

Math 210 Lecture Notes Bo Li, Fall 2013

Section 3.2 Best Approximations. Orthogonality

Examples Consider $X = L^2(-1, 1)$, $f(x) = e^x + |x|$.

$P_2 = \{\text{all (real) polynomials of degree } \leq 2\}$

e.g. $p(x) = 2x^2 - x + 3$ is a "vector" in P_2 .

$g(x) = -x + 1$ is a "vector" in P_2

Note that $P_2 \subset X$. minimization

An optimization (or variational) problem:

Find $p \in P_2$ such that

$$\|p - f\| = \min_{g \in P_2} \|g - f\|.$$

Consider $M = \{p(x) = ax^2 + bx + c \in P_2 : \text{all } a, b, c \geq 0\}$.

Another minimization problem.

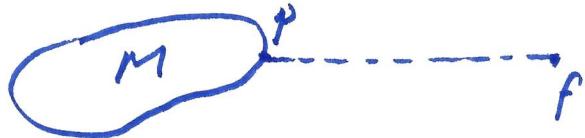
Find $p \in M$ such that

$$\|p - f\| = \min_{g \in M} \|g - f\|.$$

In general, let X be a Hilbert space. [So, depending what X is, a vector of X can be a 3-component vector, or a matrix, or a function, or other quantity.] Let M be a nonempty subset of X : $\phi \neq M \subset X$. Let $f \in X$ (but in general $f \notin M$).

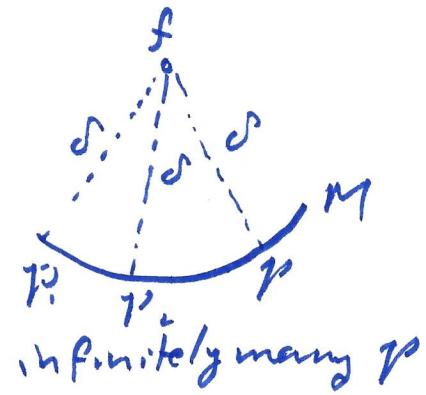
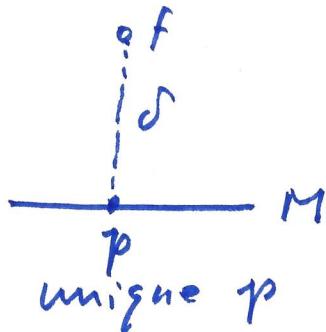
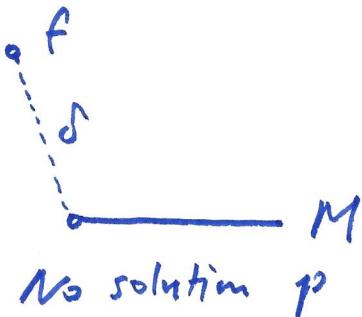
Best Approximation Problem: Find $p \in M$ such that

$$\|p - f\| = \min_{g \in M} \|g - f\|.$$



Existence? Uniqueness? Solution method?

$$\text{Let } \delta = \inf_{g \in M} \|g - f\|$$



Definition A subset Y of a Hilbert space is closed, if $y_n \in Y$ ($n=1, 2, \dots$) and $y_n \rightarrow y$ for some vector $y \in X$ then $y \in Y$. [That is Y itself is "complete;" it contains all its limits.]

Here is one of the commonly used version of theorem of best approximations in a Hilbert space.

Theorem Let X be a Hilbert space. Let M be a closed subspace of X . Let $f \in X$. Then there exists ~~a unique~~ a unique $p \in M$ such that

$$\|p - f\| = \min_{g \in M} \|g - f\|.$$

Moreover, p is characterized by

$$p - f \perp M$$

i.e., $\langle p - f, g \rangle = 0$ for any $g \in M$.

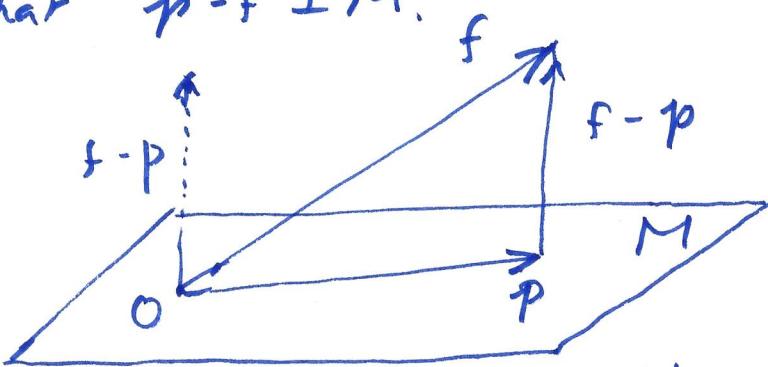
$$\text{and } \|p - f\|^2 = \|f\|^2 - \|p\|^2.$$

Definition p is the best approximation of f in M .

Note: "M is a subspace of X" means that M is closed with respect to the addition and scalar multiplication of X.

$$x, y \in M, \lambda \in F \Rightarrow x+y \in M, \lambda x \in M.$$

Note: The unique minimizer $p \in M$ satisfies that $p-f \perp M$.



Step 1 we first prove the existence.

Pf of Thm Let $\delta = \inf_{q \in M} \|q-f\|$. Let $\varepsilon_k \in M$ ($k=1, 2, \dots$) be such that $\|\varepsilon_k - f\| \downarrow \delta$.

We show that $\{\varepsilon_k\}$ is a Cauchy sequence.

Facts: ① Law of parallelogram

$$\begin{aligned} & \|(\varepsilon_j - f) - (\varepsilon_k - f)\|^2 + \|(\varepsilon_j - f) + (\varepsilon_k - f)\|^2 \\ & + 2\|\varepsilon_j - f\|^2 + 2\|\varepsilon_k - f\|^2 \end{aligned}$$

$$\textcircled{2} \quad \left\| \frac{1}{2}(\varepsilon_j + \varepsilon_k) - f \right\| \geq \delta \quad \text{since } \frac{1}{2}(\varepsilon_j + \varepsilon_k) \in M.$$

Now,

$$\begin{aligned} \|\varepsilon_j - \varepsilon_k\|^2 &= \|(\varepsilon_j - f) - (\varepsilon_k - f)\|^2 \\ &= -\|(\varepsilon_j - f) + (\varepsilon_k - f)\|^2 + 2\|\varepsilon_j - f\|^2 + 2\|\varepsilon_k - f\|^2 \\ &= -4\left\| \frac{1}{2}(\varepsilon_j + \varepsilon_k) - f \right\|^2 + 2\|\varepsilon_j - f\|^2 + 2\|\varepsilon_k - f\|^2 \end{aligned}$$

$$\leq -4\delta^2 + 2\|g_j - f\|^2 + 2\|g_k - f\|^2$$

$$\rightarrow -4\delta^2 + 2\delta^2 + 2\delta^2 = 0 \quad \text{as } j, k \rightarrow \infty.$$

Hence, $\{g_k\}$ is a Cauchy sequence.

Since X is a Hilbert space, there exists $p \in X$ such that $\lim_{k \rightarrow \infty} g_k = p$.

Since all $g_k \in M$ ($k=1, 2, \dots$) and M is closed, $p \in M$.

Since $\lim_{k \rightarrow \infty} \|g_k - f\| = \delta$ (we chose $g_k \in M$ to satisfy this.)

$$\lim_{k \rightarrow \infty} \|g_k - p\| = 0 \quad (\text{i.e., } \lim_{k \rightarrow \infty} g_k = p)$$

we have $\delta \leq \|p - f\| \leq \|p - g_k\| + \|g_k - f\| \rightarrow 0 + \delta = \delta$.
↑
since $p \in M$

Thus $\|p - f\| = \delta = \inf_{g \in M} \|g - f\|, \quad \forall p \in M.$
 $= \min_{g \in M} \|g - f\|.$

Step 2

We now prove the uniqueness.

Suppose both p and \hat{p} are in M and

$$\|p - f\| = \delta = \inf_{g \in M} \|g - f\|.$$

$$\|\hat{p} - f\| = \delta = \inf_{g \in M} \|g - f\|.$$

By similar arguments, we have

$$0 \leq \|p - \hat{p}\|^2 = \|(p - f) - (\hat{p} - f)\|^2$$

$$= -\|(p - f) + (\hat{p} - f)\|^2 + 2\|p - f\|^2 + 2\|\hat{p} - f\|^2$$

$$= -4\|\frac{1}{2}(p + \hat{p}) - f\|^2 + 2\delta^2 + 2\delta^2$$

$$\leq -4\delta^2 + 4\delta^2 = 0. \quad \text{So, } \|p - \hat{p}\| = 0, \quad p = \hat{p}.$$

Step 3. We show that the minimizer $p \in M$ is characterized by $p \in M$, $p \perp f \perp M$.

[We may assume $f \notin M$. Otherwise $p = f$.]

If $f - p \notin M$ then there exists $g \in M$ s.t. $\langle f - p, g \rangle \neq 0$. Clearly $g \neq 0$. So, we may assume $\|g\| = 1$. [Otherwise, replace g by $g/\|g\|$.] Denote $z = f - p$. Note that $\|z\| = \delta$.

We have for $\alpha \in \mathbb{R}$

$$\begin{aligned} \|z - \alpha g\|^2 &= \langle z - \alpha g, z - \alpha g \rangle \\ &= \epsilon \|z\|^2 - \alpha \langle g, z \rangle - \bar{\alpha} \langle \overline{g}, z \rangle + |\alpha|^2 \quad (\text{since } \|g\|^2 = 1) \end{aligned}$$

Let $\alpha = \langle g, z \rangle \neq 0$. Then $\alpha g \in M$, $p + \alpha g \in M$.

$$\begin{aligned} \delta^2 &\leq \|f - (p + \alpha g)\|^2 = \|z - \alpha g\|^2 \\ &= \|z\|^2 - |\langle g, z \rangle|^2 - |\langle \overline{g}, z \rangle|^2 + |\alpha|^2 \\ &= \delta^2 - |\langle g, z \rangle|^2 < \delta^2 \end{aligned}$$

Impossible. So, $f - p \perp g \quad \forall g \in M$.

Conversely, suppose $p \in M$ and $f - p \perp M$. We prove that $\|f - p\| \leq \|f - g\|$ for any $g \in M$, $g \neq p$.

For any $g \in M$, $g \neq p$: $p - g \in M$. So, $f - p \perp p - g$.

$$\begin{aligned} \text{Thus, Pythagorean Thm: } \|f - g\|^2 &= \|(f - p) + (p - g)\|^2 \\ &= \|f - p\|^2 + \|p - g\|^2 \end{aligned}$$

Step 4 $f - p \perp p$

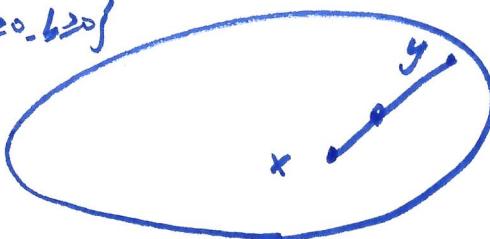
$$\begin{aligned} \text{So, } \|f\|^2 &= \|f - p + p\|^2 \\ &= \|f - p\|^2 + \|p\|^2 \end{aligned}$$

$> \|f - p\|^2$. Q.E.D.

Definition A subset $A(\neq \emptyset)$ of a vector space X is convex, if for any $x, y \in A$ and any $\lambda \in (0, 1)$

$$\lambda x + (1-\lambda)y \in A.$$

Example. $A = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : a \geq 0, b \geq 0 \right\}$
 $X = \mathbb{R}^2$



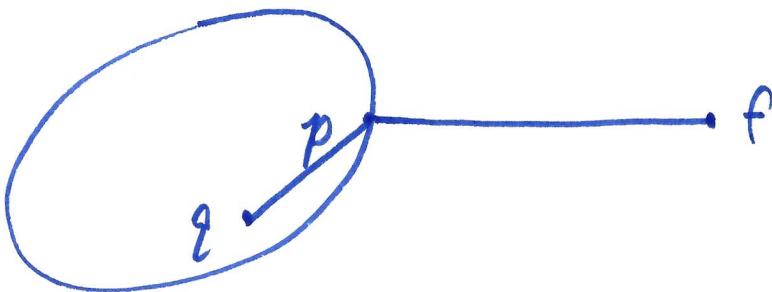
The previous theorem of best approximation can be generalized to convex subset M .

Theorem Let M be a nonempty convex subset of a Hilbert space X . Let $f \in X$. Then there exists a unique $p \in M$ such that

$$\|p - f\| = \min_{q \in M} \|q - f\|.$$

Moreover $p \in M$ is characterized by

$$\langle f - p, q - p \rangle \leq 0 \text{ for any } q \in M.$$



Definition Let X be a Hilbert space. Let M be a nonempty subspace of X . For any $f \in X$, denote by $P_M f \in M$ the best approximation of f .

So, ~~$P_M f$~~ is a mapping. $P_M f \in M$ and $f - P_M f \perp M$. Note: $P_M^2 = P_M$

The next question is how to compute (or find) the best approximation. In general, this is a hard problem. But if M is a finite-dimensional space and we know an orthonormal basis of M , then the calculation is relatively simple.

Theorem Let X be a Hilbert space and M a nonempty subspace of X with $\dim M = n < \infty$. Suppose $\{u_1, \dots, u_n\}$ is an orthonormal basis for M . Then for any $f \in X$, the best approximation of f in M is

$$P_M f = \langle f, u_1 \rangle u_1 + \dots + \langle f, u_n \rangle u_n.$$

Proof $P_M f \in M$ is defined by $\|P_M f - f\| = \min_{g \in M} \|g - f\|$, and is characterized by $P_M f - f \perp M$. Since $\{u_1, \dots, u_n\}$ is an orthonormal basis for M , we have

$$P_M f = c_1 u_1 + c_2 u_2 + \dots + c_n u_n$$

for some numbers c_1, c_2, \dots, c_n .

$$P_M f - f \perp M \implies P_M f - f \perp u_j \quad (j=1, 2, \dots, n)$$

$$\iff \langle P_M f - f, u_j \rangle = 0 \quad (j=1, 2, \dots, n)$$

$$\text{So, } \langle P_M f, u_j \rangle = \langle f, u_j \rangle \quad (j=1, 2, \dots, n)$$

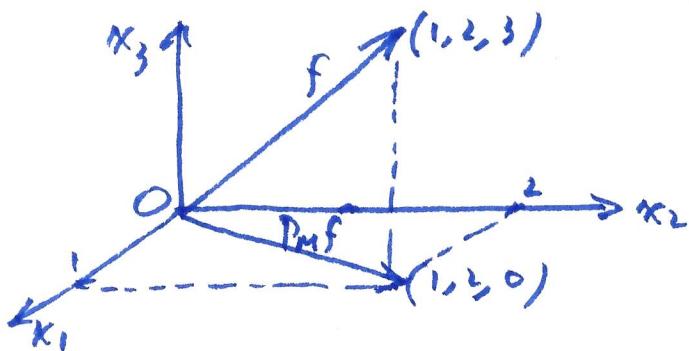
$$\begin{aligned} \text{For a fixed } j, \quad \langle f, u_j \rangle &= \langle P_M f, u_j \rangle = \langle c_1 u_1 + \dots + c_n u_n, u_j \rangle \\ &= c_1 \langle u_1, u_j \rangle + c_2 \langle u_2, u_j \rangle + \dots + c_n \langle u_n, u_j \rangle \end{aligned}$$

$$\begin{aligned} &= c_j \langle u_j, u_j \rangle \quad (\text{since } \langle u_k, u_j \rangle = 0 \text{ if } k \neq j) \\ &= c_j^2 \quad (\text{since } \langle u_j, u_j \rangle = \|u_j\|^2 = 1) \end{aligned}$$

Thus, $P_M f = \langle f, u_1 \rangle u_1 + \dots + \langle f, u_n \rangle u_n$. Q.E.D.

Example 1 $X = \mathbb{R}^3$, $f = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $M = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} : x_1, x_2 \in \mathbb{R} \right\}$.
 $\dim M = 2$. $u_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $u_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ $\{u_1, u_2\}$ is an orthonormal basis for M . (We use the standard inner product.). Let us find the best approximation of f in M .

$$\begin{aligned} P_M f &= \langle f, u_1 \rangle u_1 + \langle f, u_2 \rangle u_2 \\ &= 1 \cdot u_1 + 2 u_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}. \end{aligned}$$



Example 2 $X = \mathbb{R}^4$, $f = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$, $M = \text{span}\{v_1, v_2, v_3\}$
 $v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, $v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$. v_1, v_2, v_3 are linearly independent. They form a basis for M . The best approximation, $P_M f$, of f in M is a linear combination of v_1, v_2, v_3 :

$$P_M f = a_1 v_1 + a_2 v_2 + a_3 v_3.$$

How to find a_1, a_2, a_3 ? Use $P_M f - f \perp M$

$$\langle P_M f, v_j \rangle = \langle f, v_j \rangle \quad j = 1, 2, 3.$$

$$\langle P_M f, v_1 \rangle = \langle v_1, v_1 \rangle a_1 + \langle v_2, v_1 \rangle a_2 + \langle v_3, v_1 \rangle a_3 = \langle f, v_1 \rangle = 1$$

$$\langle P_M f, v_2 \rangle = \langle v_1, v_2 \rangle a_1 + \langle v_2, v_2 \rangle a_2 + \langle v_3, v_2 \rangle a_3 = \langle f, v_2 \rangle = 3$$

$$\langle P_M f, v_3 \rangle = \langle v_1, v_3 \rangle a_1 + \langle v_2, v_3 \rangle a_2 + \langle v_3, v_3 \rangle a_3 = \langle f, v_3 \rangle = 6$$

$$A = \left[\langle v_i, v_j \rangle \right]_{i,j=1}^3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} \quad \text{Solve } A \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix}$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 3 \\ 1 & 2 & 3 & 6 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 3 \end{array} \right]$$

$$\rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{array} \right] \quad \begin{aligned} a_1 &= -1 \\ a_2 &= -1 \\ a_3 &= 3. \end{aligned}$$

$$\text{So, } P_M f = -v_1 - v_2 + 3v_3 = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

As a check: $M = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

$$\text{So, easy to see } P_M f = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

A different way is to Gram-Schmidt orthogonalize v_1, v_2, v_3 to obtain orthonormal vectors u_1, u_2, u_3 . They form an orthonormal basis for M .

$$\text{Then. } P_M f = \langle f, u_1 \rangle u_1 + \langle f, u_2 \rangle u_2 + \langle f, u_3 \rangle u_3.$$

Example Let $X = L^2(-1, 1)$, $M = P_2 = \{\text{all (real) polynomials of deg. } \leq 2\}$, $f(x) = x^3$ ($-1 \leq x \leq 1$)

We want to find the best approximation of f in P_2 . Denote it by $p = p(x) \in P_2$.

Note $P_2 = \text{span} \{ p_0, p_1, p_2 \}$

$$\text{Let } p(x) = a_0 p_0(x) + a_1 p_1(x) + a_2 p_2(x)$$

$$\langle p, p_j \rangle = \langle f, p_j \rangle, j = 0, 1, 2$$

$p_0(x) = 1$	linearly indeps.
$p_1(x) = x$	
$p_2(x) = x^2$	

$$\left\langle p_0, p_j \right\rangle a_0 + \left\langle p_1, p_j \right\rangle a_1 + \left\langle p_2, p_j \right\rangle a_2 = \left\langle f, p_j \right\rangle$$

$$j = 0, 1, 2$$

3 linear equations for 3 unknowns a_0, a_1, a_2 .

$$\langle p_0, p_0 \rangle = \int_1^1 1 dx = 2 \quad \langle p_0, p_1 \rangle = \int_1^1 x^0 x^1 dx = 0 = \langle p_1, p_0 \rangle$$

$$\langle p_2, p_0 \rangle = \langle p_0, p_2 \rangle = \int_1^1 x^2 dx = \frac{2}{3} \quad \langle p_1, p_1 \rangle = \langle p_2, p_1 \rangle = \int_1^1 x^3 dx = 0$$

$$\langle p_2, p_2 \rangle = \int_1^1 x^4 dx = \frac{2}{5} \quad \langle p_1, p_2 \rangle = \int_1^1 x^4 dx = \frac{2}{5}$$

$$\langle f, p_0 \rangle = \int_1^1 x^3 dx = 0, \quad \langle f, p_1 \rangle = \int_1^1 x^4 dx = \frac{2}{5}, \quad \langle f, p_2 \rangle = 0.$$

$$A = \begin{bmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{5} & 0 & \frac{2}{5} \end{bmatrix} \quad A \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{2}{5} \\ 0 \end{bmatrix}$$

$$a_0 = 0, \quad a_1 = \frac{3}{5}, \quad a_2 = 0.$$

$$p(x) = \frac{3}{5} p_1(x) = \frac{3}{5} x.$$

From these examples, we see a general procedure.
We describe it in the next Theorem.

Theorem Let X be a ^(real) Hilbert space. Let M be an n -dimensional subspace of X . Let $f \in X$. Suppose $\{u_1, \dots, u_n\}$ is a basis for M . Then the best approximation of f in M is

$$P_M f = a_1 u_1 + a_2 u_2 + \dots + a_n u_n$$

where a_1, a_2, \dots, a_n are the unique solution of the system of linear equations

$$G(u_1, \dots, u_n) \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \langle f, u_1 \rangle \\ \vdots \\ \langle f, u_n \rangle \end{bmatrix}.$$

where $G(u_1, \dots, u_n) = [\langle u_j, u_k \rangle]_{j,k=1}^n$ is an $n \times n$

matrix (called the Gram matrix of u_1, \dots, u_n).

If $\{u_1, u_2, \dots, u_n\}$ is an orthonormal basis for M , then $G(u_1, \dots, u_n) = I$ is the identity matrix. Hence

$$a_j = \langle f, u_j \rangle \quad (j=1, \dots, n) \text{ and}$$

$$P_M f = \sum_{j=1}^n \langle f, u_j \rangle u_j.$$