Singular value decomposition

(Singular value decomposition)

Every $m \times n$ matrix A can be decomposed as $A = U\Sigma V^T$, where

- Σ is a (rectangular) diagonal matrix with nonnegative entries, $(m \times n)$

The diagonal entries σ_i are called the **singular values** of A.

- U is orthogonal, $(m \times m)$
- V is orthogonal. $(n \times n)$

Comment. If A is symmetric, then the singular value decomposition is already provided by the spectral theorem (the diagonalization of A). Moreover, in that case, $V = U$.

Important observations. If $A = U\Sigma V^T$, then $A^T A = V\Sigma^T \Sigma V^T$.

- Note that $\Sigma^T \Sigma$ is an $n \times n$ diagonal matrix. Its entries are σ_i^2 (the squares of the entries in Σ).
- $A^T A$ is a symmetric matrix! (Why?!) Hence, by the spectral theorem, we are able to find V and $\Sigma^T \Sigma$.

In other words, V is obtained from the (orthonormally chosen) eigenvectors of $A^T A$. Likewise, the entries of $\Sigma^T \Sigma$ are the eigenvalues of $A^T A$; their square roots are the entries of Σ, the singular values.

Finally, the equation $AV = U\Sigma$ allows us to determine U. How?! (Hint: $A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$)

This results in the following **recipe** to determine the SVD $A = U\Sigma V^T$ for any matrix A.

Find an orthonormal basis of eigenvectors $\boldsymbol{v}_i$ of $A^T A$. Let λ_i be the eigenvalue of $\boldsymbol{v}_i$.

- V is the matrix with columns $\boldsymbol{v}_i$.
- Σ is the diagonal matrix with entries $\sigma_i = \sqrt{\lambda_i}$.
- U is the matrix with columns $\boldsymbol{u}_i = \frac{1}{\sigma_i} A\boldsymbol{v}_i$. If needed, fill in additional columns to make U orthogonal.

Example 163. Determine the SVD of $A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$.

Solution. $A^T A = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$ has 8-eigenvector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and 2-eigenvector $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Since $A^T A = V\Sigma^2 V^T$ (here, $\Sigma^T \Sigma = \Sigma^2$), we conclude that $V = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \sqrt{8} & \\ & \sqrt{2} \end{bmatrix}$.

From $A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$, we find $\boldsymbol{u}_1 = \frac{1}{\sigma_1} A\boldsymbol{v}_1 = \frac{1}{\sqrt{8}}\begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Likewise, $\boldsymbol{u}_2 = \frac{1}{\sigma_2} A\boldsymbol{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Hence, $U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Check that, indeed, $A = U\Sigma V^T$!

Comment. For applications, it is common to arrange the singular values in decreasing order like we did.

Comment. If we had chosen $V = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}$ instead, then $U = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \sqrt{8} & \\ & \sqrt{2} \end{bmatrix}$.

As with diagonalization, there are choices! (A lot fewer choices though.) This is another perfectly fine SVD. In fact, it's what Sage computes below.

Sage. Let's have Sage do the work for us. In Sage, the SVD is currently only implemented for floating point numbers. (RDF is the real numbers as floating point numbers with double precision)

```
Sage] A = matrix(RDF, [[2,2],[-1,1]])
Sage] U,S,V = A.SVD()
Sage] U
```

$$\begin{bmatrix} -1.0 & 1.11022302463\times 10^{-16} \\ 8.64109131471\times 10^{-17} & 1.0 \end{bmatrix}$$

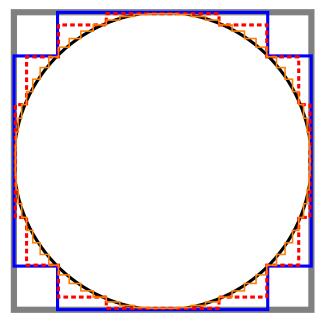
```
Sage] S
```

$$\begin{bmatrix} 2.82842712475 & 0.0 \\ 0.0 & 1.41421356237 \end{bmatrix}$$

```
Sage] V
```

$$\begin{bmatrix} -0.707106781187 & -0.707106781187 \\ -0.707106781187 & 0.707106781187 \end{bmatrix}$$

Remark 164. (April Fools' Day!) π is the perimeter of a circle enclosed in a square with edge length 1. The perimeter of the square is 4, which approximates π. To get a better approximation, we "fold" the vertices of the square towards the circle (and get the blue polygon). This construction can be repeated for even better approximations and, in the limit, our shape will converge to the true circle. At each step, the perimeter is 4, so we conclude that $\pi = 4$, contrary to popular belief.



Can you pin-point the fallacy in this argument?

Comment. We'll actually come back to this. It's related to linear algebra in infinite dimensions.