The Best Interpolation Problem

Oksana Bihun*

July 17, 2007

Continuous Optimization Workshop, MSRI
Notes on Project 4

1 The Best Interpolation Problem

Consider the problem of minimizing the functional

$$I(u) = \int_0^1 u^2(t)\,dt \tag{1.1}$$

over the admissible set

$$A = \{u \in L^2[0,1] : J(u) = C\}, \tag{1.2}$$

where the vector valued linear function $J : L^2[0,1] \to R^n$ is defined to be

$$J(u)_k = J_k(u) = \int_0^1 \psi_k(t)u(t)\,dt \tag{1.3}$$

for given functions $\psi_k \in L^2[0,1]$. The vector $C \in R^n$ belongs to the range of J, i.e. $C_i = \int_0^1 \psi_i(t)\hat{x}(t)\,dt$ for some $\hat{x} \in L^2[0,1]$.

Note that the admissible set A is convex, closed and nonempty. The existence of minimum of the functional I over the set A follows from the existence of the closest point to the origin in the closed convex set $A \subset L^2[0,1]$.

References for such problems are in books such as: [6, 3, 1, 4, 5]

2 A Necessary Condition for the Minimizer. The Exact Solution

We will derive a necessary condition for a function $u \in A$ to be a minimum.

Theorem 2.1 (Lagrange Multipliers Rule) *Suppose that the set of constraints* $J(u) = C$ *in minimization problem*

$$\min I(u) \quad \text{subject to} \quad u \in A \tag{2.1}$$

*University of Missouri-Columbia

satisfies the following property. There exists a set of $L^2[0,1]$ *functions* $w_1, \ldots, w_n$ *such that the matrix* $Y \in R^{n\times n}$ *with the components*

$$Y_{ij} = J_i(w_j) = \int_0^1 \psi_i(t) w_j(t)\, dt \tag{2.2}$$

is nonsingular.

Suppose that $u \in A$ *is a solution of problem (2.1). Then there exists a vector* $\lambda \in R^n$ *such that for every* $v \in L^2[0,1]$

$$\int_0^1 u(t)v(t)\, dt = \sum_{j=1}^n \lambda_j \int_0^1 \psi_j(t)v(t)\, dt. \tag{2.3}$$

Proof. This proof uses the idea from the proof of theorem 2 from chapter 8.4 of the book "Partial Differential equations" by L. C. Evans [2].

Let $v \in L^2[0,1]$. Consider the function $j : R^{n+1} \to R^n$ with components

$$\begin{aligned} j_k(\tau, \sigma_1, \ldots, \sigma_n) &= J_k(\tau v + \sigma_1 w_1 + \ldots + \sigma_n w_n) && (2.4)\\ &= \int_0^1 \psi_k(t)(\tau v(t) + \sigma_1 w_1(t) + \ldots + \sigma_n w_n(t))\, dt, && (2.5) \end{aligned}$$

where $k = 1, \ldots, n$ and the functions $w_1, \ldots w_n \in L^2[0,1]$ satisfy the condition from the hypothesis. Note that $j(0,\ldots,0) = 0 \in R^n$ and $D_\sigma j(0,\ldots,0) = Y$ is nonsingular. Therefore, by the Implicit Function Theorem, there exists a differentiable function $\phi : R \to R^n$ defined in a neighborhood $Q = \{\tau \in R : |\tau| \le \tau_0\}$ of the origin such that

$$j(\tau, \phi(\tau)) = 0 \tag{2.6}$$

for all $\tau \in Q$. Moreover,

$$\phi'(0) = -[D_\sigma j(0,\ldots,0)]^{-1} D_\tau j(0,\ldots,0) = -Y^{-1} b, \tag{2.7}$$

where $b \in R^n$ is the vector with the components $b_k = \int_0^1 \psi_k(t)v(t)\, dt$. Now, consider the variation $u + \tau v + \phi(\tau)\cdot w$ of the function u. Because $J(u + \tau v + \phi(\tau)\cdot w) = J(u) + j(\tau, \phi(\tau)) = C$, this vatiation belongs to the admissible set A. Therefore, the function $i(\tau) := I(u + \tau v + \phi(\tau)\cdot w)$ attains its local minimum at the origin, which implies the equality $i'(0) = 0$. The latter equality can be recast in the form

$$\int_0^1 u(t)\big(v(t) + \phi'(0)\cdot w(t)\big)\, dt = 0.$$

Hence,

$$\int_0^1 u(t)v(t)\, dt = -\int_0^1 \phi'(0)\cdot w(t)\, dt. \tag{2.8}$$

Using expression (2.7) for $\phi'(0)$, we find that

$$\phi'(0)\cdot w = -Y^{-1} b\cdot w = -\sum_{j=1}^n \sum_{i=1}^n [Y^{-1}]_{ij} \int_0^1 \psi_j(s)v(s)\, ds w_i. \tag{2.9}$$

Substituting this expression into equation (2.8), we obtain

$$\begin{aligned}\int_0^1 u(t)v(t)\,dt &= \int_0^1 \sum_{j=1}^n \sum_{i=1}^n [Y^{-1}]_{ij} \int_0^1 \psi_j(s)v(s)w_i(t)\,ds\,dt \\ &= \sum_{j=1}^n \int_0^1 \sum_{i=1}^n [Y^{-1}]_{ij} w_i(t)\,dt \int_0^1 \psi_j(s)v(s)\,ds. \qquad (2.10)\end{aligned}$$

Define $\lambda_j = \int_0^1 \sum_{i=1}^n [Y^{-1}]_{ij} w_i(t)\,dt$. Then equation (2.10) can be rewritten in the form

$$\int_0^1 u(t)v(t)\,dt = \sum_{j=1}^n \lambda_j \int_0^1 \psi_j(t)v(t)\,dt \qquad (2.11)$$

as required.

■

Remark 2.2 *The latter theorem holds if we assume that the functions* ψ_i, $i = 1, \ldots, k$, *are such that their Gramm matrix* $Y = \{\langle \psi_i, \psi_j \rangle_{L^2[0,1]}\}_{i,j=1}^n$ *is nonsingular.*

Corollary 2.3 *Suppose that the Gramm matrix* Y *of the constraint functions* ψ_i *given by* $Y = \{\langle \psi_i, \psi_j \rangle_{L^2[0,1]}\}_{i,j=1}^n$ *is nonsingular. Then there exists a unique minimizer of the functional* I *in the admissible set* A.

Moreover, the constant vector $\lambda = (\lambda_1, \ldots, \lambda_n)^T$ *satisfying relation (2.3) from the previous theorem is given by the formula*

$$\lambda = Y^{-1}C,$$

where $C = (C_1, \ldots, C_n)^T$.

The minimum of the functional I *over the admissible set* A *is given by the formula*

$$I_{\min} = \min_{u \in A} I(u) = C^T Y^{-1} C$$

and is attained on the function $u = \sum_{i=1}^n \lambda_i \psi_i$, *where* $\lambda = Y^{-1}C$.

Proof. As it was already mentioned, the existence of a minimizer follows from the fact that the admissible set A is a nonempty closed convex set: there exists a closest point to the origin in A.

Let us prove the uniqueness. Suppose that $u_1, u_2 \in A$ minimize the functional I over A. Then $I(u_1) = I(u_2) = I_{\min}$. Let us apply theorem 2.1 to the function u_1. There exists $\lambda \in R^n$ such that for all $v \in L^2[0,1]$

$$\int_0^1 u_1(t)v(t)\,dt = \sum_{i=1}^n \lambda_i \int_0^1 \psi_i(t)v(t)\,dt. \qquad (2.12)$$

Set $v = \psi_k$, then

$$\int_0^1 u_1(t)\psi_k(t)\,dt = \sum_{i=1}^n \lambda_i \int_0^1 \psi_i(t)\psi_k(t)\,dt$$

for $k = 1, \ldots, n$ or, equivalently,

$$Y\lambda = C,$$

which proves that $\lambda = Y^{-1}C$.

By repeating the same argument for u_2, we obtain that

$$\int_0^1 u_2(t)v(t)\,dt = \sum_{i=1}^{n} \lambda_i \int_0^1 \psi_i(t)v(t)\,dt \tag{2.13}$$

for all $v \in L^2[0,1]$, where $\lambda = Y^{-1}C$. Comparing equations (2.12) and (2.13), we conclude that $\int_0^1 u_1(t)v(t)\,dt = \int_0^1 u_2(t)v(t)\,dt$ for all $v \in L^2[0,1]$, which implies that $u_1 = u_2$ as functions in $L^2[0,1]$ (almost everywhere). By setting $v = u_1$ in (2.12), we obtain

$$I_{\min} = I(u_1) = \int_0^1 u_1^2(t)\,dt = \sum_{i=1}^{n} \lambda_i C_i = C^T Y^{-1} C.$$

It remains to check that $u = \sum_{i=1}^{n} \lambda_i \psi_i$ is the unique minimizer of I in the admissible set A. It is easy to see that $u \in A$. Indeed,

$$\int_0^1 \psi_k(t)u(t)\,dt = \sum_{i=1^n} \lambda_i Y_{ik} = C_k.$$

The value of the functional I on u is

$$I(u) = \int_0^1 (\sum_{i=1}^{n} \lambda_i \psi_i(t))^2\,dt \tag{2.14}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} Y_{ij}\lambda_i\lambda_j. \tag{2.15}$$

Substituting $\lambda = Y^{-1}C$ into the latter equality, it is easy to obtain that

$$I(u) = C^T Y^{-1} C = I_{\min}.$$

The proof is complete.

■

Example 2.4 *The Exact Solution for the Case of One Constraint.*

Consider the problem of minimizing the functional $I(u) = \int_0^1 u^2(t)\,dt$ *over the admissible set* $A = \{u \in L^2[0,1] : \int_0^1 \psi(t)u(t)\,dt = C\}$, *where* $0 \neq \psi \in L^2[0,1]$ *and* $C = \int_0^1 \psi(t)\hat{x}(t)\,dt$ *for some* $\hat{x} \in L^2[0,1]$. *Without loss of generality, assume that* $C \geq 0$. *By corollary 2.3, the following lemma holds.*

Lemma 2.5 *Consider the optimization problem (2.1), where the admissible set*

$$A = \{u \in L^2[0,1] : \int_0^1 \psi(t)u(t)\,dt = C\}$$

with $0 \neq \psi \in L^2[0,1]$ *and* $C = \int_0^1 \psi(t)\hat{x}(t)\,dt \geq 0$*. Then the function* $u = \frac{C}{\|\psi\|_2^2}\psi$ *minimizes the functional* I *over the admissible set* A*.*

It is easy to prove the lemma without using corollary 2.3 or theorem 2.1. Indeed, for $u = \frac{C}{\|\psi\|_2^2}\psi$ *we have that* $I(u) = I(\frac{C}{\|\psi\|_2^2}\psi) = \frac{C^2}{\|\psi\|_2^2}$*. On the other hand, for every* $w \in A$ *the following inequality holds:*

$$C = \int_0^1 \psi(t)w(t)\,dt \leq \Big(\int_0^1 w^2(t)\,dt\Big)^{1/2}\Big(\int_0^1 \psi^2(t)\,dt\Big)^{1/2}. \tag{2.16}$$

Recall that $C \geq 0$ *(we have assumed it in the beginning; we can always recast the problem to an equivalent one with* $C \geq 0$*). Hence, squaring inequality (2.16), we obtain*

$$I(u) = \frac{C^2}{\int_0^1 \psi^2(t)\,dt} \leq \int_0^1 w^2(t)\,dt = I(w)$$

for all $w \in A$*, which proves that* $u = \frac{C}{\|\psi\|_2^2}\psi$ *is the global minimum of* I *over the admissible set* A*.*

References

[1] F. Deutsch. *Best Approximation in Inner Product Spaces.* Springer–Verlag, New York, 2001.

[2] L. Evans. *Partial Differential Equations*, American Mathematical Society, Providence, 1998.

[3] R. B. Holmes. *A Course on Optimization and Best Approximation.* Springer–Verlag, Berlin, 1972.

[4] Harold S. Shapiro. *Smoothing and approximation of functions.* Van Nostrand Reinhold Co., New York, 1969. Revised and expanded edition of mimiographed notes (Matscience Report No. 55), Van Nostrand Reinhold Mathematical Studies.

[5] Harold S. Shapiro. *Topics in approximation theory.* Springer-Verlag, Berlin, 1971. With appendices by Jan Boman and Torbjörn Hedberg, Lecture Notes in Math., Vol. 187.

[6] I. Singer. *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces.* Springer–Verlag, Berlin, 1970.