5. A BEST APPROXIMATION PROBLEM

Consider the following problem:

PROBLEM. Given the real symmetric $n\times n$ matrix B and the three subspaces L_1, L_2, and L_3 of R^n, find the (unique) real symmetric $n\times n$ matrix A which is closest to B in Frobenius norm (Hilbert-Schmidt norm) and which is negative semidefinite (nsd) on L_1, positive semidefinite (psd) on L_2, and 0 on L_3.

Solution. First, it is clear that A must be 0 on $L_1\cap L_2$. Thus, we can rewrite the problem so that A must be: 0 on $L_3+L_1\cap L_2$; nsd on L_1', any complementary subspace of $L_1\cap L_2+L_1\cap L_3$ in L_1; and psd on L_2', any complementary subspace of $L_1\cap L_2+L_2\cap L_3$ in L_2. Now set $L_3'=L_3+L_1\cap L_2$, and let: P_1 be the projection on L_1' along any complementary subspace of R^n which contains $L_2'+L_3'$; P_2 be the projection on L_2' along any complementary subspace which contains $L_1'+L_3'$; and P_3 be the projection on L_3' along any complementary subspace which contains $L_1'+L_2'$. We now define the unitary diagonalizations

$$\begin{aligned} P_iBP_i^t &= U_iD_iU_i^t \\ &= U_iD_i^+U_i^t+U_iD_i^-U_i^t, \quad i=1,2,3, \end{aligned}$$

where the U_i are the unitary matrices of eigenvectors, D_i are the diagonal matrices of eigenvalues, and D_i^+ and D_i^- are the diagonal matrices of positive and negative eigenvalues. We let S be the cone of $n\times n$ psd matrices, in the space $Y=R^{(n^2+n)/2}$ of $n\times n$ real symmetric matrices represented by their distinct triangular parts. The norm in Y is given by the Euclidean inner product

$$\langle A,B\rangle=\operatorname{tr}AB,$$

the trace of the matrix product AB. We define the projection in Y,

$$\mathcal{P}\cdot=I\cdot-P_4\cdot P_4,$$

where P_4 is the (orthogonal) projection on $L_1'+L_2'+L_3'$. Let us now show that the solution is

$$A=U_1D_1^-U_1^t+U_2D_2^+U_2^t+\mathcal{P}B. \tag{5.1}$$

Choose the matrices E_1, E_2, and E_3 so that

$$L_i' = \mathcal{R}(E_i),$$

where $\mathcal{R}(\ )$ denotes range space. The matrix X in Y is nsd (psd) on L_1' (L_2') if and only if $E_1^t X E_1$ ($E_2^t X E_2$) is nsd (psd) on all of R^n, since

$$(E_i^t X E_i y, y) = (X(E_i y), (E_i y)) \qquad \text{for} \quad y \text{ in } R^n.$$

Now we can rewrite the problem as the abstract convex program

$$(\mathrm{P}) \qquad \begin{cases} \text{minimize} & p(X) = \tfrac{1}{2}\|X-B\|^2 = \tfrac{1}{2}\operatorname{tr}(X-B)^2 \\ \text{subject to} & g_1(X) = E_1^t X E_1 \in -S, \\ & g_2(X) = -E_2^t X E_2 \in -S, \\ & g_3(X) = E_3^t X E_3 \in \{0\}. \end{cases}$$

The generalized Lagrange multiplier theorem states that if A satisfies the constraints and

$$0 = \nabla p(A) + \nabla\langle \lambda_1, g_1(A)\rangle + \nabla\langle \lambda_2, g_2(A)\rangle + \nabla\langle \lambda_3, g_3(A)\rangle,$$
$$0 = \langle \lambda_i, g_i(A)\rangle, \qquad i = 1,2,3, \tag{5.2}$$

for some matrices λ_1, λ_2 in S^+, λ_3 in $\{0\}^+ = Y$, where ∇ denotes gradient, then A solves (P).

Let A be as in (5.1) and set

$$\lambda_1 = E_1^\dagger U_1 D_1^+ U_1^t E_1^{t\dagger},$$

$$\lambda_2 = E_2^\dagger U_2 D_2^- U_2^t E_2^{t\dagger},$$

$$\lambda_3 = E_3^\dagger U_3 D_3 U_3^t E_3^{t\dagger},$$

where $E_i^\dagger$ denotes the generalized inverse of E_i (see e.g. [4]) which satisfies

$$E_i E_i^\dagger = P_i.$$

Then

$$\begin{aligned} g_1(A) &= E_1^t A E_1 \\ &= E_1^t\left(U_1 D_1^- U_1^t + U_2 D_2^+ U_2^t + \mathcal{P}B\right)E_1 \\ &= E_1^t P_1^t\left(U_1 D_1^- U_1^t + P_2^t U_2 D_2^+ U_2^t P_2\right)P_1 E_1 \\ &= E_1^t U_1 D_1^- U_1^t E_1, \qquad \text{since} \quad P_2 P_1 = 0 \\ &\in -S, \qquad \text{since} \quad D_1^- \text{ is nsd.} \end{aligned}$$

Similarly

$$g_2(A) = -E_2^t U_2 D_2^+ U_2^t E_2 \in -S$$

and

$$g_3(A) = E_3^t A E_3 = 0.$$

Moreover, both λ_1 and λ_2 are in $S^+ = S$ (see e.g. [6]), since both D_1^+ and $-D_2^-$ are psd, while $\lambda_3 \in \{0\}^+ = Y$. We have left to show that (5.2) holds, or equivalently, after differentiating, that

$$B = A + E_1\lambda_1 E_1^t - E_2\lambda_2 E_2^t + E_3\lambda_3 E_3^t,$$

$$\operatorname{tr} \lambda_i E_i^t A E_i = 0, \qquad i = 1,2,3.$$

Now

$$\begin{aligned} A + E_1\lambda_1 E_1^t - E_2\lambda_2 E_2^t + E_3\lambda_3 E_3^t &= A + U_1 D_1^+ U_1^t + U_2 D_2^- U_2^t P_3 B P_3 \\ &= U_1 D_1 U_1^t + U_2 D_2 U_2^t + P_3 B P_3 + \mathcal{P}B \\ &= P_1 B P_1 + P_2 B P_2 + P_3 B P_3 + (I - P_4) B (I - P_4) \\ &= B; \end{aligned}$$

while

$$\begin{aligned}\operatorname{tr}\lambda_1 E_1^t A E_1 &= \operatorname{tr} E_1 \lambda_1 E_1^t A\\ &= \operatorname{tr} U_1 D_1^+ U_1^t A\\ &= 0,\end{aligned}$$

since the projections are mutually complementary and $D_1^+ D_1^- = 0$. Similarly

$$\operatorname{tr}\lambda_i E_i^t A E_i = 0, \qquad i = 2,3.$$

Uniqueness follows from the strict convexity of the objective function $p(X)$. ∎

REMARK 5.1. We were able to use the standard Lagrange multiplier theorem in the above, even though Slater's condition fails for (P). The cone S of psd matrices is very well behaved in general. In fact, if K is a face of S, then

$$S^+ + K^\perp \text{ is closed.} \tag{5.3}$$

For (see [2]) there exists a projection P in S such that $PS = K$, and moreover

$$S^+ + K^\perp = S + \mathfrak{N}(P),$$

(where $\mathfrak{N}$ denotes null space) is closed if and only if PS is closed (see [11]). Recall that the condition (5.3) is the one in Corollary 2.1 which allows one to replace $(S^f)^+$ by S^+ in Theorem 2.1.

REMARK 5.2. It is well known (see e.g. [12, p. 222]) that the Lagrange multipliers are sensitivity coefficients, i.e., if a solves the original program (P) in Section 2 with Lagrange multipler λ, while a_z solves the perturbed program (P_z) with the perturbed constraint $g(x) \in z - S$, then

$$p(a) - p(a_z) \leqslant \lambda z. \tag{5.4}$$

For the above matrix problem, suppose that we allow the following perturbation:

$$A \text{ must be "almost" nsd on } L_1,$$