Numerical analysis: Least Squares Approximation

Dr. Dmitry Batenkov

April 24, 2021

1 Norms induced by inner products

Let V be a vector space over a field $\mathbb{F} \subseteq \mathbb{C}$, $u, v \in V$, and let $\langle \cdot, \cdot \rangle = (\cdot, \cdot)$ **be an inner product** (see Lecture 6).

Definition 1. A **norm** induced by the inner product $\langle \cdot, \cdot \rangle$ is defined as

$$\|u\| = \sqrt{\langle u, u \rangle}.$$

Claim 1 (Cauchy-Schwarz inequality)*.* For any inner product

$$|\langle u, v \rangle| \leq \|u\| \cdot \|v\|.$$

Proof. We do not present a proof here, but the reader can consult any standard text. For example, a beautiful book "Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities". □

Claim 2*.* The induced norm $\|\cdot\|$ as defined above indeed satisfies the axioms of the norm, i.e.

1. $\forall v \in V$ we have $\|v\| \geq 0$, and furthermore $\|v\| = 0$ iff $v = 0$.
2. $\|\alpha v\| = \sqrt{(\alpha v, \alpha v)} = \sqrt{|\alpha|^2 \|v\|^2} = |\alpha| \|v\| \ \forall \alpha \in \mathbb{F}, \forall v \in V$
3. $\|u + v\| \leq \|u\| + \|v\|$ (triangle inequality).

Proof.

□

Claim 3 (Generalized Pythagoras theorem). If $\langle u, v\rangle = 0$ then $\|u+v\|^2 = \|u\|^2 + \|v\|^2$.

Proof.

$\square$

2 Least squares approximation

Given: vectors (or functions) $\{\phi_0, \phi_1, \ldots, \phi_n\}$.

Definition 2. The vectors $\{\phi_0, \phi_1, \ldots, \phi_n\}$ are *linearly independent* if $c_0\phi_0 + c_1\phi_1 + \ldots + c_n\phi_n = 0$ implies $c_0 = \cdots = c_n = 0$.

Definition 3. The functions $\{\phi_0(x), \phi_1(x), \ldots, \phi_n(x)\}$ defined on an interval I (open, closed, semi-infinite or infinite) are *linearly independent* if $c_0\phi_0(x) + c_1\phi_1(x) + \ldots + c_n\phi_n(x) = 0$ for all $x \in I$ implies $c_0 = \cdots = c_n = 0$.

Claim 4. $\langle g, \phi_i\rangle = 0$ for all $i = 0, 1, \ldots, n$ if and only if $\langle g, \sum_{i=0}^{n} c_i\phi_i\rangle = 0$ for all choices of $c_0, \ldots, c_n$.

Problem (The least squares approximation problem)**.** Let V be an inner product space over a field $\mathbb{F} \subseteq \mathbb{C}$, with inner product $\langle\cdot,\cdot\rangle$, with $\{\phi_0, \phi_1, \ldots, \phi_n\} \subset V$ a linearly independent set of vectors (functions). The **least squares approximation** of some $f \in V$ is the best approximation to f from the subspace $W = \text{span}\{\phi_0, \phi_1, \ldots, \phi_n\}$ with respect to the induced norm, i.e.

$$f^{LS} = \arg\min_{w\in W} \|f - w\|. \tag{1}$$

Theorem 1 (The least squares approximation)**.** *If the coefficients $c_0^*, \ldots, c_n^*$ (here $*$ does **not** mean complex conjugate) satisfy*

$$\sum_{k=0}^{n} c_k^* \langle\phi_k, \phi_i\rangle = \langle f, \phi_i\rangle \qquad i = 0, 1, \ldots, n \tag{2}$$

then the vector $f^* = \sum_{i=0}^{n} c_i^*\phi_i$ is the least squares approximation to$f$, i.e. $f^* = f^{LS}$ where f^{LS} is given by (1).

Definition 4. The (system of) linear equations (2) are called **the normal equations,** since it can be written in the form

$$\langle f^* - f, \phi_i \rangle = 0 \qquad i = 0, 1, \ldots, n, \tag{3}$$

and so by Claim 4 we have that the error (or the *residual*) $f^* - f$ is orthogonal to all vectors in the subspace W.

Definition 5. The matrix of inner products $G = \begin{pmatrix} (\phi_0, \phi_0) & (\phi_1, \phi_0) & \ldots & (\phi_n, \phi_0) \\ (\phi_0, \phi_1) & (\phi_1, \phi_1) & \ldots & (\phi_n, \phi_1) \\ \ldots & \ldots & \ldots & \ldots \\ (\phi_0, \phi_n) & (\phi_1, \phi_n) & \ldots & (\phi_n, \phi_n) \end{pmatrix}$ is called the **Gram matrix** associated with the set $\{\phi_0, \phi_1, \ldots, \phi_n\}$.

Claim 5. The matrix G is non-singular, therefore the system of normal equations is uniquely solvable.

Proof. Suppose $G\vec{c} = 0$ for some vector of coefficients $\vec{c} = (c_0, \ldots, c_n)$. Consider $g = \sum_{i=0}^{n} c_i \phi_i$. From $G\vec{c} = 0$ it follows that $\langle g, \phi_i \rangle = 0$ for all $i = 0, \ldots, n$, and by Claim 4 g is orthogonal to all vectors in W, in particular to itself: $\langle g, g \rangle = 0$ and therefore $g = 0$.

Therefore the homogeneous system has a unique solution, proving the claim. ☐

Proof of theorem 1. Let $w = \sum_{i=0}^{n} c_i \phi_i \in W$ any vector, then

$$w - f = (w - f^*) + (f^* - f) = \underbrace{\sum_{i=0}^{n} (c_i - c_i^*)\, \phi_i}_{:=g} + (f^* - f)\,.$$

By (3) $\langle f^* - f, g \rangle = 0$ and therfore by the Generalized Pythagoras Theorem

$$\|w - f\|^2 = \|g + (f^* - f)\,\|^2 = \|g\|^2 + \|\,(f^* - f)\,\|^2 \geq \|\,(f^* - f)\,\|^2,$$

with equality if and only if $c_i = c_i^*$ for all $i = 0, 1, \ldots, n$. Therefore by definition $f^* = f^{LS}$ is the (unique) least squares approximation to f. ☐

Remark 1. If the system $\{\phi_0, \phi_1, \ldots, \phi_n\}$ is orthogonal, i.e. $\langle \phi_i, \phi_k \rangle = 0$ if $i \neq k$, then the normal equations reduce to

$$c_k^* \langle \phi_k, \phi_k \rangle = \langle f, \phi_k \rangle \Longrightarrow c_k^* = \frac{\langle f, \phi_k \rangle}{\langle \phi_k, \phi_k \rangle} \qquad k = 0, 1, \ldots, n.$$

Example 1. $V = C[-1,1]$ over $\mathbb{F} = \mathbb{R}$ with inner product $\langle f, g \rangle = \int_{-1}^{1} f(x) g(x)\, dx$. Let $\phi_k = P_k(x)$, the $k - th$ Legendre polynomial, for $k = 0, 1, \ldots, n$. Let us find the least squares approximation to some $f \in V$. Since the system is orthogonal, the solution is given by $f^* = \sum_{k=0}^{n} c_k^* P_k(x)$ where

$$c_k^* = \frac{\int_{-1}^{1} f(x) P_k(x)\, dx}{\int_{-1}^{1} P_k^2(x)\, dx} \qquad k = 0, 1, \ldots, n.$$

Example 2. Let f be given by its samples at the grid points $x_1 = 1, x_2 = 3, x_3 = 4, x_4 = 6, x_5 = 7$ as follows:

x_i	1	3	4	6	7
$f(x_i)$	-2.1	-0.9	-0.6	0.6	0.9

Find the least squares approximation to f of the form $f^* = c_0^* + c_1^* x$.

Solution. We identify a function $f(x)$ with the vector $\vec{f}$ of its samples at $x_1, \ldots, x_5$. The inner product space is isomorphic to $\mathbb{R}^5$, and we are looking for an approximation from the two-dimensional subspace spanned by $\phi_0(x) = 1 = (1,1,1,1,1)$ and $\phi_1(x) = x = (1,3,4,6,7)$. In other words, $V = \mathbb{R}^5$ and $\left\langle \vec{f}, \vec{g} \right\rangle = \vec{g}^T \vec{f} = \sum_{k=1}^{5} f(x_k) g(x_k)$. Computing the inner products, the normal equations take the form

$$\begin{bmatrix} 5 & 21 \\ 21 & 111 \end{bmatrix} \begin{bmatrix} c_0 * \\ c_1 * \end{bmatrix} = \begin{bmatrix} -2.1 \\ 2.7 \end{bmatrix}$$

whose solution is $c_0^* = -2.542, c_1^* = 0.5033$, i.e.

$$f^*(x) = -2.542 + 0.5033x.$$