SEMI-DEFINITE PROGRAMMING USING `OPTIMIZE()` IN R

JAN DE LEEUW

ABSTRACT. Meet the abstract. This is the abstract.

1. PROBLEM

The problem is to minimize a function of the form $g(\theta) = f(C(\theta))$ where $C(\theta) = G_0 + \sum_{s=1}^{p} \theta_s G_s$, under the condition that $C(\theta)$ is positive semi-definite. Here $G_0, \cdots, G_p$ are known symmetric matrices.

Many problems in multivariate analysis, such as fitting mixed linear models or fitting minimum trace and minimum rank factor analysis models, are of this particular form.

2. RELAXATION ALONG A LINE

Suppose $C(\theta)$ is the current positive semi-definite solution in an iterative process. We improve $C(\theta)$ by changing one coordinate at a time.

Consider the line $C_s(\epsilon) = C(\theta) + \epsilon G_s$. To simplify, suppose $C(\theta)$ is positive definite. Define λ_i to be the eigenvalues of $C^{-1}(\theta)G_s$.

Then $C_s(\epsilon)$ is positive semi-definite if and only if $1 + \epsilon\lambda_i \geq 0$ for all i. Or, equivalently, ϵ must satisfy

$$-\frac{1}{\max\{\lambda_i \mid \lambda_i > 0\}} \leq \epsilon \leq -\frac{1}{\min\{\lambda_i \mid \lambda_i < 0\}}. \tag{1}$$

Date: Monday 12th April, 2010 — 13h 10min — Typeset in LUCIDA BRIGHT.
Key words and phrases. Template, LaTeX.

If λ_+ and λ_- are the largest and smallest eigenvalues then, equivalently,

$$-\frac{1}{\lambda_+} \leq \epsilon \leq +\infty \text{ if } \lambda_+ > 0 \text{ and } \lambda_- \geq 0 \tag{2a}$$

$$-\frac{1}{\lambda_+} \leq \epsilon \leq -\frac{1}{\lambda_-} \text{ if } \lambda_+ > 0 \text{ and } \lambda_- < 0 \tag{2b}$$

$$-\infty \leq \epsilon \leq -\frac{1}{\lambda_-} \text{ if } \lambda_+ \leq 0 \text{ and } \lambda_- < 0. \tag{2c}$$

Of course the signs of the largest and smallest eigenvalues are the same as the signs of the largest and smallest eigenvalues of G_s, i.e. they do not change during the iterations.

If $C(\theta)$ is singular we have to be more careful. Suppose $K_\perp$ is an orthonormal basis for its null space, and suppose $K_\perp' G_s K_\perp$ has both positive and negative eigenvalues. Then $C(\theta + \epsilon e_s)$ is positive semi-definite if and only if $\epsilon = 0$.

The idea of the algorithm is to use the `optimize()` function in R to find the optimum ϵ in the feasible interval (2a). We go through all coordinates to complete a single coordinate descent iteration cycle, and we compute cycles until convergence. A trivial extension is the case in which there are in addition bound constrains on the θ_s. Just intersect the interval of the bound constraints with the interval defined by (2a).

3. Examples

3.1. **Minimum Trace Factor Analysis (MTFA).** Suppose R is a correlation matrix. Define $G_0 = R - I$ and $G_s = e_s e_s'$, where the e_s are unit vectors. Thus $C(\theta)$ is R with its diagonal replaced by the elements of θ. We impose bound constraints $0 \leq \theta_s \leq 1$. In this case the G_s are positive semidefinite, or rank one. The largest eigenvalue of $C^{-1}(\theta)G_s$ is $e_s' C^{-1}(\theta) e_s$. Thus the interval over which we minimize is

$$-\max(\theta_s, \frac{1}{e_s'C^{-1}(\theta)e_s}) \leq \epsilon \leq 1 - \theta_s.$$

In MTFA we minimize $g(\theta) = \mathbf{tr}\ C(\theta)$. Obviously we do not need `optimize()` in this case, because $g(\theta + \epsilon e_s) = g(\theta) + \epsilon$, and thus the optimum is attained at

$$\hat{\epsilon} = -\max(\theta_s, \frac{1}{e_s'C^{-1}(\theta)e_s}).$$

It is of some interest that, by the Lagrange theorem, the rank of $C(\theta + \hat{\epsilon}e_s)$ is less than the rank of $C(\theta)$, and thus $C(\theta + \hat{\epsilon}e_s)$ is always singular. The vector in the null space is $C^{-1}(\theta)e_s$. The result in the appendix that, because $e_s e_s' \gtrsim 0$, we must have $\epsilon \geq 0$, and thus

3.2. **MLM.** The (profile, concentrated) likelihood function for a mixed linear model is

$$n \log \min_{\beta}(y - X\beta)'(Z\Theta Z' + I)^{-1}(y - X\beta) + \log \mathbf{det}(Z\Theta Z' + I)$$

Appendix A. Positive Semi-definite Perturbations

Suppose A and B are symmetric matrices, with $A \gtrsim 0$. We want to determine the interval of all ϵ such that $A + \epsilon B \gtrsim 0$.

Let $K_\perp$ be an orthonormal basis for the null-space of A and K be a set of eigenvectors corresponding to the non-null space. Write B in the form

$$B = \begin{bmatrix} K & K_\perp \end{bmatrix} \begin{bmatrix} B_{11} & B_{10} \\ B_{01} & B_{00} \end{bmatrix} \begin{bmatrix} K' \\ K'_\perp \end{bmatrix}$$

Also suppose Λ^2 is the diagonal matrix of the non-zero eigenvalues of A. Then $A + \epsilon B$ is positive semi-definite if and only if

$$C(\epsilon) \triangleq \begin{bmatrix} I + \epsilon\Lambda^{-1}B_{11}\Lambda^{-1} & \epsilon\Lambda^{-1}B_{10} \\ \epsilon B_{01}\Lambda^{-1} & \epsilon B_{00} \end{bmatrix}$$

is positive semi-definite. Now apply Albert [1969, Theorem 1]. This says that $C(\epsilon) \gtrsim 0$, with $\epsilon \neq 0$, if and only if

$$\epsilon B_{00} \gtrsim 0, \tag{3a}$$

$$B_{00}B_{00}^{+}B_{01} = B_{01}, \tag{3b}$$

$$I + \epsilon\Lambda^{-1}[B_{11} - B_{10}B_{00}^{+}B_{01}]\Lambda^{-1} \gtrsim 0. \tag{3c}$$

It follows that there exists an $\epsilon \neq 0$ such that $A + \epsilon B \gtrsim 0$ if and only if we have (3b) and B_{00} is either positive semi-definite or negative semi-definite. Bounds on ϵ in that case are given by (3c).

References

A. Albert. Conditions for Positive and Nonnegative Definiteness in Terms of Pseudoinverses. *SIAM Journal on Applied Mathematics*, 17:434–440, 1969.

Department of Statistics, University of California, Los Angeles, CA 90095-1554

E-mail address, Jan de Leeuw: `deleeuw@stat.ucla.edu`

URL, Jan de Leeuw: `http://gifi.stat.ucla.edu`