



planetmath.org

Math for the people, by the people.

best approximation in inner product spaces

Canonical name	BestApproximationInInnerProductSpaces
Date of creation	2013-03-22 17:32:16
Last modified on	2013-03-22 17:32:16
Owner	asteroid (17536)
Last modified by	asteroid (17536)
Numerical id	12
Author	asteroid (17536)
Entry type	Feature
Classification	msc 41A65
Classification	msc 46C05
Classification	msc 46N10
Classification	msc 49J27
Classification	msc 41A52
Classification	msc 41A50
Defines	approximation by polynomials
Defines	best fitting lines

The study of best approximations in inner product spaces has a very elegant treatment with profound consequences. Most of the theory of Hilbert spaces depends on this study and several approximation problems are better understood using this techniques and results.

For example: least square fitting, linear regression, approximation of functions by polynomials, among many other problems, can be seen as particular cases of the general study of best approximation in inner product spaces.

Some of the above problems are going to be discussed later in this entry.

## 1 Existence and Uniqueness

Our fundamental result on the existence and uniqueness of best approximations is the following (we postpone its proof to this attached <http://planetmath.org/ProofOfExistence>

**Theorem** - Let  $X$  be an inner product space and  $A \subseteq X$  a <http://planetmath.org/CompleteMetricSpace> convex and non-empty subset. Then for every  $x \in X$  there exists a unique <http://planetmath.org/BestApproximation> best approximation of  $x$  in  $A$ , i.e. there exists a unique element  $a_0 \in A$  such that

$$\|x - a_0\| = d(x, A) = \inf_{a \in A} \|x - a\|.$$

## 2 Geometric Interpretation

The following result gives a very geometric interpretation of the best approximation when  $A$  is a subspace of  $X$ . We also postpone its proof to an attached entry.

**Theorem** - Let  $X$  be an inner product space,  $A \subseteq X$  a subspace and  $x \in X$ . The following statements are equivalent:

- $a_0 \in A$  is the best approximation of  $x$  in  $A$ .
- $a_0 \in A$  and  $x - a_0 \perp A$ .

Thus, the best approximation of  $x$  in a subspace  $A$  is just the orthogonal projection of  $x$  in  $A$ .

### 3 Calculation of Best Approximations

When the  $A$  is a complete subspace of  $X$ , the best approximation can be "calculated" explicitly. Recall that, in this case,  $A$  becomes an Hilbert space (since it is complete) and therefore it has an orthonormal basis.

Again, we postpone the proof of the next result to an attached entry.

**Theorem -** Let  $X$  be an inner product space and  $A \subseteq X$  a complete subspace. Let  $(e_i)_{i \in J}$  be an orthonormal basis for  $A$ . Then for every  $x \in X$  the best approximation  $a_0 \in A$  of  $x$  in  $A$  is given by

$$a_0 = \sum_{i \in J} \langle x, e_i \rangle e_i .$$

One can also write the best approximation in of any other basis (not necessarily an orthonormal one). For simplicity we present here how that can be done when  $A$  is a finite dimensional subspace of  $X$ .

**Theorem -** Let  $X$  be an inner product space and  $A \subseteq X$  a finite dimensional subspace. Let  $v_1, \dots, v_n$  be a basis for  $A$ . Then for every  $x \in X$  the best approximation  $a_0 \in A$  of  $x$  in  $A$  is given by

$$a_0 = \sum_{i=1}^n a_0^i v_i$$

where the coefficients  $a_0^i$  are the solutions of the system of equations

$$\begin{pmatrix} \langle v_1, v_1 \rangle & \cdots & \langle v_1, v_n \rangle \\ \vdots & \ddots & \vdots \\ \langle v_n, v_1 \rangle & \cdots & \langle v_n, v_n \rangle \end{pmatrix} \begin{pmatrix} a_0^1 \\ \vdots \\ a_0^n \end{pmatrix} = \begin{pmatrix} \langle x, v_1 \rangle \\ \vdots \\ \langle x, v_n \rangle \end{pmatrix} .$$

*Remark—* The above matrix is a symmetric <http://planetmath.org/PositiveDefinitepositi> definite matrix, which implies that the system has a unique solution as expected.

### 4 Applications

There are several applications of the above results. We explore two of them in the following.

#### 4.0.1 - Approximation of functions by polynomials :

Suppose we want to find a polynomial of degree  $\leq n$  that approximates in the best possible way a given function  $f$ . We are in fact trying to find a point in the subspace of polynomials of degree  $\leq n$  that is closest to  $f$ , i.e. we are trying to find the best approximation of  $f$  in that subspace.

For example, let  $f \in L^2([0, 1])$ . Consider the basis  $v_k(t) = t^k$ ,  $0 \leq k \leq n$ , of the subspace of polynomials of degree  $\leq n$ .

The best approximation of  $f$  by these polynomials is the function  $a_0(t) = a_0^1 + a_0^1 t + \dots + a_0^n t^n$ , where the coefficients  $a_0^1, \dots, a_0^n$  are the solutions of the system

$$\begin{pmatrix} 1 & \dots & \frac{1}{n+1} \\ \vdots & \ddots & \vdots \\ \frac{1}{n+1} & \dots & \frac{1}{2n+1} \end{pmatrix} \begin{pmatrix} a_0^1 \\ \vdots \\ a_0^n \end{pmatrix} = \begin{pmatrix} \int_0^1 f(t) dt \\ \vdots \\ \int_0^1 t^n f(t) dt \end{pmatrix}.$$

*Remark*— Instead of polynomials we could approximate  $f$  by any other of functions using the same procedure.

#### 4.0.2 - Best Fitting Lines :

Suppose we want to find the line that best fits some given points  $(t_1, y_1), \dots, (t_n, y_n)$ , i.e. the affine function  $a_0(t) = \alpha t + \beta$  that minimizes  $\sum_{k=1}^n |a_0(t_k) - y_k|^2$ .

We are then led to consider the inner product

$$\langle f, g \rangle = \sum_{k=1}^n f(t_k)g(t_k)$$

in the space of functions  $h : \{t_1, \dots, t_n\} \rightarrow \mathbb{R}$ .

With this setting we are then looking for the best approximation of the function  $f(t_k) = y_k$  on the subspace of affine functions.

A base for the subspace of affine functions is given by the functions  $v_1(t) = 1$  and  $v_2(t) = t$ .

The best approximation of  $f$  on this space is the function  $a_0(t) = \beta + \alpha t$ , where the coefficients  $\beta, \alpha$  are the solutions of the system

$$\begin{pmatrix} n & \sum_{k=1}^n t_k \\ \sum_{k=1}^n t_k & \sum_{k=1}^n t_k^2 \end{pmatrix} \begin{pmatrix} \beta \\ \alpha \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^n y_k \\ \sum_{k=1}^n y_k t_k \end{pmatrix}.$$

Thus, the function  $a_0(t) = \beta + \alpha t$  obtained by the above procedure provides the line that best fits the data  $(t_1, y_1), \dots, (t_n, y_n)$ .