8. SINGULAR VALUE DECOMPOSITION

8.1 Introduction

A matrix $A \in \mathbf{F}^{nxn}$ can be transformed to the Jordan canonical form using similarity transformations. Next we try to find the simplest possible "almost-diagonal" form for a general *mxn* matrix *A* by using unitary transformation matrices. We end up with a decomposition that has unbelievable theoretical and practical value, as we shall see.

8.2 Constructing singular value decomposition

Theorem 8.2.1. Let $A \in \mathbf{F}^{mxn}$. There exists unitary *mxm* and *nxn* matrices *U* and *V* such that

$$U^*AV = \Lambda,$$

where Λ is an *mxn* real diagonal matrix where

$$\Lambda = diag(\sigma_1, \sigma_2, \ldots, \sigma_r, 0, \ldots, 0) = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_r & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix},$$

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0.$$

Proof. A^*A is a positive semidefinite *nxn* Hermitian matrix. It has real nonnegative eigenvalues. Listing these eigenvalues in descending order we get

$$\sigma(A^*A) = \left\{ \sigma_1^2, \sigma_2^2, \ldots, \sigma_r^2, \sigma_{r+1}^2, \ldots, \sigma_n^2 \right\}.$$

We assume that $n-r$ of them are zeros; $\sigma_{r+1}{}^2$, $\sigma_{r+2}{}^2, \ldots, \sigma_n{}^2 = 0$. The rest of the eigenvalues $\sigma_1{}^2$, $\sigma_2{}^2, \ldots, \sigma_r{}^2$ are positive. We form a unitary *nxn* matrix consisting of the orthonormal eigenvectors $\boldsymbol{v}_i$ corresponding to $\sigma_i{}^2$.

$$V = \left[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \right].$$

These vectors form a basis of $\mathbf{F}^n$. We try to form a basis for $\mathbf{F}^m$ also. First we choose the vectors

$$\boldsymbol{u}_i = \frac{A\boldsymbol{v}_i}{\sigma_i} \quad i = 1, 2, \ldots, r.$$

They are orthonormal, because

$$< \boldsymbol{u}_i, \boldsymbol{u}_j >= \frac{< \boldsymbol{v}_i, A * A\boldsymbol{v}_j >}{(\sigma_i \sigma_j)} = \sigma_j^2 \frac{< \boldsymbol{v}_i, \boldsymbol{v}_j >}{(\sigma_i \sigma_j)} = \delta_{ij},$$

where δ_{ij} is Kronecker's delta. We extend the set formed by the vectors $\boldsymbol{u}_i$, i=1,...,r, to an orthonormal basis $\{\boldsymbol{u}_1,\boldsymbol{u}_2,..,\boldsymbol{u}_r,\boldsymbol{u}_{r+1},..,\boldsymbol{u}_m\}$ of $\mathbf{F}^m$ and choose the unitary matrix U to be

$$U = \left[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r, \boldsymbol{u}_{r+1}, \ldots, \boldsymbol{u}_m\right].$$

That is all we need to do! Now we just show that $A = U\Lambda V^*$, or $U^*AV = \Lambda$.

$$U * AV = U * A\left[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\right] = U * \left[A\boldsymbol{v}_1, A\boldsymbol{v}_2, \ldots, A\boldsymbol{v}_n\right]$$

$$= U * \left[\sigma_1 \boldsymbol{u}_1, \sigma_2 \boldsymbol{u}_2, \ldots, \sigma_r \boldsymbol{u}_r, A\boldsymbol{v}_{r+1}, \ldots, A\boldsymbol{v}_n\right]$$

$$= U * \left[\sigma_1 \boldsymbol{u}_1, \sigma_2 \boldsymbol{u}_2, \ldots, \sigma_r \boldsymbol{u}_r, \mathbf{0}, \ldots, \mathbf{0}\right]$$

$$= \left[\sigma_1 U * \boldsymbol{u}_1, \sigma_2 U * \boldsymbol{u}_2, \ldots, \sigma_r U * \boldsymbol{u}_r, \mathbf{0}, \ldots, \mathbf{0}\right]$$

$$= \left[\sigma_1 \boldsymbol{e}_1, \sigma_2 \boldsymbol{e}_2, \ldots, \sigma_r \boldsymbol{e}_r, \mathbf{0}, \ldots, \mathbf{0}\right] = \Lambda$$

We used the result $\mathcal{N}(A^*A) = \mathcal{N}(A)$, which implies $A\boldsymbol{v}_i = \mathbf{0}$, when $i = r+1,\ldots,n$. Direct computations show us that $U^*\boldsymbol{u}_i = \boldsymbol{e}_i$, i=1,...,r. ❑

Definition* 8.2.1** The decomposition $A = U\Lambda V^*$ constructed above is ***singular value decomposition or abbreviated to ***SVD***. The diagonal elements σ_i of the matrix Λ are called ***singular values.***

NB. If $A \in \mathbf{R}^{m\mathrm{x}n}$ *,* then U and V are orthogonal matrices.

Example 9.2.1 Let

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

then

$$A * A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix},$$

and

$$\sigma(A^*A) = \{10,0\}, \quad \sigma_1^2 = 10, \quad \sigma_2^2 = 0$$

so

$$\Lambda = \begin{pmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{pmatrix}.$$

We compute the orthonormal eigenvectors of A^*A

$$\sigma_1^2 = 10, \qquad \boldsymbol{v}_1 = \frac{1}{\sqrt{2}}[1,1]^T$$

$$\sigma_2^2 = 0, \qquad \boldsymbol{v}_2 = \frac{1}{\sqrt{2}}[1,-1]^T,$$

we see that r=1 and

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

Next we form the matrix U

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{10}} A\boldsymbol{v}_1 = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

the vector $\boldsymbol{u}_2$ is chosen so that it is orthonormal to $\boldsymbol{u}_1$, for example

$$\boldsymbol{u}_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

will do. Hence

$$U = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}.$$

We double-check the result.

$$U\Lambda V^* = \frac{1}{\sqrt{2}}\frac{1}{\sqrt{5}}\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}.$$

Exercise 8.2.1 Prove, that the matrices A^*A and AA^* have the same nonzero eigenvalues $\sigma_1^2, \sigma_2^2 \ldots, \sigma_r^2$ and that the column vectors of V are the orthonormal eigenvectors of A^*A and that the column vectors of U are the orthonormal eigenvectors of AA^*. How can these results be used to construct the singular value decomposition?

Note that singular value decomposition is not unique. Eigenvectors corresponding to equal eigenvalues can be chosen in any order (in V). Also the vectors that are needed when we extend the basis to form U can be chosen in any order. Singular values are unique however.

***Exercise* 8.2.2** Let A be positive, semidefinite and hermitian. Prove the claim: The singular value decomposition of A is its spectral representation.

***Exercise* 8.2.3** What happens in the previous exercise if the matrix A is Hermitian but not necessarily positive semidefinite.

Computation times. Numerically SVD is not computed as above (see Golub 1983). The computation time of SVD depends on what parts of the decomposition we want to have. If we want to have the whole decomposition it takes approximately $\boldsymbol{13n}^3$ flops for an nxn matrix.

8.3 Finding bases for the range and null space by using SVD

***Theorem* 8.3.1** Let $A \in \mathbf{F}^{mxn}$ and $A = U\Lambda V^*$ its singular value decomposition. We write the matrices U and V in the terms of their columns

$$U = \left[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\right],$$

$$V = \left[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\right],.$$

Now

a) $rank(A) = r$

b) $\mathcal{R}\left(A\right) = span\left\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r\right\}$

c) $\mathcal{N}(A) = span\left\{\boldsymbol{v}_{r+1}, \boldsymbol{v}_{r+2}, \ldots, \boldsymbol{v}_n\right\}$

Proof. b) $(\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r)$ are orthonormal eigenvectors, so they are linearly independent. Let $\boldsymbol{y} \in \mathcal{R}(A)$, there exists an $\boldsymbol{x}$ such that

$$\boldsymbol{y} = A\boldsymbol{x} = U\Lambda V^*\boldsymbol{x} = \left[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\right] \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_r & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} V^*\boldsymbol{x} =$$

$$= \left[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m \right] \begin{bmatrix} \sigma_1 z_1 \\ \sigma_2 z_2 \\ \vdots \\ \sigma_r z_r \\ 0 \\ 0 \end{bmatrix} = \sigma_1 z_1 \boldsymbol{u}_1 + \sigma_2 z_2 \boldsymbol{u}_2 + \ldots + \sigma_r z_r \boldsymbol{u}_r \quad ,$$

where $\boldsymbol{z} = V^* \boldsymbol{x}$.

We have shown that $\boldsymbol{y} \in span\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r\}$, so

$$\mathcal{R}(A) \subset span \left\{ \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r \right\}.$$

For the other direction we assume that $\boldsymbol{y} \in span\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r\}$. Now

$$\boldsymbol{y} = \sum_{i=1}^{r} \alpha_i \boldsymbol{u}_i.$$

We choose $\boldsymbol{x} = V [\alpha_1/\sigma_1, \alpha_2/\sigma_2, \ldots, \alpha_r/\sigma_r, 0, .., 0]^T$. Direct computation shows, that $\boldsymbol{y} = A\boldsymbol{x}$, hence $\boldsymbol{y} \in \mathcal{R}(A)$. Now we have shown that $span\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r\} \subset \mathcal{R}(A)$ and b) is proven.

The vector set $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r\}$ is an orthonormal basis for $\mathcal{R}(A)$ and so

$$rank\,(A) = dim\ \mathcal{R}(A) = r.$$

This is just the claim a).

c) We denote, as earlier, $\boldsymbol{x} = V\boldsymbol{z}$.

$$A\boldsymbol{x} = \mathbf{0} \Leftrightarrow U \Lambda V^* \boldsymbol{x} = \mathbf{0} \Leftrightarrow \Lambda V^* \boldsymbol{x} = \mathbf{0} \Leftrightarrow$$

$$\Lambda \boldsymbol{z} = \mathbf{0} \Leftrightarrow \boldsymbol{z} = \left[0, 0, \ldots, 0, z_{r+1}, \ldots, z_n \right]^T,$$

hence

$$\boldsymbol{x} = V\boldsymbol{z} = \left[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \right] \begin{bmatrix} 0 \\ 0 \\ z_{r+1} \\ \vdots \\ z_n \end{bmatrix} = \sum_{i=r+1}^{n} z_i \boldsymbol{v}_i \ .$$

The last equation shows that $\boldsymbol{x}$ is a linear combination of the vectors $\boldsymbol{v}_i$ if and only if

$$\boldsymbol{x} \in span\{\boldsymbol{v}_{r+1},\ldots,\boldsymbol{v}_n\},$$

so

$$\mathcal{N}(A) = span\{\boldsymbol{v}_{r+1}, \boldsymbol{v}_{r+2}, \ldots, \boldsymbol{v}_n\}. \square$$

Corollary 8.3.1. The vectors $\boldsymbol{v}_i$ are orthonormal and so we have proven the rank-nullity theorem for the second time.

$$dim \;\; \mathcal{N}(A) = n - r = n - rank(A).$$

Corollary 8.3.2. The matrix $A \in \mathbf{F}^{\mathrm{nxn}}$ is nonsingular $\Leftrightarrow$ $rank(A) = n \Leftrightarrow \sigma_n > 0$. The inverse of A in terms of SVD is

$$A^{-1} = (U\Lambda V^*)^{-1} = V\Lambda^{-1}U^*$$

$$= V\begin{bmatrix} 1/\sigma_1 & 0 & .. & 0 \\ 0 & 1/\sigma_2 & .. & 0 \\ 0 & .. & 1/\sigma_r & 0 \\ 0 & & .. & 1/\sigma_n \end{bmatrix}U^*$$

This is not exactly the singular value decomposition (why?), but almost.

8.4 Matrix norms and SVD

Theorem 8.4.1 Let $A \in \mathbf{F}^{\mathrm{mxn}}$. Now

a) $\|A\| = \sigma_1$.

If $A \in \mathbf{F}^{\mathrm{nxn}}$ is nonsingular, then

b) $\sigma_n \|\boldsymbol{x}\| \le \|A\boldsymbol{x}\| \le \sigma_1 \|\boldsymbol{x}\| \quad \forall \;\; \boldsymbol{x} \in \mathbf{F}^{\mathrm{n}}$.

c) $\|A^{-1}\| = 1/\sigma_n$

Proof. In this proof we will repeatedly use the result that multiplication by a unitary matrix does not change the value of the norm.

a) $$\|A\| = \max_{\|\boldsymbol{x}\|=1} \|A\boldsymbol{x}\| = \max_{\|\boldsymbol{x}\|=1} \|U^* A\boldsymbol{x}\| = \max_{\|V\boldsymbol{y}\|=1} \|U^* AV\boldsymbol{y}\|$$

$$= \max_{\|\boldsymbol{y}\|=1} \|\Lambda \boldsymbol{y}\| = \max_{\|\boldsymbol{y}\|=1} \sqrt{\sigma_1^2 |y_1|^2 + \sigma_2^2 |y_2|^2 + \ldots + \sigma_r^2 |y_r|^2} = \sigma_1.$$

The last equation can be written in the following way: because σ_1 is the greatest singular value,

$$\sqrt{\sigma_1^2 |y_1|^2 + \sigma_2^2 |y_2|^2 + \ldots + \sigma_{\mathbf{r}}^2 |y_{\mathbf{r}}|^2} \le \sigma_1 \sqrt{|y_1|^2 + \ldots + |y_{\mathbf{n}}|^2}$$

$$= \sigma_1 \|\boldsymbol{y}\|,$$

hence

$$\max_{\|\boldsymbol{y}\|=1} \|\Lambda \boldsymbol{y}\| \le \sigma_1.$$

By choosing $\boldsymbol{y} = \boldsymbol{e}_1$ we see that the upper bound is attained and so

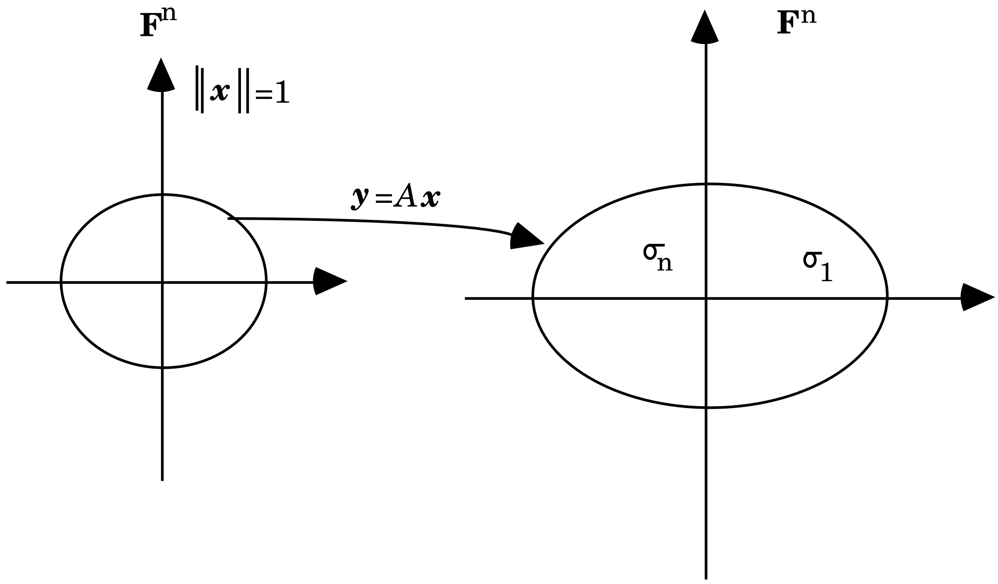
$$\max_{\|\boldsymbol{y}\|=1} \|\Lambda \boldsymbol{y}\| = \sigma_1.$$

b) Because $\|A\boldsymbol{x}\| \le \|A\|\|x\|$, and $\|A\| = \sigma_1$ (see a)) we get the upper bound $\|\boldsymbol{A}x\| \le \sigma_1 \|x\|$.

The proof of the lower bound is similar to the proof of a)

$$\|A\boldsymbol{x}\| \ge \sigma_n \sqrt{|y_1|^2 + |y_2|^2 + \ldots + |y_n|^2} = \sigma_n \|\boldsymbol{y}\| = \sigma_n \|V^* \boldsymbol{x}\| = \sigma_n \|\boldsymbol{x}\|.$$

c) Here we use the result that if $A = U\Lambda V^*$ is SVD then $A^{-1} = V\Lambda^{-1}U^*$. We continue the proof as we did in a). ❑



Picture 8.4.1. If the column vectors of V and U are chosen to be basis vectors of $\mathbf{F}^n$ and $\mathbf{F}^m$, then the matrix is transformed to Λ. In this transformation a unitball of $\mathbf{F}^n$ is transformed to an ellipsoid in $\mathbf{F}^m$, where the semi-axes are of length $\sigma_1, \sigma_2, \ldots, \sigma_n$.

Exercise 8.4.1 Form the singular value decompositions and compute the norms of the matrices

$$a)\ A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},\ b)\ A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix},\ c)\ A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Exercise 8.4.2 Use singular value decomposition to prove that

$$\text{a)}\quad \|\boldsymbol{A}\| = \max_{\|x\|\le 1, \|y\|\le 1} \left| \frac{<x, \boldsymbol{A}y>}{\|y\|\ \|x\|} \right| \quad \text{ja b)}\quad \|A\| = \|A^*A\|^{1/2}.$$

Exercise 8.4.3 Let A be a nonsingular matrix. Find an upper bound for the norm of B so that $A+B$ is nonsingular.

Hint: If $\|B\| < \sigma_n$, where σ_n is the smallest singular value of A, then $A+B$ is nonsingular.

8.5 Approximation of a matrix

Theorem 8.5.1 Let $A \in \mathbf{F}^{n\mathrm{x}n}$, $rank(A) = r$ and $A = U\Lambda V^*$ its SVD. The best approximation of A (with respect to matrix norm) using matrices with rank $< r$, is

$$B = U \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_{r-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} V^*$$

Proof. We denote

$$A = U\Lambda_A V^* \quad \Lambda_A = diag(\sigma_1, \sigma_2, ..., \sigma_{r-1}, \sigma_r, 0, ..., 0)$$

$$B = U\Lambda_B V^* \quad \Lambda_B = diag(\sigma_1, \sigma_2, ..., \sigma_{r-1}, 0, ..., 0).$$

The difference between the matrices A and B is

$$\|A - B\| = \max_{\|\boldsymbol{x}\|=1} \|(A - B)\boldsymbol{x}\| = \max_{\|\boldsymbol{x}\|=1} \left\|U\left(\Lambda_A - \Lambda_B\right)V^*\boldsymbol{x}\right\|$$

$$= \max_{\|V\boldsymbol{y}\|=1} \left\|\left(\Lambda_A - \Lambda_B\right)\boldsymbol{y}\right\| = \max_{\|V\boldsymbol{y}\|=1} \sigma_r |y_r| = \sigma_r,$$

We choose some other matrix F, with rank $k < r$. So rank($\mathcal{N}(F)$) = n-k. Let $\{\boldsymbol{x}_1, \boldsymbol{x}_2, .., \boldsymbol{x}_{n-k}\}$ be an orthonormal basis of $\mathcal{N}(F)$. The intersection of the spaces

$$\mathcal{N}(F) \cap span\{\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_r\}$$

is $\neq \{\mathbf{0}\}$, or otherwise

$$dim(\mathbf{F}^{\mathrm{n}}) \geq dim\ \mathcal{N}(F) + r = n - k + r > n.$$

Let $\boldsymbol{x} \in \mathrm{N}(F) \cap span\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r\}$ so, that $\|\boldsymbol{x}\| = 1$. For this $\boldsymbol{x}$

$$A\boldsymbol{x} - F\boldsymbol{x} = A\boldsymbol{x} = A\sum_{i=1}^{r} < \boldsymbol{v}_i, \boldsymbol{x} > \boldsymbol{v}_i = \sum_{i=1}^{r} < \boldsymbol{v}_i, \boldsymbol{x} > A\boldsymbol{v}_i$$

$$= \sum_{i=1}^{r} \sigma_i < \boldsymbol{v}_i, \boldsymbol{x} > \boldsymbol{u}_i$$

Furthermore because $\{\boldsymbol{u}_i\}$ is an orthonormal basis

$$\|A\boldsymbol{x} - F\boldsymbol{x}\|^2 = \|A\boldsymbol{x}\|^2 = \sum_{i=1}^{r} \sigma_i^2 \left|\langle\boldsymbol{v}_i, \boldsymbol{x}\rangle\right|^2 \geq \sum_{i=1}^{r} \sigma_r^2 \left|\langle\boldsymbol{v}_i, \boldsymbol{x}\rangle\right|^2$$

$$= \sigma_r^2 \sum_{i=1}^{r} \left| < \boldsymbol{v}_i, \boldsymbol{x} > \right|^2 = \sigma_r^2 \quad ,$$

(because $\|\boldsymbol{x}\| = 1$) and so

$$\|A - F\| = \max_{\|\boldsymbol{x}\|=1} \|A\boldsymbol{x} - F\boldsymbol{x}\| \geq \sigma_r. \quad ◆$$

***Corollary* 8.5.1** We can generalize the previous result in the following way:

$$\min_{rank(F)=k} \|A - F\| = \|A - A_k\| = \sigma_{k+1} \quad ,$$

where

$$A_k = U \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & 0 \\ 0 & & \sigma_k & \\ & & & 0 \end{bmatrix} V^*$$

is the best approximation of A by using matrices of rank k.

8.6 Computing the rank of a matrix

The rank of a matrix A can be computed fairly reliably by using singular value decomposition. Lets say that the rank is r. Due to numerical errors linearly dependent rows and columns may become linearly independent and the rank of the matrix may grow. In this kind of situation A has r eigenvectors that clearly differ from zero and the rest of the singular values are close to zero. If we set the very small singular values to be zero we get the right rank. Of course this method requires consideration and the right kind of estimations because there are matrices with very small singular values.