2/\sqrt{6} & 0 \\ 1/\sqrt{6} & 1/\sqrt{2} \end{bmatrix}$$

$$A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1/\sqrt{6} & -1/\sqrt{2} \\ 2/\sqrt{6} & 0 \\ 1/\sqrt{6} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

```

A <- square(6, 2, -7, 6)
list2env(setNames(eigen(t(A) %*% A), c("Sigma", "V")), envir = globalenv())

<environment: R_GlobalEnv>

Sigma <- sqrt(Sigma[Sigma > 0])
U <- A %*% V %*% diag(x = 1/Sigma)

eqn <- paste("A =", paste(sapply(list(U, diag(x = Sigma), t(V)), mat2latex,
  sink = TRUE), collapse = " "))
print_eqn(eqn)

```

$$A = \begin{bmatrix} -0.894427190999916 & -0.447213595499958 \\ 0.447213595499958 & -0.894427190999916 \end{bmatrix} \begin{bmatrix} 10 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} -0.447213595499958 & 0.894427190999916 \\ -0.894427190999916 & -0.447213595499958 \end{bmatrix}$$

11.

12.

13.

14.

15.

The singular values of A are the inverses of those of A^{-1}

16.

See above.

17.

Watch as I blunder into proving the generalized inverse is the least-squares solution:

$$\begin{aligned}
A^T A x &= A^T b \\
x &= (A^T A)^{-1} A^T b \\
&= (V \Sigma^T U^T U \Sigma V^T)^{-1} V \Sigma^T U^T b \\
&= (\Sigma V^T)^{-1} (V \Sigma^T)^{-1} V \Sigma^T U^T b \\
&= V \Sigma^+ (\Sigma^T)^+ V^T V \Sigma \\
&= V \Sigma^+ (\Sigma^T)^+ \Sigma^T U^T b \\
&= V \Sigma^+ U^T b
\end{aligned}$$

Of course, the diagonal of Σ^+ is $1/\sigma_i$.

18.

```
A_plus <- t(square(3, 4, -4, 3)) %*% matrix(c(1/2, 0, 0, 1, 0, 0, 0, 0), nrow = 2) %*%
  (1/10 * t(square(rep(1, 6), -1, -1, 1, -1, 1, -1, 1, -1, -1, 1)))

b <- 1:4
x <- c(A_plus %*% b, 0, 0)
```

19.

As shown above, the solution x is a linear combination of V , where each coefficient is the dot product of b with u_i divided by σ_i (the length of Au_i). This is roughly a sum of normalized projections.

20.

The polar decomposition is just:

$$A = (UV^T)(V\Sigma V^T)$$

where the first term is clearly orthogonal and the second symmetric.

In the reverse form

$$A = (V\Sigma V^T)^T(UV^T)^T$$

21.

Polar decomp:

```
A <- square(6, -7, 2, 6)
list2env(setNames(eigen(t(A) %*% A), c("Sigma", "V")), envir = global_env())
```

```
<environment: R_GlobalEnv>
```

```
Sigma <- diag(x = sqrt(Sigma[Sigma > 0]))
U <- sweep(A %*% V, diag(Sigma), `/,`, MARGIN = 2)
Q <- U %*% t(V)
S <- V %*% Sigma %*% t(V)
Q %*% S
```

```
      [,1] [,2]
[1,]     6     2
[2,]    -7     6
```

22.

a. The cross-products matrix

$$\begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}$$

can be decomposed by SVD into three linear transformations $A_3 A_2 A_1$. Then the polar decomposition is:

$$A = (A_1 A_3)(A_3^T A_2 A_3$$

This matrix first projects onto the plane shared by the two vectors. If $v = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$, then the SVD is:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

In polar form:

$$A = \left(\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \right) \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \right)$$

Cribbing from an earlier exercise involving the cross product, the transformation projects onto the plane of the second two coordinates of the vector x being crossed with v , scales by the singular values of v , then rotates $\pi/2$ counterclockwise about c_1

23.

$$\begin{aligned} AA^T &= U \Sigma V^T V \Sigma^T U^T \\ &= U \Sigma \Sigma^T U^T \end{aligned}$$

This is an eigenbasis because $\Lambda = \Sigma \Sigma^T$. This shows that the eigenvalues (up to r) of $A^T A$ and AA^T are the same.

24.

They are the absolute values of the eigenvalues, the square roots of the eigenvalues of $A^T A$, which are square of those of A .

25.

Let u be a vector of V . Then $Au = \sigma u_i \implies \|Au\| = \sigma_i$. The σ s account for the scaling factor applied to every vector in A 's basis, so if u is not a member of V $\|Au\|$ must fall in that range.

26.

The logic for a general $n \times m$ is the same - the singular values are the upper and lower bound on the lengths of the transformed unit vector.

27.

The singular values are the square roots of the eigenvalues of $A^T A$. The eigenvalues of $A^T A$ are the squares of those of A . Since the singular values are just the absolute values of the eigenvalues in this case, the absolute values of the eigenvalues must fall in their range. ## 28.

$$\begin{aligned}\det(A^T A) &= \det(A^T) \det(A) \\ &= \det A^2 \\ \prod_{i=1}^n \sigma_i &= \sqrt{\det A^2} = \det A\end{aligned}$$

29.

This is obvious from the definitions: σ_i scales the product of u_i and v_i^T that forms a column of A .

30.

The eigenvalues of A are $6 \pm \sqrt{14}i$, with absolute values $5\sqrt{10}$

$$\begin{aligned}A &= \begin{bmatrix} 6 & 2 \\ -7 & 6 \end{bmatrix} \\ \Sigma &= \begin{bmatrix} 10 & 0 \\ 0 & 5 \end{bmatrix} \\ V &= \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} \\ U &= \frac{1}{\sqrt{5}} \begin{bmatrix} -1 & -2 \\ -2 & 1 \end{bmatrix}\end{aligned}$$

Which implies

$$A = 2 \begin{bmatrix} -1 \\ -2 \end{bmatrix} \begin{bmatrix} 2 & -1 \end{bmatrix} + \begin{bmatrix} -2 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix}$$

which is actually the formula for $-A$ because I screwed up the signs of the eigenvalues. ## 31.

Any matrix may be decomposed:

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

which may be written as a linear combination:

$$A = \sigma_1 u_1 v^T + \cdots + \sigma_r u_r v^T$$

where each term is a rank-one matrix, as an outer product of two vectors scaled by a nonzero σ_i . Any σ beyond σ_r yield rank-0 matrices that can be ignored.

32.

33.

Yes, that implies the eigenvalues are all ± 1 , which is true only of orthogonal matrices. ## 34.

$U = V$ in the singular combination if $A^T A$ and $A A^T$ have exactly the same eigenvectors. Since they have the same eigenvalues anyway, this happens only if $A^T = A$ - that is, A is symmetric.

35.

Assuming full column rank, this just returns e_i :

$$\begin{aligned} & (A^T A)^{-1} A^T u_i \\ & V \Sigma^+ U^T u_i \\ & V \Sigma^+ e_i \qquad \frac{1}{\sigma_i} v_i \end{aligned}$$

36.

This projects u_i into A 's image. Since U up to u_m is a basis for the image already, this resolves to u_i again if $i \leq m$, 0 otherwise (since $u_{i>m}$ are in $\ker A^T$).

$$\begin{aligned} & A(A^T A)^{-1} A^T u_i \\ & = U \Sigma V^T V \Sigma^+ e_i \\ & = U e_i \\ & = u_i \end{aligned}$$