Chapter 7

Hybrid methods for solving the educational testing problem

7.1 Introduction

In this chapter new methods for solving the educational testing problem are introduced. The methods described here depend upon both projection and l_1 SQP methods using a hybrid method. The hybrid method works in two stages. First stage is the projection method which converges globally so is potentially reliable but often converges at slow order. Meanwhile in the second stage there is l_1 SQP methods, in particular the method described in Section 6.4, which converges at second order if the correct rank r^* is given. The main disadvantage of the l_1 SQP methods are that they require the correct r^*. A hybrid method is one which switches between these methods and aims to combine their best features. To apply an l_1 SQP method requires a knowledge of the rank r^* and this knowledge can also be gained from the progress of the projection method. Hybrid methods can work well but there is one disadvantage. If the positive definite matrix have the same rank as the optimal positive semi–definite matrix in which the l_1 SQP method works well, then most of the time will be taken up in the first stage, using the projection method. If this converges slowly then the hybrid method will not solve the problem effectively. Thus it is important to ensure that the second stage method is used to maximum effect. Hence in the algorithm of Section 7.3 the l_1 SQP method is applied first.

In Sections 7.2 and 7.3 two new methods are described. Firstly, there is the projection–l_1SQP method, which starts with the projection method to determine the rank $r^{(k)}$ and continues with the l_1 SQP method. Secondly, the l_1 SQP–projection method is described, which solves the problem by the l_1 SQP method and uses the projection method to update the rank. Numerical results and comparisons are given in Section 7.4.

As with the methods of Chapter 6 it is easy to move from one method to the other in either direction. This in contrast to Chapter 4 where some special techniques were developed to enable this to be done.

7.2 Projection–l_1SQP method

The main disadvantage of the l_1 SQP method is finding the exact rank r^*, since it is not known in advance it is necessary to estimate it by an integer $r^{(k)}$. It is suggested that the best estimate of the matrix rank $r^{(k)}$ is obtained by carrying out some iterations of the projection method given in Section 6.3. This is because the projection method is a globally convergent method.

The method in this section follows a similar strategy as that in Section 4.3.

Consider Λ_r in (5.2.4), then at the solution the number of eigenvalues in Λ_r is equal to the rank r^*. Thus

$$No.\ \Lambda_r^* \ = \ r^* \tag{7.2.1}$$

where $No.\ \Lambda$ is the number of positive eigenvalues in Λ. A similar equation to (7.2.1) is used to calculate an estimated rank $r^{(k)}$ given by

$$No.\ \Lambda_r^{(k)} \ = \ r^{(k)}.$$

where Λ_r is given by (5.2.4). The range of error is relatively small. Then the l_1 SQP method will be applied to solve the problem as described in Section 6.4.

Another consideration is τ how to be chosen, if τ is close to the boundary of the condition (6.3.4) then the equation

$$No.\ \Lambda_r^{(k)} \ = \ r^*$$

may satisfied in the first few iterations. Experiments proved this fact see Table 6.5.1.

The projection–l_1 SQP algorithm can be described as follows.

Algorithm 7.2.1

Given any positive definite matrix $F = F^T \in \Re^{n \times n}$, let s be a positive integer. Then the following algorithm solves the educational testing problem

i. Let $F^{(0)} = F$

ii. Choose τ to be close to the boundary of the condition (6.3.3).

iii. Apply Algorithm 6.3.1 until

$$No.\ \Lambda_r^{(k)} = No.\ \Lambda_r^{(k+j)} \quad j = 1, 2, \ \ldots, s \tag{7.2.2}$$

iv. $r^{(k)} = No.\ \Lambda_r^{(k)}$

v. Use the result vector $\mathbf{x}$ from Algorithm 6.3.1 as an initial vector for l_1 SQP method

vi. Apply l_1 SQP method to solve the problem with $r = r^{(k)}$.

$$If$$

$$\|D_2(\mathbf{x})\| \le \epsilon \ \ for\ some\ small\ \epsilon$$

$$Then$$

$$F^* = F^{(k)}, \ \ r^* = r^{(k)} \ \ and \ \ terminate$$

$$Endif$$

vii. Apply one inner iteration of the Algorithm 6.3.1.

viii. Go to (iv).

The integer s in Algorithm 7.2.1 can be any positive number. If it is small then the rank $r^{(k)}$ may not be accurately estimated, however the number of iterations taken by projection method is small. In the other hand if s is large then a more accurate rank is obtained but the projection method needs more iterations.

The advantage of using the projection method as the first stage of the projection–l_1 SQP method is that if $F^{(0)}$ is positive semi–definite (singular) then the projection method terminates at the first iteration. Moreover it gives the best estimate to $r^{(k)}$.

Another way of estimating the rank $r^{(k)}$ is suggested by Fletcher [1985] and it was given in the end of Section 5.4, equation (5.4.3).

7.3 l_1SQP–Projection method

Starting with projection method has the advantage of knowing if the given matrix is a positive semi–definite (singular) or not, and it gives the best estimate for the matrix rank $r^{(k)}$. However sometimes it takes many iterations before equation (7.2.2) is satisfied, especially if τ is chosen to be small, this means slow convergence since the projection method is slow converges method. In this method an algorithm starts with the l_1SQP method with an estimated rank $r^{(k)}$ is considered. Then one iteration of the projection method will be calculated after every stage of the l_1SQP–projection algorithm the resulting vector $\mathbf{x}^{(k)}$ will be used as an initial vector to the next stage, thus the vector $\mathbf{x}^{(k)}$ is updated at every stage from the previous one.

The method in this section follows a similar strategy as that in Section 4.4.

Now the l_1SQP–projection algorithm can be described as follows.

Algorithm 7.3.1

Given any positive definite matrix $F = F^T \in \Re^{n\times n}$ the following algorithm solves the educational testing problem

i. Let $F^{(0)} = F$

ii. Choose $r^{(k)}$ (small as possible based on one of Section 7.2 strategies).

iii. Apply l_1 SQP method if $\|D_2(\mathbf{x})\| \leq \epsilon$ for some small ϵ, terminates.

iv. Use the result $\mathbf{x}^{(k)}$ as an initial vector for projection method (Algorithm 6.3.1).

v. Choose τ to be close to the boundary of the condition (6.3.3), $(\tau = \sum x_i^{(k)})$.

vi. Apply one iteration of the projection method.

vii. $r^{(k)} = No.\ \Lambda_r^{(k)}$.

viii. Use the result $\mathbf{x}^{(k)}$ as an initial vector for l_1 SQP method.

ix. Go to (iii).

Another advantage of this algorithm is that if the rank is not correct then instead of adding one to $r^{(k)}$ it goes back to the projection method to provide a better estimate to $r^{(k)}$. This will increase or decrease $r^{(k)}$ nearer to r^*, therefore variables will be added to or subtracted from the problem. The new variables are estimated using the projection method.

Another advantage is that at every stage only one iteration of projection method is used giving a faster converging algorithm.

Example 7.3.2

An example of this algorithm for $n = 5$

$$F = F^{(0)} = \begin{bmatrix} 0 & 5 & 4 & 3 & 1 \\ 5 & 0 & 6 & 3 & 3 \\ 4 & 6 & 0 & 6 & 4 \\ 3 & 3 & 6 & 0 & 5 \\ 1 & 3 & 4 & 5 & 0 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} 10 \\ 10 \\ 10 \\ 10 \\ 10 \end{bmatrix}.$$

If we use projection method to estimate $r^{(0)}$ as in the previous section with $\tau = -100$ then $r^{(0)} = 1$, apply the l_1SQP algorithm then it terminates with

$$\mathbf{x} = [10/3 \quad 7.5 \quad 4.8 \quad 2.7 \quad 0.3], \qquad \sum_{i=1}^{5} x_i = 18.6333$$

$\sum_{i=1}^{5} x_i = 18.6333$ and $D(\mathbf{x}) \ncong 0$. Applying one iteration from the projection method with $\tau = 18$ we find that $r^{(k)} = 3$ and

$$\mathbf{x} = [3.4671 \quad 7.5652 \quad 6.1089 \quad 4.5908 \quad 2.5492].$$

Apply this as an initial vector for l_1 SQP algorithm, after 15 iterations the l_1SQP algorithm terminates with $D(\mathbf{x}) \cong 0$,

$$\mathbf{x} = [13/3 \quad 9.0 \quad 6.0 \quad 9.0 \quad 13/3]$$

and $\sum \mathbf{x} = 32.6667$. If we use the initial $\tau = 18$ instead of $\tau = -100$ in the first stage of the projection method then $r^{(0)} = 3$ and the l_1SQP algorithm will terminate directly with the same result.

7.4 Numerical results and comparisons

In this section numerical problems are obtained from the data given in Table 6.2.1, by Woodhouse [1976]. The Woodhouse data set is a 64×20 data which corresponds to 64 students

Columns which determine F	τ	TNII	$r^{(0)}$	r^*	NQP	$\sum \theta_i^*$
1,2,5,6	400	4	3	3	11	542.77356
1,3,4,5	400	2	2	2	12	633.15784
1,2,3,6,8,10	600	11	4	5	8	305.48170
1,2,4,5,6,8	600	4	4	4	13	564.46331
1–6	600	6	4	4	10	535.36227
1–8	800	13	5	6	14	641.83848
1–10	1000	15	7	8	21	690.78040
1–12	1200	23	9	9	9	747.48921
1–14	1400	25	10	12	34	671.27506
1–16	1600	22	11	14	44	663.46204
1–18	1800	20	12	15	27	747.50574
1–20	2000	29	14	18	39	820.34265

Table 7.4.1: Results for the educational testing problem from the projection–l_1SQP method of Section 7.2.

and 20 subtests. Various selections from the set of subsets of columns are used to give various test problems to form the matrix F. These subsets are those given in the first columns of Tables 7.4.1–3, the value of n is the number of elements in each subset. Equation (6.2.2) gives the formula for calculating the educational testing problems from Table 6.2.1.

The result obtained by the new method of Section 7.2 are tabulated in Table 7.4.1. In Table 7.4.1 the columns headed by NQP give the number of times that the major l_1SQP is solved.

In the projection–l_1SQP method τ needs to be estimated very close to $\sum x_i^*$, this will give us a very good estimate of the rank. Since the average size of the educational testing problem elements are more than 100, $\tau = n \times 100$ is chosen as an initial value (see Section 6.5). In Table 7.4.1 it is clear that when $n > 10$ then τ becomes very small comparing with $\sum x_i^*$ which makes the projection method estimate $r^{(k)}$ very small comparing with the correct r^*.

The result obtained by the new method of Section 7.3 are tabulated in Table 7.4.2. In the l_1SQP–projection method $r^{(k)}$ updated using one iteration of the projection method. In the projection method τ estimated using the result from the l_1SQP method. In the 1–10 case the projection method estimated $r^{(k)} = 10$ instead of $r^{(k)} = 9$.

In both Tables 7.4.1 and 7.4.2 it can be seen that the results we have are exactly the same as Fletcher [1985]. Also one or two of the variables are adjusted so that the matrix $F - diag\ \boldsymbol{\theta}$ is exactly singular and positive semi–definite.

Finally in Table 7.4.3 the four methods are compared.

Columns which determine F	$r^{(0)}$	NQP	PM$r^{(k)}$	NQP	$\sum \theta_i^*$
1,2,5,6	2	5	3	6	542.77356
1,3,4,5	2	12			633.15784
1,2,3,6,8,10	3	4	5	5	305.48170
1,2,4,5,6,8	3	6	4	4	564.46331
1–6	3	7	4	4	535.36227
1–8	5	7	6	6	641.83848
1–10	6	9	8	11	690.78040
1–12	8	3	10	9	747.48921
1–14	10	6	12	9	671.27506
1–16	11	9	14	10	663.46204
1–18	13	7	15	16	747.50574
1–20	15	5	18	21	820.34265

Table 7.4.2: Results for the educational testing problem from the l_1SQP–projection method of Section 7.3.

PM$r^{(k)}$:rank r updated from the projection method.

Columns which		PM	l_1SQP		Pl_1SQP			l_1SQPP	
determine F	r^*	TNII	$r^{(0)}$	NQP	TNII	$r^{(0)}$	NQP	$r^{(0)}$	TNQP
1,2,5,6	3	197	2	14	4	3	11	2	11
1,3,4,5	2	224	2	12	2	2	12	2	12
1,2,3,6,8,10	5	580	3	9	11	4	8	3	9
1,2,4,5,6,8	4	4994	3	13	4	4	13	3	10
1–6	4	1351	3	14	6	4	10	3	11
1–8	6	1948	5	29	13	5	14	5	13
1–10	8	2918	6	34	15	7	21	6	20
1–12	9	2403	8	29	23	9	9	8	12
1–14	12	3196	10	36	25	10	34	10	15
1–16	14	5215	11	42	22	11	44	11	19
1–18	15	14043	13	27	20	12	27	13	23
1–20	18	8255	15	39	29	14	39	15	26

Table 7.4.3: Comparing the four methods.

Pl_1SQP: the projection–l_1SQP method.
l_1SQPP: the l_1SQP–projection method.
TNQP : total number of NQP.