Important results from linear algebra

These notes contain some important results from linear algebra, which you will need in order to understand SVD. Notes below are FIRST EDITION. So please correct errors/typos by mail.

1 Matrices

Definition 1: An $m \times n$ **matrix A** consists of m rows and n columns of numbers. The number in the i'th row and j'te column is denoted $\mathbf{A}_{ij}$ or a_{ij}. An $m \times n$ matrix $\mathbf{A}$ can be multiplied with an $n \times k$ matrix $\mathbf{B}$ to get an $m \times k$ matrix $\mathbf{AB}$. The formula is

$$(\mathbf{AB})_{ij} = \sum_{s=1}^{n} a_{is} b_{sj}$$

An n-dimensional **vector x** consists of n numbers, $\mathbf{x} = (x_1, \ldots, x_n)$ and one can think of it as an $n \times 1$ matrix (column vector). The multiplication $\mathbf{Ax}$ of an $m \times n$ matrix $\mathbf{A}$ and an n dimensional vector $\mathbf{x}$ is calculated as

$$(\mathbf{Ax})_i = \sum_{j=1}^{n} a_{ij} x_j$$

Definition 2: Consider an $m \times n$ matrix $\mathbf{A}$. The **transposed** matrix, called $\mathbf{A}^T$ is the $n \times m$ matrix definered as $\mathbf{A}^T_{ij} = \mathbf{A}_{ji} \quad \forall i, j$.

Theorem 1: The following results hold

i) Consider arbitrary $m \times j$, $j \times k$ and $k \times n$ matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$. Then $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$.

ii) Consider arbitrary $m \times n$ matrix $\mathbf{A}$, n-dimensional vectors $\mathbf{x}$ and $\mathbf{y}$ and scalars a, b. We then have $\mathbf{A}(a\mathbf{x} + b\mathbf{y}) = a\mathbf{Ax} + b\mathbf{Ay}$

iii) Consider arbitrary $m \times n$ matrices $\mathbf{A}$, $\mathbf{B}$ and an arbitrary n-dimensional vectors $\mathbf{x}$. We then have that $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$ and $(\mathbf{Ax})^T = \mathbf{x}^T\mathbf{A}^T$.

The proof can be done by directly applying Definition 1 and Definition 2 and is left to the reader.

2 Linear independency. Rank of a matrix.

Definition 3: The vectors $\mathbf{x}_1, \ldots, \mathbf{x}_K$ in $\mathbb{R}^n$ are **linearly independent** if the expression

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_K\mathbf{x}_K = 0$$

implies that

$$c_1 = c_2 = \ldots = c_K = 0$$

In the following we will study $m \times n$ matrices where $m \geq n$.

Definition 4: Let $\mathbf{A}$ be an arbitrary $m \times n$ matrix where $m \geq n$. Now perceive the columns of $\mathbf{A}$ as vectors in $\mathbb{R}^m$. The maximum number of linearly independent columns in $\mathbf{A}$ is denoted the **rank** of $\mathbf{A}$. Furthermore, if the n columns of $\mathbf{A}$ all are linearly independent $\mathbf{A}$ is said to have full rank, i.e. the rank is n.

If $\mathbf{A}$ does not have have full rank it is said to be **singular**. Hence, there will be $c_1, \ldots, c_n$ not all zero so that $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_n\mathbf{x}_n = 0$ where $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ are the columns of $\mathbf{A}$. Writing $\mathbf{c} = (c_1, \ldots, c_n)$, we then get $\mathbf{Ac} = \mathbf{0_m}$, where $\mathbf{0_m}$ is the m-dimensional vector with only zero elements. We thus also get $(\mathbf{A}^T\mathbf{A})\mathbf{c} = \mathbf{A}^T(\mathbf{Ac}) = 0$. Vice versa, if $\mathbf{A}^T\mathbf{A}$ is singular, we have a nonzero $\mathbf{c}$ so that $\mathbf{A}^T\mathbf{Ac} = 0$. Hence, $0 = \mathbf{c}^T(\mathbf{A}^T\mathbf{Ac}) = (\mathbf{Ac})^T(\mathbf{Ac}) = \|\mathbf{Ac}\|^2$. Hence, $\mathbf{Ac} = 0$ and $\mathbf{A}$ is also singular. We then have the result:

Theorem 2: For an arbitrary $m \times n$ matrix $\mathbf{A}$, where $m \geq n$, the matrix $\mathbf{A}^* = \mathbf{A}^T\mathbf{A}$ is singular if and only if $\mathbf{A}$ is singular.

3 Orthogonal and orthonormal vectors and matrices

Definition 5: The vectors $\mathbf{x}_1, \mathbf{x}_2$ in a vectorspace $\mathbb{R}^n$ are said to be **orthogonal** if $\mathbf{x}_1 \cdot \mathbf{x}_2 = 0$. Furthermore, if $\mathbf{x}_1 \cdot \mathbf{x}_1 = \mathbf{x}_2 \cdot \mathbf{x}_2 = 1$ then $\mathbf{x}_1, \mathbf{x}_2$ are said to be **orthonormal**.

Definition 6: An $m \times n$ matrix $\mathbf{U}$ ($m \geq n$) is said to be **column orthonormal** if it consists of n pairwise orthonormal columns (the columns are perceived as vectors in $\mathbb{R}^n$).

If $\mathbf{U}$ is an $m \times n$ column orthonormal matrix then

$$(\mathbf{U}^T\mathbf{U})_{ij} = \sum_{k=1}^{m} \mathbf{U}_{ki}\mathbf{U}_{kj} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

Hence, $\mathbf{U}^T\mathbf{U} = \mathbf{I}_n$ where $\mathbf{I}_n$ is the $n \times n$ identitymatrix. As the identitymatrix is non-singular, we obtain from Theorem 2 that $\mathbf{U}$ is non-singular and hence we have shown

Theorem 3: Orthonormal vectors are always linearly independent.

We also define a special case of column orthonormal matrices:

Definition 7: An $n \times n$ matrix $\mathbf{V}$ is said to be **orthonormal** if it consists of n pairwise orthonormal columns (the columns are perceived as vectors in $\mathbb{R}^n$).

4 Subspaces, bases and orthonormal bases.

Definition 8: Consider $S \subseteq \mathbb{R}^n$. If for arbitrary elements $\mathbf{x}_1, \mathbf{x}_2 \in S$ and arbitrary scalars $c_1, c_2 \in \mathbb{R}$ applies that $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 \in S$ then S is said to be a *vector space*. The vector space S is also said to be a **subspace** of $\mathbb{R}^n$.

Notice: $\mathbb{R}^n$ is thus a vector space.

Definition 9: Let $S \subseteq \mathbb{R}^n$ be a vector space. The **dimension** of the vector space is the maximum number of linearly independent vectors, that can be found in S. Or stated differently: if $\mathbf{u}_1, \ldots, \mathbf{u}_K \in S$ are linearly independent and meet the condition that every $y \in S$ can be written as $y = c_1\mathbf{u}_1 + \ldots + c_K\mathbf{u}_K$ then S has dimension K. The vectors $\mathbf{u}_1, \ldots, \mathbf{u}_K$ is then said to form a **basis** for S. If $\mathbf{u}_1, \ldots, \mathbf{u}_K$ are orthonormal, $\mathbf{u}_1, \ldots, \mathbf{u}_K$ is said to form an orthonormal **basis** for S.

Notice: Hence $\mathbb{R}^n$ has dimension n.

Assume that the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ form an orthonormal basis for $\mathbb{R}^n$. We know then that an arbitrary $\mathbf{x}$ can be written as

$$\mathbf{x} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \ldots \alpha_n \mathbf{u}_n$$

We call $(\alpha_1, \ldots, \alpha_n)$ the **coordinates** of $\mathbf{x}$ w.r.t. the base $\mathbf{u}_1, \ldots, \mathbf{u}_n$ We now immediately get for all i that

$$\mathbf{x} \cdot \mathbf{u}_i = \mathbf{x} \cdot (\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \ldots \alpha_n \mathbf{u}_n) = \alpha_i$$

Hence the coordinates of $\mathbf{x}$ w.r.t. the orthonormal basis $\mathbf{u}_1, \ldots, \mathbf{u}_n$ are given by the simple formulas $\alpha_i = \mathbf{x} \cdot \mathbf{u}_i$ for $i = 1, \ldots, n$.

Theorem 4: Let $\mathbf{u}_1 \ldots, \mathbf{u}_K$ be orthonormal vectors and let

$$\mathbf{x} = \sum_{k=1}^{K} \alpha_k \mathbf{u}_k$$

Then $\|\mathbf{x}\|^2 \equiv \mathbf{x} \cdot \mathbf{x} = \sum_{k=1}^{K} \alpha_k^2$.

Proof:

$$\mathbf{x} \cdot \mathbf{x} = (\sum_{i=1}^{K} \alpha_i \mathbf{u}_i) \cdot (\sum_{j=1}^{K} \alpha_j \mathbf{u}_j) = \sum_{i=1}^{K} \sum_{j=1}^{K} \alpha_i \alpha_j (\mathbf{u}_i \cdot \mathbf{u}_j) = \sum_{k=1}^{K} \alpha_k^2$$

where orthonormality of the $\mathbf{u}_j$'s was used to obtain the last equality.

4.1 The Gram-Schmidt method

Orthonormal bases are very convenient. Luckily a systematic way of constructing an orthonormal basis for a subspace from an arbitrary basis exists: Assume that the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k$ are linearly independent and let S be the subspace for which these vectors form a basis. Now we can construct a set of orthonormal vectors $\mathbf{e}_1, \ldots, \mathbf{e}_k$, which form an orthonormal basis for S in this way:

The Gram-Schmidt method:

$\mathbf{e}_1 := \mathbf{x}_1/\|\mathbf{x}_1\|$

For $i := 2, \ldots, k$ do {

$\quad \mathbf{e}_i := \mathbf{x}_i - \sum_{j=1}^{i-1}(\mathbf{x}_i \cdot \mathbf{e}_j)\mathbf{e}_j$

$\quad \mathbf{e}_i := \mathbf{e}_i/\|\mathbf{e}_i\|$

}

5 Range, null space and SVD

Definition 10: Let $\mathbf{A}$ be an arbitrary $m \times n$ ($m \geq n$) matrix. The function $f(\mathbf{x}) = \mathbf{Ax}$ is then said to be a **linear mapping** from $\mathbb{R}^n \to \mathbb{R}^m$. The **range** of the linear mapping is the set $B(\mathbf{A}) \subseteq \mathbb{R}^m$, that meets the condition that for any $\mathbf{y} \in B(\mathbf{A})$ exists an $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{Ax} = \mathbf{y}$. **The null space** of the linear mapping is the set $N(\mathbf{A}) \subseteq \mathbb{R}^n$, that meets the condition that for any $\mathbf{x} \in N(\mathbf{A})$ it applies that $\mathbf{Ax} = 0$.

From Theorem 1 ii), we immediately see that $B(\mathbf{A})$ is a subspace of $\mathbb{R}^m$ and $N(\mathbf{A})$ is a subspace of $\mathbb{R}^n$. We are then ready to show the following result:

Theorem 5 Let $\mathbf{u}_1 \ldots, \mathbf{u}_K$ be an arbitrary orthonormal basis for $B(\mathbf{A})$. Then the least squares solution $\mathbf{x}$ that minimizes $\|\mathbf{Ax} - \mathbf{b}\|$ satisfies

$$\mathbf{Ax} = \sum_{k=1}^{K}(\mathbf{b} \cdot \mathbf{u}_k)\mathbf{u}_k \equiv \mathbf{b}_{LS}$$

Proof: First, we observe directly that $(\mathbf{Ax}-\mathbf{b}) \cdot \mathbf{u}_k = 0$ for $k = 1, \ldots, K$ and hence that $\mathbf{Ax} - \mathbf{b}$ is orthogonal to $B(\mathbf{A})$. We write $\mathbf{Ax} - \mathbf{b} = d\mathbf{u}^{\perp}$ where $\mathbf{u}^{\perp}$ is a unit vector and $d = \|\mathbf{Ax} - \mathbf{b}\|$. Consider an arbitrary $\mathbf{y} \in \mathbb{R}^n$ and let $\mathbf{z} = \mathbf{Ay} - \mathbf{Ax}$. As $B(\mathbf{A})$ is a subspace, we must have $\mathbf{z} \in B(\mathbf{A})$ and hence we can write $\mathbf{z} = \sum_{k=1}^{K} z_k\mathbf{u}_k$. We then get $\mathbf{Ay} - \mathbf{b} = (\mathbf{Ay} - \mathbf{Ax}) + (\mathbf{Ax} - \mathbf{b}) = d\mathbf{u} + \sum_{k=1}^{K} z_k\mathbf{u}_k$. By Theorem 4, we now get

$$\|\mathbf{Ay} - \mathbf{b}\| = \sum_{k=1}^{K} z_k^2 + d^2$$

If $\mathbf{z} \neq 0$, i.e. $\mathbf{Ay} \neq \mathbf{Ax}$, we get $\|\mathbf{Ay} - \mathbf{b}\| > \|\mathbf{Ax} - \mathbf{b}\|$ which proves that $\mathbf{Ax} = \sum_{k=1}^{K}(\mathbf{b} \cdot \mathbf{u}_k)\mathbf{u}_k$ is the nearest point in $B(\mathbf{A})$.

Theorem 6 Consider an arbitrary $m \times n$ matrix $\mathbf{A}$. Then we can write $\mathbf{A}$ as $\mathbf{A} = \mathbf{UWV}^T$, where $\mathbf{U}$ is an $m \times n$ column othonormal matrix, $\mathbf{V}$ is an $n \times n$ orthonormal matrix and $\mathbf{W}$ is an $n \times n$ diagonal matrix having non-negative diagonal elements $w_1, \ldots w_n$ ordered such that $w_1 \geq w_2 \geq ... \geq w_n$. This is said to be a **Singular Value Decomposition** (SVD) of $\mathbf{A}$.

This Theorem is rather complicated to show. Luckily we have an algorithm to construct the SVD (see Numerical Recipes). The following Theorem is easier to show:

Theorem 7 Consider an arbitrary $m \times n$ matrix $\mathbf{A}$ and assume that for $\mathbf{W}$ it applies that $w_1, \ldots, w_K$ are positive and $w_{K+1}, \ldots, w_n$ are equal to zero. Then it applies that

i) $N(\mathbf{A})$ has dimension $n - K$ and the last $n - K$ columns of $\mathbf{V}$ form an orthonormal basis for $N(\mathbf{A})$.

ii) $B(\mathbf{A})$ has dimension K and the first K columns of $\mathbf{U}$ form an orthonormal basis for $B(\mathbf{A})$.

iii) The SVD solution $\mathbf{x} = \mathbf{V}\tilde{\mathbf{W}}^{-1}\mathbf{U}^T\mathbf{b}$, where $[\tilde{\mathbf{W}}^{-1}]_{jj} = 0$ if $\mathbf{W}_{jj} = 0$, otherwise $[\tilde{\mathbf{W}}^{-1}]_{jj} = 1/\mathbf{W}_{jj}$, is the least squares solution to $\mathbf{Ax} = \mathbf{b}$. (Notice that it then follows that if all of the $\mathbf{W}_{jj}$'s are positive, i.e. $\mathbf{A}$ has full rank it applies that $\mathbf{x} = \mathbf{VW}^{-1}\mathbf{U}^T\mathbf{b}$ is the least squares solution to $\mathbf{Ax} = \mathbf{b}$).

Proof: As $\mathbf{V}$ is orthonormal the columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are orthonormal and thus form an orthonormal basis for $\mathbb{R}^n$. Then we can write an arbitrary $\mathbf{x} \in \mathbb{R}^n$ as $\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n$. Now we can calculate

$$\begin{aligned} \mathbf{Ax} &= \mathbf{UWV}^T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n) \\ &= c_1\mathbf{UWV}^T\mathbf{v}_1 + c_2\mathbf{UWV}^T\mathbf{v}_2 + \ldots c_n\mathbf{UWV}^T\mathbf{v}_n \end{aligned}$$

We notice that $\mathbf{V}^T\mathbf{v}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ where the 1 is at the i'th position. Insertion gives us

$$\mathbf{Ax} = c_1w_1\mathbf{u}_1 + c_2w_2\mathbf{u}_2 + \ldots c_nw_n\mathbf{u}_n$$

where $\mathbf{u}_1, \ldots, \mathbf{u}_n$ are the columns of $\mathbf{U}$. This proves i) and ii).

In order to show iii), i.e. that the SVD solution is the same as the least squares solution we exploit that we know that the first K columns of $\mathbf{U}$ form an orthonormal basis for the range of $\mathbf{A}$. I.e. that the nearest point in $B(\mathbf{A})$ (the least squares mapping) according to Theorem 5 is given by $\mathbf{b}_{LS} = \sum_{j=1}^{K} (\mathbf{u}_j \cdot \mathbf{b})\mathbf{u}_j$. Now let $\mathbf{x}$ be the SVD solution. I.e.

$$\begin{aligned}
\mathbf{Ax} &= (\mathbf{UWV}^T)(\mathbf{V}[\tilde{\mathbf{W}}^{-1}]\mathbf{U}^T\mathbf{b}) = \mathbf{U}(\mathbf{W}[\tilde{\mathbf{W}}^{-1}]\mathbf{U}^T\mathbf{b}) \\
&= [\mathbf{u}_1 \ldots \mathbf{u}_K | \mathbf{u}_{K+1} \ldots \mathbf{u}_n] \begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{b} \\ \mathbf{u}_2 \cdot \mathbf{b} \\ \ldots \\ \mathbf{u}_K \cdot \mathbf{b} \\ - - - \\ 0 \\ \ldots \\ 0 \end{bmatrix} \\
&= \sum_{j=1}^{K} (\mathbf{u}_j \cdot \mathbf{b})\mathbf{u}_j = \mathbf{b}_{LS}
\end{aligned}$$

6 Error analysis for m linear equations in n unknowns with $m \geq n$

There are two types of errors that should be checked, namely the residual error and the error on the result $\mathbf{x}\|$.

6.1 Residual error

The residual error should be computed as a relative error, namely

$$\epsilon_{residual} = \frac{\|\mathbf{Ax} - \mathbf{b}\|}{\|\mathbf{b}\|} \tag{1}$$

If $m = n$, the residual error should be very close to zero unless the matrix is near singular. For $m > n$, the linear equations are typically from some sort of

fitting problem such as a Least Squares Problem. The value $\epsilon_{residual}$ indicates how good the fitting model is. It is easy to see that a random fitting model would produce $\epsilon_{residual} \simeq \sqrt{\frac{m-n}{m}}$. If your result is not much better than that, you should consider the quality of your model.

6.2 Error on the result $\|\mathbf{x}\|$

Even though solving a set of linear equations seem very deterministic, it is relevant to consider the error $\delta\mathbf{x}$ on the result $\mathbf{x}$. In typical applications, there are two very different sources to this error.

The first source is the error on the right hand side $\delta\mathbf{b}$. The error $\delta\mathbf{b}$ is typically is some kind of measurement error and therefore may be quite large.

The second source is the error on the matrix $\delta\mathbf{A}$ which is typically due to the real number precision. Hence, $\|\delta\mathbf{A}\|$ is mostly of the order $\|\delta\mathbf{A}\| \simeq 10^{-18}$.

6.2.1 SVD analysis of impact on inaccuracy on result from errors on right hand side

The impact from measurement errors or other errors on the right hand side to the inaccuracy on the result is analyzed in NR. Assuming that we can estimate the size of the inaccuracy on each element of the right hand size, we can incorporate them in $\mathbf{A}$ and $\mathbf{b}$ as outlined in Eq.15.4.4 and Eq.15.4.5 respectively. That is

$$\begin{aligned} \mathbf{A}_{ij} &:= \mathbf{A}_{ij}/\sigma_i \quad i = 1, \ldots m, \quad j = 1, \ldots n \\ \mathbf{b}_i &:= \mathbf{b}_i/\sigma_i \quad i = 1, \ldots m \end{aligned}$$

where σ_i is the inaccuracy on $\mathbf{b}_i$.

The resulting covariances on the least squares estimate is then given in Eq.15.4.19 and Eq.15.4.20 respectively. In our notation these equations are

$$\sigma^2(x_j) = \sum_{i=1}^{n} \left(\frac{V_{ji}}{w_i} \right)^2 \tag{2}$$

and

$$\mathrm{Cov}(x_j, x_k) = \sum_{i=1}^{n} \left(\frac{V_{ji} V_{ki}}{w_i^2} \right) \tag{3}$$

It should however be noticed that this result does not rely on the problem being a Least Squares problem. Hence, it applies to any problem

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

where we assume that $\mathbf{A}$ has full rank. We just need to assume that the equations have already been scaled with the uncertainties as outlined above.

The error estimate $\delta\mathbf{x}$ is then purely given by the SVD matrices using Eq.15.4.19

$$[\delta\mathbf{x}]_j \simeq \sqrt{\sum_{i=1}^{n} \left(\frac{V_{ji}}{w_i} \right)^2} \quad j = 1, \ldots, n \tag{4}$$

For Least Squares problem this estimate holds both for solving with SVD and using Normal Equations with Cholesky or LU.

6.2.2 SVD analysis of impact on inaccuracy on result from errors on the matrix elements

Errors on the matrix $\mathbf{A}$ are mostly purely due to roundoff, and are therefore often neglected, which can be dangerous. Again, we look at the SVD representation for

$$(\mathbf{A} + \delta\mathbf{A})(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b}$$

For simplicity, we restrict ourselves to inaccuracies $\delta\mathbf{A}$ that are in the range of $\mathbf{A}$. Hence, $\delta\mathbf{A} = \mathbf{U}\mathbf{U}^T\delta\mathbf{A}$. Formulating using SVD, we get

$$\mathbf{U}\mathbf{W}\mathbf{V}^T(\mathbf{I} + \mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\delta\mathbf{A})(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b}$$

We now write the explicit SVD solution

$$(\mathbf{I} + \mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\delta\mathbf{A})(\mathbf{x} + \delta\mathbf{x}) = \mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\mathbf{b}$$

and hence

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

which yields

$$\delta\mathbf{x} = [(\mathbf{I}+\mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\delta\mathbf{A})^{-1}-\mathbf{I}]\mathbf{x}$$

We first remember that $\|\mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\delta\mathbf{A}\| \leq \|\mathbf{W}^{-1}\|\|\delta\mathbf{A}\|$. If we assume that $\|\delta\mathbf{A}\|$ is small enough so that $\|\mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\delta\mathbf{A}\| < 1$, we can apply the geometric series result

$$(\mathbf{I}+\mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\delta\mathbf{A})^{-1} = \mathbf{I}+\sum_{n=1}^{\infty}(-\mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\delta\mathbf{A})^n$$

Here, we have used a generalization of the result for geometric series: If $|x| < 1$, we have $1+x+x^2+x^3+\ldots = \frac{1}{1-x}$. It generalizes to matrices. If $\mathbf{B}$ is a square matrix and $\|\mathbf{B}\| \leq 1$, we have $\mathbf{I}+\mathbf{B}+\mathbf{B}^2+\mathbf{B}^3+\ldots = (\mathbf{I}-\mathbf{B})^{-1}$.

Insertion yields

$$\begin{aligned}\delta\mathbf{x} &= \left(\sum_{n=1}^{\infty}(-\mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\delta\mathbf{A})^n\right)\mathbf{x} \\ &= -\mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\delta\mathbf{A}\left(\sum_{n=0}^{\infty}(-\mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\delta\mathbf{A})^n\right)\mathbf{x}\end{aligned}$$

Hence,

$$\begin{aligned}\|\delta\mathbf{x}\| &\leq \|\mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\delta\mathbf{A}\|\left(\sum_{n=0}^{\infty}\|\mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\delta\mathbf{A}\|^n\right)\|\mathbf{x}\| \\ &= \frac{\|\mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\delta\mathbf{A}\|}{1-\|\mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\delta\mathbf{A}\|}\|\mathbf{x}\| \\ &\leq \frac{\|\mathbf{W}^{-1}\|\|\delta\mathbf{A}\|}{1-\|\mathbf{W}^{-1}\|\|\delta\mathbf{A}\|}\|\mathbf{x}\|\end{aligned}$$

We then obtain the following bound on the relative error

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{\|\mathbf{W}^{-1}\|\|\delta\mathbf{A}\|}{1-\|\mathbf{W}^{-1}\|\|\delta\mathbf{A}\|} \tag{5}$$

Hence, for $\|\mathbf{W}\|^{-1}\|\delta\mathbf{A}\| << 1$, the maximal impact from roundoff errors approximates to $\|\mathbf{W}\|^{-1}\|\delta\mathbf{A}\|$. However, if this value become close to 1, the error on x become potentially unbounded and the result hence useless.

If we compute $\mathbf{A}^T\mathbf{A}$ using SVD, we get

$$\mathbf{A}^T\mathbf{A} = (\mathbf{U}\mathbf{W}\mathbf{V}^T)^T(\mathbf{U}\mathbf{W}\mathbf{V}^T) = (\mathbf{V}\mathbf{W}\mathbf{U}^T)(\mathbf{U}\mathbf{W}\mathbf{V}^T) = \mathbf{V}\mathbf{W}^2\mathbf{V}^T$$

which is itself the SVD of $\mathbf{A}^T\mathbf{A}$.

We then get the SVD solution to $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$ as

$$\begin{aligned}(\mathbf{x}+\delta\mathbf{x}) &= (\mathbf{I}+\mathbf{V}\mathbf{W}^{-2}\mathbf{V}^T\delta\mathbf{A})^{-1}\mathbf{V}\mathbf{W}^{-2}\mathbf{V}^T(\mathbf{V}\mathbf{W}\mathbf{U}^T\mathbf{b}) \\ &= (\mathbf{I}+\mathbf{V}\mathbf{W}^{-2}\mathbf{V}^T\delta\mathbf{A})^{-1}\mathbf{V}\mathbf{W}^{-1}\mathbf{U}^T\mathbf{b} \\ &= (\mathbf{I}+\mathbf{V}\mathbf{W}^{-2}\mathbf{V}^T\delta\mathbf{A})^{-1}\mathbf{x}\end{aligned}$$

Performing the derivation completely equivalent to the above, we get

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{\|\mathbf{W}^{-2}\|\|\delta\mathbf{A}\|}{1-\|\mathbf{W}^{-2}\|\|\delta\mathbf{A}\|} \tag{6}$$

Here we see the problem with the Normal Equations. If for example $w_n = 10^{-9}$ for $\mathbf{A}$, we get $\|\mathbf{W}^{-2}\| = 10^{-18}$ and hence $\|\mathbf{W}^{-2}\|\|\delta\mathbf{A}\|$ becomes around one for a double precision number representation. This is exactly the problem we see in Filip with the Normal Equations.