1. $[\mathbf{Q}, \mathbf{R}] = \texttt{qr}(\mathbf{A})$
2. compute $\mathbf{Q}^\top \mathbf{b}$ and set $\mathbf{b}^* = (\mathbf{Q}^\top \mathbf{b})([1:n])$
3. solve $\mathbf{R}^* \mathbf{x} = \mathbf{b}^*$ with back substitution

slide 28

Remark 5.12 *The QR-factorization can also be used in the case $m = n$ to solve a linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ with the following three steps:*

1. *compute the QR-factorization of $\mathbf{A}$*
2. *solve $\mathbf{Q}\mathbf{y} = \mathbf{b}$ by computing $\mathbf{y} = \mathbf{Q}^\top \mathbf{b}$*
3. *solve $\mathbf{R}\mathbf{x} = \mathbf{y}$ by back substitution*

The cost is about twice that of the procedure using an LU-factorization. It is, however, preferred if $\kappa(\mathbf{A})$ is large.

slide 29

■

Example 5.13 *Consider*

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ \varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 2 \\ \varepsilon \\ \varepsilon \end{pmatrix}, \qquad \mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \mathbf{A}^\top \mathbf{A} = \begin{pmatrix} 1+\varepsilon^2 & 1 \\ 1 & 1+\varepsilon^2 \end{pmatrix}.$$

Note $\mathbf{A}\mathbf{x} = \mathbf{b}$ so that $\mathbf{x}$ is the exact solution of the least squares problem. We note $\kappa(\mathbf{A}^\top \mathbf{A}) = \frac{2}{\varepsilon^2} + 1$ so that $\mathbf{A}$ is ill-conditioned for small ε. In `matlab`*:*

```
>> e = 1e-7;
>> A = [1 1; e 0; 0 e]; b = [2;e;e];
>> x = (A'*A)\(A'*b) %solution using normal equations
x =
   1.011235955056180
   0.988764044943820
>> [Q,R] = qr(A) ;
>> bb=Q'*b ;
>> xx = R(1:2,1:2)\bb(1:2) %solution using QR-factorization
xx =
   1.000000000000000
   1.000000000000000
```

The method using the normal equations yields a solution with two digits of accuracy (consistent with $\kappa(\mathbf{A}^\top \mathbf{A}) \approx 10^{14}$) whereas the method based on the QR-factorization yields the correct solution. ■

$$\mathbf{A} = \begin{pmatrix} * & * \\ * & * \\ * & * \end{pmatrix} = \mathbf{U} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{pmatrix} \mathbf{V}^\top$$

Figure 5.1: structure of the SVD of an 3×2 matrix; $\mathbf{U}$, $\mathbf{V}$ are orthogonal

5.3 underdetermined systems

The system (5.1) is underdetermined if $m < n$. Let us assume that $\mathbf{A}$ has full rank, i.e., it has m linearly independent rows. Then (5.1) has a solution. However, the solution is not unique. One way to fix the solution is to seek the *minimum norm solution*, i.e., to find $\mathbf{x}^*$ such that

$$\|\mathbf{x}^*\|_2 = \min\{\|\mathbf{y}\|_2 \,|\, \mathbf{Ay} = \mathbf{b}\}.$$

A convenient tool to solve this minimization problem is the *singular value decomposition* (SVD) of $\mathbf{A}$.

5.3.1 SVD

The SVD is a very important tool in the analysis of matrices. Without proof, we state its existence:

Theorem 5.14 (SVD) *Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ (m, n arbitrary). Then there exist $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min\{m,n\}} \geq 0$ and orthogonal matrices $\mathbf{U} \in \mathbb{R}^{m\times m}$, $\mathbf{V} \in \mathbb{R}^{n\times n}$, and $\Sigma \in \mathbb{R}^{m\times n}$ with $\Sigma_{ij} = \delta_{ij}\sigma_i$, $\sigma_i \geq 0$, such that*

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^\top, \tag{5.4}$$

The values σ_i are called the *singular values*, the columns of $\mathbf{U}$ the left singular vectors and the columns of the $\mathbf{V}$ the right singular vectors.
The SVD of a matrix $\mathbf{A}$ reveals many important properties of $\mathbf{A}$:

Exercise 5.15 *Let the singular values σ_i be sorted in descending order. Then:*

1. *Let $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_{\min\{m,n\}} = 0$. Then r is the* rank *of $\mathbf{A}$. (If all singular values are positive, then the matrix $\mathbf{A}$ has full rank).*
2. *The columns $\mathbf{U}(:, [1:r])$ form an orthogonal basis of the range of $\mathbf{A}$.*
3. *The columms $\mathbf{V}(:, [r+1:n])$ form an orthogonal basis of the kernel of $\mathbf{A}$. The columns $\mathbf{V}(:, [1:r])$ form an orthogonal basis of $(\ker \mathbf{A})^\perp$, the orthogonal complement of the kernel of $\mathbf{A}$.* ■

Exercise 5.16 *Let $\mathbf{U}\Sigma\mathbf{V}^\top$ be the SVD of $\mathbf{A}$. Show: the eigenvalues of $\mathbf{A}^\top\mathbf{A}$ are the eigenvalues of the diagonal matrix $\Sigma^\top\Sigma$ and those of $\mathbf{AA}^\top$ the eigenvalues of the diagonal matrix $\Sigma\Sigma^\top$. What can you say about the eigenvectors of the matrices $\mathbf{A}^\top\mathbf{A}$ and $\mathbf{AA}^\top$?* ■

Remark 5.17 *In* `matlab`/`python`*, the SVD is available as* `svd`/`numpy.linalg.svd`. ■

For $r = \text{rank}(\mathbf{A})$, we introduce the matrices

$$\widetilde{\mathbf{U}} = \mathbf{U}(:, [1:r]), \qquad \widetilde{\mathbf{V}} = \mathbf{V}(:, [1:r]), \qquad \widetilde{\Sigma} = \Sigma([1:r], [1:r]), \qquad \mathbf{V}' := \mathbf{V}(:, [r+1, n]).$$

We note that $\mathbf{A} = \widetilde{\mathbf{U}}\widetilde{\Sigma}\widetilde{\mathbf{V}}^\top$ and $\widetilde{\Sigma}$ is invertible. This factorization of $\mathbf{A}$ is called the *reduced SVD*. We also note that the columns ov $\mathbf{V}'$ span the kernel of $\mathbf{A}$.

Remark 5.18 *A slightly different interpretation of the SVD is obtained by writing it as* $\mathbf{AV} = \mathbf{U}\Sigma$*. Writing* $\mathbf{V} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$*,* $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_m)$*, this means* $\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i$*,* $i = 1, \ldots, r$*, where* $r = \text{rank}(\mathbf{A})$*. That is, we have found pairwise orthogonal vectors* $\mathbf{v}_i$ *that are mapped under* $\mathbf{A}$ *to an orthogonal basis of the range of* $\mathbf{A}$*.* ∎

Exercise 5.19 *Show:* $\mathbf{V}(\mathbf{V}')^\top \mathbf{x}$ *is the orthogonal projection of* $\mathbf{x}$ *onto* $\operatorname{Ker} \mathbf{A}$*.* $\widetilde{\mathbf{V}}\widetilde{\mathbf{V}}^\top \mathbf{x}$ *is the orthogonal projection of of* $\mathbf{x}$ *onto* $(\operatorname{Ker} \mathbf{A})^\perp$*. Analogously,* $\widetilde{\mathbf{U}}\widetilde{\mathbf{U}}^\top \mathbf{x}$ *is the orthgonal projection of* $\mathbf{x}$ *onto* $\operatorname{Im} \mathbf{A}$*.* ∎

5.3.2 Finding the minimum norm solution using the SVD

Let $m \leq n$ and assume (for simplicity) that $\mathbf{A}$ has full rank, i.e., $r = \text{rank}(\mathbf{A}) = m$. Then $\widetilde{\mathbf{U}} = \mathbf{U}$ and the reduced SVD then takes the form $\mathbf{A} = \mathbf{U}\widetilde{\Sigma}\widetilde{\mathbf{V}}^\top$. We observe that

$$\widetilde{\mathbf{x}}^* := \widetilde{\mathbf{V}}\widetilde{\Sigma}^{-1}\mathbf{U}^\top \mathbf{b}$$

satisfies $\mathbf{A}\mathbf{x}^* = \mathbf{b}$ since

$$\mathbf{A}\widetilde{\mathbf{x}}^* = \mathbf{U}\widetilde{\Sigma}\widetilde{\mathbf{V}}^\top \widetilde{\mathbf{V}}\widetilde{\Sigma}^{-1}\mathbf{U}^\top \mathbf{b} = \mathbf{U}\widetilde{\Sigma}\widetilde{\Sigma}^{-1}\mathbf{U}^\top \mathbf{b} = \mathbf{U}\mathbf{U}^\top \mathbf{b} = \mathbf{b}$$

We note that every solution $\mathbf{x}$ of $\mathbf{A}\mathbf{x} = \mathbf{b}$ has the form $\mathbf{x} = \widetilde{\mathbf{x}}^* + \mathbf{V}'\mathbf{y}$ for a $\mathbf{y} \in \mathbb{R}^{n-r}$. We also note that $\widetilde{\mathbf{x}}^*$ is orthogonal to $\ker \mathbf{A}$ (which is spanned by $\mathbf{V}'$). That is: for every solution $\mathbf{x}$ of $\mathbf{A}\mathbf{x} = \mathbf{b}$ we have

$$\|\mathbf{x}\|^2 = \|\widetilde{\mathbf{x}}^*\|^2 + \|\mathbf{V}'\mathbf{y}\|^2,$$

which is obviously minimized by $\mathbf{y} = 0$. Hence, $\widetilde{\mathbf{x}}^*$ is the sought minimum norm solution.

5.3.3 Solution of the least squares problem with the SVD

The least squares problem could, alternatively to using the QR-factorization, also be solved with the SVD:

Exercise 5.20 *Assume that* $m \geq n$ *and that an SVD of* $\mathbf{A}$ *(with full rank) is given. Formulate a method to compute the solution of (5.2).* Remark: *Since computing an SVD is more expensive than computing a* QR*-factorization, this is rarely done in practice.* ∎

5.3.4 Further properties of the SVD

Exercise 5.21 *Let* $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^\top$ *be the SVD of a matrix* $\mathbf{A}$*. Show:*

(a) $\|\mathbf{A}\|_F^2 = \sum_i \sigma_i^2$*, where the* Frobenius norm *of* $\mathbf{A}$ *is given by* $\|\mathbf{A}\|_F^2 = \sum_{i,j} |\mathbf{A}_{ij}|^2$.

(b) $\|\mathbf{A}\|_2^2 = \max_i \sigma_i^2 = \sigma_1^2$.

We have:

Theorem 5.22 *Let* $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^\top$ *be the SVD of the matrix* $\mathbf{A}$ *with rank* r*. Let the singular values be sorted in descending order. Then for every* $\nu \in \{1, \ldots, r\}$ *the matrix* $\mathbf{A}_\nu := \mathbf{U}(:, [1 : \nu])\Sigma([1 : \nu], [1 : \nu])\mathbf{V}(:, [1 : \nu])^\top$ *satifies*

$$\|\mathbf{A} - \mathbf{A}_\nu\|_2 = \min_{\mathbf{B}\in\mathbb{R}^{m\times n}:\ \mathrm{rank}(\mathbf{B})=\nu} \|\mathbf{A} - \mathbf{B}\|_2,$$
$$\|\mathbf{A} - \mathbf{A}_\nu\|_F = \min_{\mathbf{B}\in\mathbb{R}^{m\times n}:\ \mathrm{rank}(\mathbf{B})=\nu} \|\mathbf{A} - \mathbf{B}\|_F.$$

slide 30

Remark 5.23 *The SVD can be used to determine the rank of a matrix by checking the number of non-zero singular values. In practice, one has to select a cut-off* $\varepsilon > 0$ *(typically a little larger than machine precision) and defines the rank* $r = \#\{\sigma_i \mid \sigma_i \geq \varepsilon\}$. ■

5.3.5 The Moore-Penrose Pseudoinverse (CSE)

We consider the least squares problem without conditions on m, n, and the rank of $\mathbf{A}$:

$$\text{find } \mathbf{x} \in \mathbb{R}^n \text{ s.t. } \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2 \leq \|\mathbf{A}\mathbf{y} - \mathbf{b}\|_2 \quad \forall \mathbf{y} \in \mathbb{R}^n. \tag{5.5}$$

This problem has solutions but possibly more than one. To enforce uniqueness, we seek again the "minimum norm" solution, i.e., the $\mathbf{x}^* \in \mathbb{R}^n$ with the smallest norm. We have:

Theorem 5.24 *Let* $\mathbf{A} \in \mathbb{R}^{m\times n}$ *with* $\mathrm{rank}\,\mathbf{A} = r$*. Let* $\mathbf{A} = \widetilde{\mathbf{U}}\widetilde{\Sigma}\widetilde{\mathbf{V}}^\top$ *be the economy size SVD of* $\mathbf{A}$*. Then* $\mathbf{x}^* := \mathbf{A}^+\mathbf{b}$ *with the* Moore-Penrose pseudoinverse

$$\mathbf{A}^+ := \widetilde{\mathbf{V}}\widetilde{\Sigma}^{-1}\widetilde{\mathbf{U}}^\top \tag{5.6}$$

is the minimum norm solution of the least squares problem (5.5).

Before proving Theorem 5.24, we formulate a representation of the orthogonal projection onto a subspace, which takes a particularly simple form if an orthonormal basis of the space is available:

Lemma 5.25 *Let* $\mathbf{V} \in \mathbb{R}^{n\times k}$ *have orthonormal columns. Then the map* $\mathbf{x} \mapsto \mathbf{V}\mathbf{V}^\top\mathbf{x}$ *is the orthogonal projection onto the subspace* $\mathcal{V}$ *spanned by the columns of* $\mathbf{V}$*. If* $\widetilde{\mathbf{V}} \in \mathbb{R}^{n\times(n-k)}$ *is such that* $(\mathbf{V}, \widetilde{\mathbf{V}})$ *is an orthogonal matrix (i.e., the space* $\widetilde{\mathcal{V}}$ *spanned by the columns of* $\widetilde{\mathbf{V}}$ *is the orthogonal complement of* $\mathcal{V}$*) then*

$$\mathbf{x} = \mathbf{V}\mathbf{V}^\top\mathbf{x} + \widetilde{\mathbf{V}}\widetilde{\mathbf{V}}^\top\mathbf{x} \qquad \forall \mathbf{x} \in \mathbb{R}^n. \tag{5.7}$$

Proof: We recall that the orthogonal projection $P\mathbf{x} \in \mathcal{V}$ of $\mathbf{x}$ onto $\mathcal{V}$ is characterized by

$$(\mathbf{x} - P\mathbf{x}, \mathbf{y})_2 = 0 \qquad \forall \mathbf{y} \in \mathcal{V}. \tag{5.8}$$

We now check that $P\mathbf{x} := \mathbf{V}\mathbf{V}^\top\mathbf{x}$ satisfies (5.8). We note that $\mathbf{V}\mathbf{V}^\top\mathbf{x} \in \mathcal{V}$ and that any $\mathbf{y}' \in \mathcal{V}$ can be written as $\mathbf{y}' = \mathbf{V}\mathbf{y}$ for some $\mathbf{y} \in \mathbb{R}^k$. We compute for arbitrary $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^k$:

$$(\mathbf{x} - \mathbf{V}\mathbf{V}^\top\mathbf{x}, \mathbf{V}\mathbf{y})_2 = (\mathbf{V}^\top(\mathbf{x} - \mathbf{V}\mathbf{V}^\top\mathbf{x}), \mathbf{y})_2 = ((\mathbf{V}^\top\mathbf{x} - \underbrace{\mathbf{V}^\top\mathbf{V}}_{=\mathrm{I}}\mathbf{V}^\top\mathbf{x}), \mathbf{y})_2 = 0,$$

which shows (5.8). Similarly, $\widetilde{\mathbf{V}}\widetilde{\mathbf{V}}^\top\mathbf{x}$ is the orthogonal projection of $\mathbf{x}$ onto the space $\widetilde{\mathcal{V}}$. By construction $\mathbf{x} - \mathbf{V}\mathbf{V}^\top\mathbf{x}$ is in the orthogonal complement of $\mathcal{V}$, i.e., in the space $\widetilde{\mathcal{V}}$. Hence, by the projection property $\widetilde{\mathbf{V}}\widetilde{\mathbf{V}}^\top(\mathbf{x} - \mathbf{V}\mathbf{V}'\mathbf{x}) = \mathbf{x} - \mathbf{V}\mathbf{V}'\mathbf{x}$. Since $\widetilde{\mathbf{V}}^\top\mathbf{V} = 0$, we obtain (5.7) by rearranging the terms. □

Proof of Theorem 5.24: We decompose $\mathbf{b}$ into its component in $\operatorname{Im}\mathbf{A}$ and the rest using Lemma 5.25:

$$\mathbf{b} = \widetilde{\mathbf{U}}\widetilde{\mathbf{U}}^\top\mathbf{b} + \mathbf{U}'(\mathbf{U}')^\top\mathbf{b}, \qquad \mathbf{U}' := \mathbf{U}(:, [r+1:n]).$$

Next, we compute for arbitrary $\mathbf{x} \in \mathbb{R}^n$

$$\begin{aligned}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 &= \|\widetilde{\mathbf{U}}\widetilde{\Sigma}\widetilde{\mathbf{V}}^\top\mathbf{x} - \mathbf{b}\|_2^2 = \|\widetilde{\mathbf{U}}(\widetilde{\Sigma}\widetilde{\mathbf{V}}^\top\mathbf{x} - \widetilde{\mathbf{U}}^\top\mathbf{b}) + \mathbf{U}'(\mathbf{U}')^\top\mathbf{b}\|_2^2 \\ &= \|\widetilde{\mathbf{U}}(\widetilde{\Sigma}\widetilde{\mathbf{V}}^\top\mathbf{x} - \widetilde{\mathbf{U}}^\top\mathbf{b})\|_2^2 + \|\mathbf{U}'(\mathbf{U}')^\top\mathbf{b}\|_2^2 \\ &= \|\widetilde{\Sigma}\widetilde{\mathbf{V}}^\top\mathbf{x} - \widetilde{\mathbf{U}}^\top\mathbf{b}\|_2^2 + \|\mathbf{U}'(\mathbf{U}')^\top\mathbf{b}\|_2^2\end{aligned}$$

This expression is minimal if we can find $\mathbf{x}$ such that

$$\widetilde{\mathbf{V}}^\top\mathbf{x} = \widetilde{\Sigma}^{-1}\widetilde{\mathbf{U}}^\top\mathbf{b}. \tag{5.9}$$

(We will see at the end of the proof that indeed such $\mathbf{x}$ exist.) Let us now seek the $\mathbf{x}^*$ from all $\mathbf{x}$ satisfying (5.9) with minimal norm. We write again with Lemma 5.25

$$\mathbf{x} = \widetilde{\mathbf{V}}\widetilde{\mathbf{V}}^\top\mathbf{x} + \mathbf{V}'(\mathbf{V}')^\top\mathbf{x}.$$

Hence, any $\mathbf{x}$ that satisfies (5.9) has to satisfy

$$\|\mathbf{x}\|_2^2 = \|\widetilde{\mathbf{V}}\widetilde{\mathbf{V}}^\top\mathbf{x}\|_2^2 + \|\mathbf{V}'(\mathbf{V}')^\top\mathbf{x}\|_2^2 \overset{(5.9)}{=} \|\widetilde{\mathbf{V}}\widetilde{\Sigma}\widetilde{\mathbf{U}}^\top\mathbf{b}\|_2^2 + \|\mathbf{V}'(\mathbf{V}')^\top\mathbf{x}\|_2^2.$$

We see that $\mathbf{x}^*$ with the smallest norm should be such that $(\mathbf{V}')^\top\mathbf{x}^* = 0$. Then, we get

$$\mathbf{x}^* \overset{(\mathbf{V}')^\top\mathbf{x}^*=0}{=} \widetilde{\mathbf{V}}\widetilde{\mathbf{V}}^\top\mathbf{x}^* \overset{(5.9)}{=} \widetilde{\mathbf{V}}\widetilde{\Sigma}^{-1}\widetilde{\mathbf{U}}^\top\mathbf{b}.$$

Indeed, this $\mathbf{x}^*$ satisfies $(\mathbf{V}')^\top\mathbf{x}^* = 0$ as well as (5.9). Hence, we have found the unique minimum norm solution. □

Let us interpret the Moore-Penrose pseudoinverse. To that end, we let us restrict $\mathbf{A}$ to $(\operatorname{Ker}\mathbf{A})^\perp$, which we denote by $\mathbf{A}_K$ to emphasize that the domain of definition and range has changed:

$$\begin{aligned}\mathbf{A}_K : (\operatorname{Ker}\mathbf{A})^\perp &\to \operatorname{Im}\mathbf{A} \\ \widetilde{\mathbf{V}}z &\mapsto \mathbf{A}\widetilde{\mathbf{V}}z = \widetilde{\mathbf{U}}\widetilde{\Sigma}\widetilde{\mathbf{V}}^\top\widetilde{\mathbf{V}}z = \widetilde{\mathbf{U}}\widetilde{\Sigma}z\end{aligned}$$

This map is a bijection. Indeed, since the columns of $\widetilde{\mathbf{U}}$ and $\widetilde{\mathbf{V}}$ are linearly independent, the inverse $\mathbf{A}_K^{-1}$ is easily read off to be:[2]

$$\mathbf{A}_K^{-1} : \widetilde{\mathbf{U}}\zeta \mapsto \widetilde{\mathbf{V}}\widetilde{\Sigma}^{-1}\zeta$$

We now consider

$$\begin{array}{ccccc} \mathbb{R}^m & \overset{\text{ortho. Proj.}}{\longrightarrow} & \operatorname{Im}\mathbf{A} & \overset{\mathbf{A}_K^{-1}}{\longrightarrow} & \operatorname{Ker}\mathbf{A} \\ \mathbf{b} & \mapsto & \widetilde{\mathbf{U}}(\widetilde{\mathbf{U}}^\top\mathbf{b}) & \mapsto & \widetilde{\mathbf{V}}\widetilde{\Sigma}^{-1}\widetilde{\mathbf{U}}^\top\mathbf{b} \end{array}$$

This is precisely $\mathbf{A}^+$! Hence, the Moore-Penrose pseudoinverse takes from a vector $\mathbf{b}$ its component in $\operatorname{Im}\mathbf{A}$ and then applies the well-defined inverse $\mathbf{A}_K^{-1}$ that maps from $\operatorname{Im}\mathbf{A}$ to $(\operatorname{Ker}\mathbf{A})^\perp$.

Exercise 5.26 *Let* $\operatorname{rank}\mathbf{A} = r$. *Show:* $\|\mathbf{A}^+\|_2 = \sigma_r^{-1}$. ■

5.3.6 Further remarks

- The Moore-Penrose pseudoinverse is the inverse of $\mathbf{A}$ if $\mathbf{A} \in \mathbb{R}^{n\times n}$ is invertible.
- In general, $\mathbf{A}^+$ shares some properties with the inverse: $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$ and $(\mathbf{A}^+)^+ = \mathbf{A}$.

Computing the SVD

The SVD is computed with variants of algorithms that compute eigenvalues and eigenvectors. Since $\mathbf{A}^\top\mathbf{A} = \mathbf{V}^\top\Sigma^\top\Sigma\mathbf{V}$ and $\mathbf{A}\mathbf{A}^\top = \mathbf{U}^\top\Sigma\Sigma^\top\mathbf{U}^\top$, one could compute the SVD by computing the eigenvalues and eigenvectors of $\mathbf{A}^\top\mathbf{A}$ or $\mathbf{A}\mathbf{A}^\top$. However, since $\mathbf{A}^\top\mathbf{A}$ and $\mathbf{A}\mathbf{A}^\top$ are typically ill conditioned, one resorts to computing the eigenvalues and eigenvectors of the symmetric matrix

$$\begin{pmatrix} 0 & \mathbf{A}^\top \\ \mathbf{A} & 0 \end{pmatrix},$$

whose eigenvalues are $\pm\sigma_i$. A popular algorithm for the SVD is → Golub-Kahan.

[2] An alternative way to see that $\mathbf{A}_K$ is a bijection is to check the dimensions: $\dim(\operatorname{Ker}\mathbf{A})^\perp = n - \dim\operatorname{Ker}\mathbf{A}$ and by a linear algebra fact $n = \dim(\operatorname{Ker}\mathbf{A})^\perp + \dim\operatorname{Im}\mathbf{A}$ so that $\dim(\operatorname{Ker}\mathbf{A})^\perp = \dim\operatorname{Im}\mathbf{A}$