Solutions to Homework #5, Math 116

Keziah Cook and Michael McElroy

November 24, 2003

Problem 3.7 (Luenberger)

Let M and N be orthogonal closed subspaces of a Hilbert space H and let $x \in H$ be given. Show that $M \oplus N$ is closed and that the orthogonal projection of x onto $M \oplus N$ is equal to $m + n$, where m and n are the orthogonal projections of x onto M and N respectively.

Solution to 3.7

Why do we care that $M \oplus N$ is closed? So that we may apply the projection theorem, and to know that a minimizer exists. We want limits in $M \oplus N$ to remain in the set. This follows from the closure of M and N.

Now, M and N are subspaces, so $0 \in M, N$. M and N are orthogonal, so for all $m' \in M$ and $n' \in N$ we have $\langle m'|n'\rangle = 0$. This is what $M \perp N$ means. We have that $\langle x - m|m - m'\rangle = 0$, but since $m - m' = m'' \in M$, we may as well write $\langle x - m|m\rangle = \langle x - m|m'\rangle = 0$. Same for N, $\langle x - n|n\rangle = \langle x - n|n'\rangle = 0$. Now, we want to show that $x - (m + n) \perp M \oplus N$, so that $\langle x - m - n|m + n - m' - n'\rangle = 0$. Expanding, we have

$$\langle x - m - n|m + n - m' - n'\rangle = \langle x - m|m - m'\rangle + \langle x - n|n - n'\rangle - \langle n|m - m'\rangle - \langle m|n - n'\rangle.$$

Since m and n are minimizers to M and N, the first two terms equal zero, and since $M \perp N$, the last two terms equal zero. □

Problem 3.8 (Luenberger)

Let $\{x_1, x_2, \ldots, x_m\}$ and $\{y_1, y_2, \ldots, y_n\}$ each be sets of linearly independent vectors in a Hilbert space H generating the subspaces M and N, respectively. Given a vector $x \in H$, find the best approximation $\hat{x}$ (i.e. minimizing $\|x - \hat{x}\|$) in the subspace $M \cap N$.

(a) Give a set of equations equivalent to the normal equations which produce $\hat{x}$.

(b) Give a geometrical interpretation of the solution method.

(c) Give a computational procedure for producing $\hat{x}$.

Solution to 3.8

(a) We need to find out what is $M \cap N$. We'll project N into M: for each i, $y_i = y_{i,M} + y_{i,M^\perp}$, and $y_{i,M} \perp y_{i,M^\perp}$. That is, $H = M \oplus M^\perp$, so each vector can be expressed as the sum of a component in M and the component not in M. In this case, $y_{i,M} = \sum_{k=1}^{m} \frac{\langle y_i|x_k\rangle}{\langle x_k|x_k\rangle} x_k$, the sum of the projections of y_i onto the basis vectors of M; $y_{i,M^\perp} = y_i - y_{i,M}$.

So $M \cap N = [y_{1,M}, y_{2,M}, \ldots, y_{n,M}]$, the span of these vectors. These vectors may not linearly independent. We could write the normal equations with them, but then we'd have a Gram matrix which was not invertible. One time or another, we'll need to rid ourselves of the redundant equations.

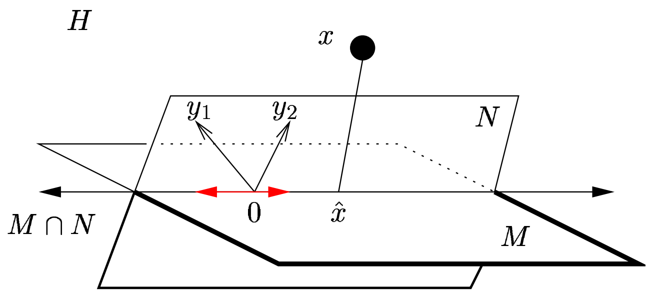

Figure 1: Projecting basis vectors for N ($\{y_1, y_2\}$) onto M to find a spanning set for $M \cap N$.

Idea 1 Let's write out the Gram matrix version. If the projection of $x \in H$ onto $M \cap N$ is $\hat{x}$ then $\hat{x} = \sum_{j=1}^{n} \alpha_j y_{j,M}$. The projection theorem gives us that $\langle x - \hat{x} | y_{j,M} \rangle = 0$ for all j. So we write

$$\begin{pmatrix} \langle y_{1,M} | y_{1,M} \rangle & \cdots & \langle y_{n,M} | y_{1,M} \rangle \\ \vdots & \ddots & \vdots \\ \langle y_{n,M} | y_{1,M} \rangle & \cdots & \langle y_{n,M} | y_{n,M} \rangle \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \langle x | y_{1,M} \rangle \\ \vdots \\ \langle x | y_{n,M} \rangle \end{pmatrix}, \qquad \text{or} \quad G\alpha = b.$$

Here the Gram matrix G may not be invertible, so there may be many solutions for α, but we only need one. Use Gaussian elimination.

Idea 2 Do Gram-Schmidt orthogonalization first to get a basis for $M \cap N$. Then write out the Gram matrix equation using this basis for $M \cap N$, invert the Gram matrix, and go home happy.

Idea 3 It does not work to project x onto M and then project that projection onto N. Consider $H = \mathbb{R}^2$, and M and N lines with bases $x_1 = (1, 0)$ and $y_1 = (1, 1)$. So N is the "$y = 0$" line and N is the "$y = x$" line. Let $x = (-2, 2)$. Since $M \cap N$ is only one point, the origin $(0, 0)$, we know that this is the problem's solution. However, projecting x onto M yields $(-2, 0)$, and projecting this onto N yields $(-1, -1)$.

This idea worked in problem 3.6 but not here. Why not? In problem 3.6, the space of polynomials of degree $\leq n$ was M and N was the subspace of M where the integral of the polynomial over the domain equalled zero. So there $M \cap N = N$, while here in 3.8 N is not necessarily a subspace of M.

(b) Geometry picture - see figure 1.

(c) Another idea is to say $M \cap N$ is a closed subspace of H, so $H = (M \cap N) \oplus (M \cap N)^{\perp}$. If we can project x into $(M \cap N)^{\perp}$, say $x_{(M \cap N)^{\perp}}$, then our solution is $x - x_{(M \cap N)^{\perp}}$ (i.e. x minus its projection into the $(M \cap N)^{\perp}$ space). We know this from problem 3.7.

What is $(M \cap N)^{\perp}$? It's the space of elements perpendicular to M or N, i.e. $(M \cap N)^{\perp} = M^{\perp} + N^{\perp}$. There is no reason to believe that we should use the direct sum $\oplus$, since if $M = N$ we have a clear contradiction (problems arise if M and N have a non-trivial intersection).

In the end, I do not see an easy way of projecting onto $(M \cap N)^{\perp} = M^{\perp} + N^{\perp}$; you still need a basis for $(M \cap N)^{\perp}$, and this could be infinite-dimensional. Our computational approaches(Gram matrix) assumes a finite-dimensional space. We can get around the dimensionality problem by first projecting x onto $M + N$, and then considering the subproblem of minimizing the distance of that projection to $M \cap N$ (see problem 3.6). Still, it is not clear to me how this is better than our approach in part a.

$\square$

Problem 3.16 (Luenberger)

Prove Parseval's equality: an orthonormal sequence $\{e_i\}_{i=1}^{\infty}$ is complete in a Hilbert space H if and only if for each x, y in H

$$(x \mid y) = \sum_{i=1}^{\infty} (x \mid e_i)\,(e_i \mid y) \tag{1}$$

Solution to 3.16

- $\rightarrow$

  By assumption $\{e_i\}_{i=1}^{\infty}$ is complete. Thus for any $x, y \in H$, there exist α_i, β_i such that $x = \sum_{i=1}^{\infty} \alpha_i e_i$ and $y = \sum_{i=1}^{\infty} \beta_i e_i$.

$$\begin{aligned}(x \mid y) &= \left(\sum_{i=1}^{\infty} \alpha_i e_i \mid \sum_{j=1}^{\infty} \beta_j e_j\right) \\ &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \alpha_i \overline{\beta_j}\,(e_i \mid e_j)\end{aligned}$$

  We can get rid of the double sum, because $(e_i \mid e_j) = 0$ for $i \neq j$

$$\begin{aligned} &= \sum_{i=1}^{\infty} \alpha_i \overline{\beta_i}\,(e_i \mid e_i) \\ &= \sum_{i=1}^{\infty} \alpha_i \overline{\beta_i}\end{aligned}$$

  as $(e_i \mid e_i) = 1$
  We note that $(x \mid e_i) = (\sum_{j=1}^{\infty} \alpha_j e_j \mid e_i) = \alpha_i$ and $(e_i \mid y) = (e_i \mid \sum_{j=1}^{\infty} \beta_j e_j) = \overline{\beta_i}$. Thus we see

$$(x \mid y) = \sum_{i=1}^{\infty} (x \mid e_i)\,(e_i \mid y)$$

  as desired.

- $\leftarrow$

  Now we need to show that whenever Parseval's identity holds, the sequence $e_{i}{}_{i=1}^{\infty}$ is complete.

  Let $\{e_i\}_{i=1}^{\infty}$ be an orthonormal sequence in H for which Parseval's identity holds. Assume it is not complete. Let M be the closed span of $\{e_i\}_{i=1}^{\infty}$. Then there exists $x \in H$ such that x is not in M. Also, we can define $\hat{x} = \sum_{i=1}^{\infty} (x \mid e_i)\, e_i$, the projection of x onto M.

Now let's look at $\|x-\hat{x}\|^2 = (x-\hat{x} \mid x-\hat{x})$. This expands to:

$$
\begin{aligned}
(x-\hat{x} \mid x-\hat{x}) &= \left(x - \sum_{i=1}^{\infty} (x \mid e_i)\, e_i \mid x - \sum_{i=1}^{\infty} (x \mid e_i)\, e_i \right) \\
&= (x \mid x) - \sum_{i=1}^{\infty} \overline{(x \mid e_i)}\,(x \mid e_i) - \sum_{i=1}^{\infty} (x \mid e_i)\,(e_i \mid x) + \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} \overline{(x \mid e_i)}\,(x \mid e_j)\,(e_i \mid e_j) \\
&= (x \mid x) - \sum_{i=1}^{\infty} (e_i \mid x)\,(x \mid e_i) - \sum_{i=1}^{\infty} (x \mid e_i)\,(e_i \mid x) + \sum_{i=1}^{\infty} (e_i \mid x)\,(x \mid e_j) \\
&= (x \mid x) - \sum_{i=1}^{\infty} (x \mid e_i)\,(e_i \mid x)
\end{aligned}
$$

But by Parseval's identity, we know $(x \mid x) = \sum_{i=1}^{\infty} (x \mid e_i)\,(e_i \mid x)$. Thus $\|x-\hat{x}\|^2 = 0$. This means $x = \hat{x}$, which implies $H = M$ and therefore $\{e_i\}_{i=1}^{\infty}$ is complete.

□

Problem 3.17 (Luenberger)

Let $\{y_1, \ldots, y_n\}$ be independent and suppose $\{e_1, e_2, \ldots, e_n\}$ are obtained from the y_i's be the Gram-Schmidt procedure. Let $\hat{x} = \sum_{i=1}^{n} \langle x|e_i\rangle e_i = \sum_{i=1}^{n} \alpha_i y_i$. Show how to obtain the α_i coefficients from the Fourier coefficients $\langle x|e_i\rangle$.

Solution to 3.17

So, $\{e_i\}$ is an orthonormal basis of $[\{y_j\}]$. We have the projection of $x \in H$ onto the subspace $[\{y_j\}]$, and we have the expansion of $\hat{x}$ in our basis $\{e_i\}$. We want to express $\hat{x}$ in terms of the non-orthonormal basis $\{y_i\}$.

Call the Fourier coefficients β_i, so that $\beta_i = \langle x|e_i\rangle$. Taking the inner product of $\hat{x}$ with any basis vector e_k yields

$$\langle \hat{x}|e_k\rangle = \left\langle \sum_{i=1}^{n} \beta_i e_i | e_k \right\rangle = \beta_k.$$

However, $\langle \hat{x}|e_k\rangle$ also equals $\langle \sum_{i=1}^{n} \alpha_i y_i | e_k\rangle = \sum_{i=1}^{n} \alpha_i \langle y_i|e_k\rangle$. Now we have an equation relating the β's and y_i's: $\beta_k = \sum_{i=1}^{n} \alpha_i \langle y_i|e_k\rangle$. We may construct a matrix equation from the β_i equations:

$$\begin{pmatrix} \langle y_1|e_1\rangle & \cdots & \langle y_n|e_1\rangle \\ \vdots & \ddots & \vdots \\ \langle y_1|e_n\rangle & \cdots & \langle y_n|e_n\rangle \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}, \qquad \text{or} \quad A\alpha = \beta.$$

So $\alpha = A^{-1}\beta$. □

Problem 3.20 (Luenberger)

Over ten months a company is to hire and fire all its employees responsible for the making of one-million widgets. Assume the company hires (and fires) at the beginning of each month and may fire all the remaining employees at the end of the tenth month (so there are eleven distinct hire/fire periods). Each employee makes 100 widgets a month, and their labor costs is proportional to their numbers. The cost of hiring and firing is proportional to the square of the number being hired/fired. Minimize the cost.

Solution to 3.20

Let u_i be the number hired on the i^{th} period. If $u_i < 0$ then the number is fired. The requirement that all are released at the end means that $\sum_{i=1}^{11} u_i = 0$. The number of workers in the j^{th} month equals $\sum_{i=1}^{j} u_i$. Hence the production requirement boils down to $10^6 = \sum_{p=1}^{11} 100 \cdot \sum_{j=1}^{p} u_j$. A better way of writing this is $10^4 = \sum_{p=1}^{11}(10-p)u_p$.

Labor costs are proportional to the number of workers which is proportional to the work done, and the work to be done is fixed (10^6 widgets), so the total number of workers in the ten periods is fixed. Hence we may ignore labor costs. Our problem is to minimize $\sum_{p=1}^{11} u_p^2$ such that $\sum_{i=1}^{11} u_i = 0$ and $\sum_{p=1}^{11}(10-p)u_p = 10^4$. Letting $u = [u_1, u_2, \ldots, u_{11}]^T$, we may rewrite our problem as

$$\text{Minimize } \|u\|^2 \quad \text{such that} \quad Au = b$$

$$\text{where } A = \begin{pmatrix} 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 10^4 \\ 0 \end{pmatrix}.$$

Minimizing the norm squared is equivalent to minimizing the norm. See the solution to 3.21 below, where I say that the smallest vector such that $Ax = b$ is $x_{min} = A^T(AA^T)^{-1}b$. Using this formula, we have

$$u_{min} = A^T(AA^T)^{-1}b \approx \begin{pmatrix} 454.5 & 363.6 & 272.7 & 181.8 & 90.9 & 0 & -90.9 & -181.8 & -272.7 & -363.6 & -454.5 \end{pmatrix}^T.$$

We want a solution with integers, and its not hard to verify that the optimal thing to do is to round down the 454.5 to 454 and round up the others, i.e.

$$u_{min} = \begin{pmatrix} 454 & 364 & 273 & 182 & 91 & 0 & -91 & -182 & -273 & -364 & -454 \end{pmatrix}^T.$$

□

Problem 3.21 (Luenberger)

> Using the projection theorem, solve the finite-dimensional problems, minimize x^TQx s.t. $Ax = b$. x is an n-vector, Q is positive definite symmetric matrix, and A is an $m \times n$ matrix.

Solution to 3.21

Luenberger theorem 2, page 65, states that the minimizer for $\|x\| = x^Tx$ given $Ax = b$ is $x_{min} = A^T(AA^T)^{-1}b$. He uses notation where $A = [y_1, y_2, \ldots y_m]^T$ and $b = [c_1, c_2, \ldots, c_m]^T$. In our problem, we have an inner product $\langle x|y\rangle = x^TQy$, and norm $\|x\| = \sqrt{\langle x|x\rangle}$. Minimizing $\|x\|$ is equivalent to minimizing $\|x\|^2$, so we see that this problem is really a minimum norm problem.

However, the equation that Luenberger gives assumes a different inner product, so we need a change of coordinates. First, a theorem about positive definite square matrices.

Theorem (Cholesky decomposition). *Given any symmetric positive definite matrix Q, we have an upper-triangular matrix R such that $Q = R^TR$.*

Since Q is non-singular so is R. Now our choice of coordinates is obvious: let $y = Rx$. Now $x^TQx = x^TR^TRx = y^Ty$. Our constraint equation $Ax = b$ becomes $b = Ax = AR^{-1}Rx = (AR^{-1})y$. Letting $C = AR^{-1}$, we have

$$y_{min} = C^T(CC^T)^{-1}b = (AR^{-1})^T\left(AR^{-1}R^{T^{-1}}A^T\right)^{-1}b = R^{T^{-1}}A^T\left(AQ^{-1}A^T\right)^{-1}b.$$

So $x_{min} = R^{-1}y_{min} = Q^{-1}A^T\left(AQ^{-1}A^T\right)^{-1}b$. □

Writing

(1)

(2) Legendre polynomials have a simple form: $P_n(x) = \frac{1}{2^n n}\left[(x^2-1)^n\right]^{(n)}$ for $n \in \mathbb{N}$. We may work them out, however, by a Gram-Schmidt procedure on the polynomials $\{1, x, x^2, \ldots\}$.

I choose to orthogonalize the polynomials first, and then to normalize them last (you may normalize whenever you want). The general equation for the n^{th} element z_n in the orthogonal sequence is

$$z_n = x^n - \sum_{i=0}^{n-1} \frac{\langle x^n | x^i \rangle}{\langle x_i | x_i \rangle} x_i.$$

Our inner product is $\langle y(x)|z(x)\rangle = \int_{-1}^{1} y(x)z(x)\,\mathrm{d}x$. Anyway, the good thing is that even functions are orthogonal to odd functions, so $1 \perp t$. Hence $z_0 = 1$ and $z_1 = t$. Furthermore 1 and t^2 are not orthogonal, so $z_2 = x^2 - \frac{\int_{-1}^{1} x^2\,\mathrm{d}x}{\int_{-1}^{1} 1\,\mathrm{d}x} 1 = x^2 - \frac{1}{3}$. Now, $1 \perp x^3$ and $x^2 - \frac{1}{3} \perp x^3$, but x and x^3 are not orthogonal. So $z_3 = x^3 - \frac{\int_{-1}^{1} x^4\,\mathrm{d}x}{\int_{-1}^{1} x^2\,\mathrm{d}x} x = t^3 - \frac{3}{5}t$.

Normalizing the vectors leads to our orthonormal sequence $\{b_i\}$, $b_n = z_n/\|z_n\|$:

$$\begin{aligned} b_0 &= \frac{\sqrt{2}}{2}, \\ b_1 &= \sqrt{\frac{3}{2}}x = \frac{\sqrt{6}}{2}x, \\ b_2 &= \sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right) = \frac{3}{4}\sqrt{10}\left(x^2 - \frac{1}{3}\right) \end{aligned}$$

and

$$b_3 = \sqrt{\frac{175}{8}}\left(x^3 - \frac{3}{5}x\right) = \frac{5}{4}\sqrt{14}\left(x^3 - \frac{3}{5}x\right).$$

Note that the actual Legendre polynomials defined above as $P_n(x)$ are not normalized: e.g. $P_2(x) = \frac{3}{2}\left(x^2 - \frac{1}{3}\right)$, not $\frac{3}{4}\sqrt{10}\left(x^2 - \frac{1}{3}\right)$.