LEAST SQUARES SOLUTION TRICKS

VESA KAARNIOJA, JESSE RAILO AND SAMULI SILTANEN

Abstract. This handout is for the course *Applications of matrix computations* at the University of Helsinki in Spring 2018. We define the linear least squares solutions and minimum norm solutions. We study their basic properties such as the relation to the normal equation and singular values. Fitting of a linear model to noisy data is also considered as an example.

Contents

1. Least squares solution and minimum norm solution

Let us define the *least squares solution* and *minimum norm solution* of the matrix equation $A\mathbf{x} = \mathbf{b}$, where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} \in \mathbb{R}^k,$$

and the matrix A has size $k \times n$.

Date: 26.3.2018.

- **Version 1.** *Suggestions and corrections could be send to jesse.railo@helsinki.fi.*

Definition 1.1. *A vector* $\widetilde{\mathbf{x}} \in \mathbb{R}^n$ *is called a* least-squares solution *of the equation* $A\mathbf{x} = \mathbf{b}$ *if*

$$\|A\widetilde{\mathbf{x}} - \mathbf{b}\| = \min_{\mathbf{x}\in\mathbb{R}^n} \|A\mathbf{x} - \mathbf{b}\|. \tag{1.1}$$

Furthermore, we give a special name for the shortest least-squares solution (in general there may be many least-squares solutions). A vector $\widetilde{\mathbf{x}}_0$ *is called the* minimum norm solution *of* $A\mathbf{x} = \mathbf{b}$ *if* $\widetilde{\mathbf{x}}_0$ *is a least-squares solution of* $A\mathbf{x} = \mathbf{b}$ *and additionally satisfies*

$$\|\widetilde{\mathbf{x}}_0\| = \min\{\|\widetilde{\mathbf{x}}\| \,:\, \widetilde{\mathbf{x}} \text{ is a least-squares solution of } A\mathbf{x} = \mathbf{b}\}. \tag{1.2}$$

The vector norm above is the Euclidean norm $\|\mathbf{x}\|^2 = x_1^2 + x_2^2 + \cdots + x_n^2$.

In the next two sections we explain how to compute these solutions in practice.

2. Normal equation and the least squares solution

Recall the definitions of the following linear subspaces related to the matrix A:

$$\begin{aligned}
\mathrm{Ker}(A) &= \{\mathbf{x} \in \mathbb{R}^n \,:\, A\mathbf{x} = 0\},\\
\mathrm{Range}(A) &= \{\mathbf{b} \in \mathbb{R}^k \,:\, \text{there exists } \mathbf{x} \in \mathbb{R}^n \text{ such that } A\mathbf{x} = \mathbf{b}\},\\
\mathrm{Coker}(A) &= (\mathrm{Range}(A))^\perp \subset \mathbb{R}^k.
\end{aligned}$$

Consider the quadratic functional $Q : \mathbb{R}^n \to \mathbb{R}$ defined by

$$Q(\mathbf{x}) = \|A\mathbf{x} - \mathbf{b}\|^2.$$

We want to find a minimizer $\widetilde{\mathbf{x}} \in \mathbb{R}^n$ for Q. In other words, we look for a vector $\widetilde{\mathbf{x}}$ for which it holds that

$$Q(\widetilde{\mathbf{x}}) = \min_{\mathbf{x}\in\mathbb{R}^n} Q(\mathbf{x}). \tag{2.1}$$

Note that Q is continuously differentiable in any variable x_j. Therefore, since $\widetilde{\mathbf{x}}$ is a minimizer, we have

$$0 = \frac{d}{dt}\|A(\widetilde{\mathbf{x}} + t\mathbf{w}) - \mathbf{b}\|^2\bigg|_{t=0}$$

for any $\mathbf{w} \in \mathbb{R}^n$. (Why?)

We use the notation $\langle \mathbf{x}, \mathbf{y}\rangle$ for the inner product between two vertical vectors $\widetilde{\mathbf{x}} \in \mathbb{R}^n$ and $\widetilde{\mathbf{y}} \in \mathbb{R}^n$. The definition is

$$\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x}^T\mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

Note that

$$\langle \mathbf{x}, \mathbf{x}\rangle = \mathbf{x}^T\mathbf{x} = x_1^2 + x_2^2 + \cdots + x_n^2 = \|\mathbf{x}\|^2.$$

Also, use matrix algebra to see that

$$\langle A\mathbf{x}, \mathbf{b}\rangle = (A\mathbf{x})^T\mathbf{b} = (\mathbf{x}^T A^T)\mathbf{b} = \mathbf{x}^T(A^T\mathbf{b}) = \langle \mathbf{x}, A^T\mathbf{b}\rangle.$$

Now use the linearity of the inner product to compute

$$\begin{aligned}
0 =& \frac{d}{dt}\|A(\widetilde{\mathbf{x}} + t\mathbf{w}) - \mathbf{b}\|^2\Big|_{t=0} \\
=& \frac{d}{dt}\langle A\widetilde{\mathbf{x}} + tA\mathbf{w} - \mathbf{b}, A\widetilde{\mathbf{x}} + tA\mathbf{w} - \mathbf{b}\rangle\Big|_{t=0} \\
=& \frac{d}{dt}\Big\{\|A\widetilde{\mathbf{x}}\|^2 + 2t\langle A\widetilde{\mathbf{x}}, A\mathbf{w}\rangle + t^2\|A\mathbf{w}\|^2 \\
& \qquad - 2t\langle \mathbf{b}, A\mathbf{w}\rangle - 2\langle A\widetilde{\mathbf{x}}, \mathbf{b}\rangle + \|\mathbf{b}\|^2\Big\}\Big|_{t=0} \\
=& \Big\{2\langle A\widetilde{\mathbf{x}}, A\mathbf{w}\rangle + 2t\|A\mathbf{w}\|^2 - 2\langle \mathbf{b}, A\mathbf{w}\rangle\Big\}\Big|_{t=0} \\
=& 2\langle A\widetilde{\mathbf{x}}, A\mathbf{w}\rangle - 2\langle \mathbf{b}, A\mathbf{w}\rangle \\
=& 2\langle A^T A\widetilde{\mathbf{x}}, \mathbf{w}\rangle - 2\langle A^T\mathbf{b}, \mathbf{w}\rangle.
\end{aligned}$$

We conclude that the identity $\langle A^T A\widetilde{\mathbf{x}}, \mathbf{w}\rangle = \langle A^T\mathbf{b}, \mathbf{w}\rangle$ holds for any nonzero $\mathbf{w} \in \mathbb{R}^n$. Therefore, the minimizing vector must satisfy

$$A^T A\widetilde{\mathbf{x}} = A^T\mathbf{b}. \tag{2.2}$$

The identity (2.2) is called the *normal equation*.

Recall that if $V \subset \mathbb{R}^n$ is a linear subspace (i.e. a vector subspace), then $V^\perp$ is defined using the rule: $\mathbf{x} \in V^\perp$ if and only if $\langle \mathbf{x}, \mathbf{w}\rangle$ for every $\mathbf{w} \in V$. Then also $V^\perp$ is a linear subspace and every $\mathbf{x} \in \mathbb{R}^n$ has the unique decomposition as $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_\perp$ where $\mathbf{x}_0 \in V$ and $\mathbf{x}_\perp \in V^\perp$. One then says that $\mathbb{R}^n = V \oplus V^\perp$ is an *orthogonal sum* of V and $V^\perp$.

Lemma 2.1. *Let $A : \mathbb{R}^n \to \mathbb{R}^k$ be a linear map. Then Ker(A^T) = Range$(A)^\perp$.*

Proof. This is a simple observation using only definitions

$$\mathbf{y} \in \text{Range}(A)^\perp \Leftrightarrow \langle \mathbf{y}, \mathbf{x}\rangle = 0 \quad \forall \mathbf{x} \in \text{Range}(A)$$

and

$$\mathbf{y} \in \text{Ker}(A^T) \Leftrightarrow A^T\mathbf{y} = 0,$$

and the formula $\langle \mathbf{y}, A\mathbf{w}\rangle = \left\langle A^T\mathbf{y}, \mathbf{w}\right\rangle$. Details are left as an exercise. □

Theorem 2.1. *Let $A : \mathbb{R}^n \to \mathbb{R}^k$ be a linear map. Then there exists at least one solution to the normal equation (2.2) for any $\mathbf{b} \in \mathbb{R}^k$.*

Proof. It is sufficient to show that $A^T\mathbf{b} \in \text{Range}(A^TA)$ for every $\mathbf{b} \in \mathbb{R}^k$, i.e. $\text{Range}(A^T) \subset \text{Range}(A^TA)$. (Notice that it is a trivial observation that $\text{Range}(A^TA) \subset \text{Range}(A^T)$ since $\text{Range}(A) \subset \mathbb{R}^k$.)

Let us decompose $\mathbf{x} \in \mathbb{R}^k$ into $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_\perp$ where $\mathbf{x}_0 \in \text{Range}(A)$ and $\mathbf{x}_\perp \in \text{Range}(A)^\perp$. Now

$$A^T\mathbf{x} = A^T(\mathbf{x}_0 + \mathbf{x}_\perp) = A^T\mathbf{x}_0 + A^T\mathbf{x}_\perp = A^T\mathbf{x}_0$$

since $\mathbf{x}_\perp \in \text{Ker}(A^T)$ by Lemma 2.1. Hence $A^T(\mathbb{R}^k) \subset A^T(\text{Range}(A))$, i.e. $\text{Range}(A^T) \subset \text{Range}(A^TA)$.

This shows that the normal equation (2.2) $A^TA\mathbf{x} = A^T\mathbf{b}$ has always a solution. □

Theorem 2.2. *Let $\widetilde{\mathbf{x}} \in \mathbb{R}^n$. Then $\widetilde{\mathbf{x}}$ is a least squares solution of $A\mathbf{x} = \mathbf{b}$ if and only if $\widetilde{\mathbf{x}}$ solves the normal equation (2.2).*

Proof. We already proved the direction "⇒" using calculus. Let us prove the other direction now. This proof is based on the orthogonal decomposition

$$\mathbb{R}^k = \text{Range}(A) \oplus \text{Range}(A)^\perp$$

and Lemma 2.1.

If $\widetilde{\mathbf{x}} \in \mathbb{R}^n$ solves the normal equation (2.2), then $A^T(A\widetilde{\mathbf{x}} - \mathbf{b}) = 0$. Therefore

$$A\widetilde{\mathbf{x}} - \mathbf{b} \in \text{Ker}(A^T) = \text{Range}(A)^\perp$$

by Lemma 2.1. This means that $\langle A\widetilde{\mathbf{x}} - \mathbf{b}, \mathbf{y}\rangle = 0$ for every $\mathbf{y} \in \text{Range}(A)$.

Let us decompose $\mathbf{b} = \mathbf{b}_0 + \mathbf{b}_\perp$ where $\mathbf{b}_0 \in \text{Range}(A)$ and $\mathbf{b}_\perp \in \text{Range}(A)^\perp$. We can conclude that

$$\langle A\widetilde{\mathbf{x}} - \mathbf{b}, A\widetilde{\mathbf{x}} - \mathbf{b}_0\rangle = 0$$

since $A\widetilde{\mathbf{x}} - \mathbf{b}_0 \in \text{Range}(A)$, and

$$\langle A\widetilde{\mathbf{x}} - (\mathbf{b}_0 + \mathbf{b}_\perp), -\mathbf{b}_\perp\rangle = \|\mathbf{b}_\perp\|^2$$

since $\mathbf{b}_\perp \in \text{Range}(A)^\perp$. Adding up these two formulas gives directly that

$$\|A\widetilde{\mathbf{x}} - \mathbf{b}\|^2 = \|\mathbf{b}_\perp\|^2 .$$

The proof is complete if we show that $\|A\mathbf{x} - \mathbf{b}\|^2 \geq \|\mathbf{b}_\perp\|^2$ for every $\mathbf{x} \in \mathbb{R}^n$. Using that $A\mathbf{x} + \mathbf{b}_0 \in \text{Range}(A)$ and $\mathbf{b}_\perp \in \text{Range}(A)^\perp$, we see that $A\mathbf{x} + \mathbf{b}_0$ and $\mathbf{b}_\perp$ are orthogonal to each others. Therefore

$$\|A\mathbf{x} - \mathbf{b}\|^2 = \|A\mathbf{x} - \mathbf{b}_0\|^2 + \|\mathbf{b}_\perp\|^2 \geq \|\mathbf{b}_\perp\|^2$$

as desired. □

The above proof may seem a bit abstract at first sight, it however has a geometric nature. A solution of the normal equation approximates perfectly the part of $\mathbf{b}$ in Range(A), and the total error is equal to the part of $\mathbf{b}$ in Range$(A)^{\perp}$. Moreover, we showed in the end of the proof that this is the best approximation of $\mathbf{b}$ that can be obtained in the range of A, i.e. a least squares solution. One can also notice that the formula (4.2) gives a concrete equivalent way to look at least squares solutions.

Corollary 2.1. *If the $n \times n$ matrix $A^T A$ is invertible, the least squares solution of $A\mathbf{x} = \mathbf{b}$ is given by*

$$\widetilde{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}. \tag{2.3}$$

If $A^T A$ is not invertible, there is no unique minimizer for Q and we cannot use formula (2.3). But even in that case we can compute the minimum norm solution!

3. Fitting a linear model to noisy data

Consider the following linear model describing the relationship between two scalar quantities $x \in \mathbb{R}$ and $x \in \mathbb{R}$:

$$y = a_0 x + b_0, \tag{3.1}$$

where $a_0, b_0 \in \mathbb{R}$ are parameters.

Assume given noisy data $y'_1, y'_2, \ldots, y'_n$ at points $x_1, x_2, \ldots, x_n$. More precisely,

$$y'_j = a x_j + b + \varepsilon_j, \tag{3.2}$$

where ε_j is some unknown error in the measurement.

We can solve for the parameters $a, b \in \mathbb{R}$ that give the model of the form (3.1) that best fits the data in the least-squares sense. Namely, write

$$A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \in \mathbb{R}^n, \qquad \mathbf{y}' = \begin{bmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_n \end{bmatrix} \in \mathbb{R}^n,$$

and consider the linear system of equations defined by

$$A \begin{bmatrix} a \\ b \end{bmatrix} = \mathbf{y}'. \tag{3.3}$$

Now in general the equation (3.3) has no solutions because of the errors in (3.2). But if the matrix $(A^T A)$ is invertible, then we can use (2.3) to compute the least-squares solution as

$$\begin{bmatrix} \widetilde{a} \\ \widetilde{b} \end{bmatrix} = (A^T A)^{-1} A^T \mathbf{y}'.$$

4. Computing the minimum norm solution

We need a method for computing minimum norm solutions. For this, write A in the form of its SVD $A = UDV^T$ as explained in Section A. Recall that the singular values are ordered from largest to smallest as shown in (A.4), and let r be the largest index for which the corresponding singular value is nonzero:

$$r = \max\{j \mid 1 \le j \le \min(k,n),\ d_j > 0\}. \tag{4.1}$$

The definition of index r is essential in the following analysis, so we will be extra-specific:

$$d_1 > 0, \quad d_2 > 0, \quad \cdots \quad d_r > 0, \quad d_{r+1} = 0, \quad \cdots \quad d_{\min(k,n)} = 0.$$

Of course, it is also possible that all singular values are zero, in which case r is not defined and A is the zero matrix, or none of the singular values may be zero.

The next result gives a method to determine the minimum norm solution.

Theorem 4.1. *Let A be a $k \times n$ matrix and denote by $A = UDV^T$ the singular value decomposition of A. The minimum norm solution of the equation $A\mathbf{x} = \mathbf{b}$ is given by $A^+\mathbf{b}$ where*

$$A^+\mathbf{b} = VD^+U^T\mathbf{b},$$

and where

$$D^+ = \begin{bmatrix} 1/d_1 & 0 & \cdots & 0 & & \cdots & 0 \\ 0 & 1/d_2 & & & & & \vdots \\ \vdots & & \ddots & & & & \\ & & & 1/d_r & & & \\ & & & & 0 & & \\ \vdots & & & & & \ddots & \vdots \\ 0 & \cdots & & & & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{n \times k}.$$

Proof. Write the singular matrix V in the form $V = [V_1 \quad V_2 \quad \cdots \quad V_n]$ and note that the column vectors $V_1, \dots, V_n$ form an orthogonal basis for $\mathbb{R}^n$. We write $\mathbf{x} \in \mathbb{R}^n$ as a linear combination $\mathbf{x} = \sum_{j=1}^n a_j V_j = V\mathbf{a}$, and our goal is to find such coefficients $a_1, \dots, a_n$ that $\mathbf{x}$ becomes a minimum norm solution.

Set $\mathbf{b}' = U^T\mathbf{b} \in \mathbb{R}^k$ and compute

$$\begin{aligned} \|A\mathbf{x} - \mathbf{b}\|^2 &= \|UDV^TV\mathbf{a} - U\mathbf{b}'\|^2 \\ &= \|D\mathbf{a} - \mathbf{b}'\|^2 \\ &= \sum_{j=1}^{r}(d_j a_j - \mathbf{b}'_j)^2 + \sum_{j=r+1}^{k}(\mathbf{b}'_j)^2, \end{aligned} \tag{4.2}$$

where we used the orthogonality of U (namely, $\|U\mathbf{x}\| = \|\mathbf{x}\|$ for any vector $\mathbf{x} \in \mathbb{R}^k$). Now since d_j and $\mathbf{b}'_j$ are given and fixed, the expression (4.2) attains its minimum when $a_j = \mathbf{b}'_j/d_j$ for $j = 1, \ldots, r$. So any $\mathbf{x}$ of the form

$$\mathbf{x} = V \begin{bmatrix} d_1^{-1}\mathbf{b}'_1 \\ \vdots \\ d_r^{-1}\mathbf{b}'_r \\ a_{r+1} \\ \vdots \\ a_n \end{bmatrix}$$

is a least-squares solution. The smallest norm $\|\mathbf{x}\|$ is clearly given by the choice $a_j = 0$ for $r < j \le n$, so the minimum norm solution is uniquely determined by the formula $\mathbf{a} = D^+\mathbf{b}'$. □

Definition 4.1. *The matrix A^+ is called the* pseudoinverse*, or the* Moore-Penrose inverse *of A.*

Appendix A. The singular value decomposition

We know from matrix algebra that any matrix $A \in \mathbb{R}^{k\times n}$ can be written in the form

$$A = UDV^T, \tag{A.1}$$

where $U \in \mathbb{R}^{k\times k}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal matrices, that is,

$$U^TU = UU^T = I, \quad V^TV = VV^T = I,$$

and $D \in \mathbb{R}^{k\times n}$ is a diagonal matrix. The right side of (A.1) is called the singular value decomposition (SVD) of matrix A, and the diagonal elements d_j are the *singular values* of A. The properties of d_j, and the columns u_i of U, and the columns V_i of V correspond to those of the SVE.

In the case $k = n$ the matrix D is square-shaped: $D = \mathrm{diag}(d_1, \ldots, d_k)$. If $k > n$ then

$$D = \begin{bmatrix} \mathrm{diag}(d_1, \ldots, d_n) \\ \mathbf{0}_{(k-n)\times n} \end{bmatrix} = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & \cdots & d_n \\ 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}, \tag{A.2}$$

and in the case $k < n$ the matrix D takes the form

$$\begin{aligned} D &= [\mathrm{diag}(d_1, \ldots, d_k), \mathbf{0}_{k\times(n-k)}] \\ &= \begin{bmatrix} d_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & d_2 & & \vdots & \vdots & & \vdots \\ \vdots & & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & d_k & 0 & \cdots & 0 \end{bmatrix}. \end{aligned} \tag{A.3}$$

The diagonal elements d_j are nonnegative and in decreasing order:

$$d_1 \geq d_2 \geq \ldots \geq d_{\min(k,n)} \geq 0. \tag{A.4}$$

Note that some or all of the d_j can be equal to zero.