Smacof Meets Chebyshev

Jan de Leeuw - University of California Los Angeles

Started October 08 2023, Version of October 09, 2023

Contents

Note: This is a working paper which will be expanded/updated frequently. All suggestions for improvement are welcome. The Rmd file, the pdf, all R files, a LaTeX version and so on are available at https://github.com/deleeuw/sChebyshev.

Notation

Conventions

Since we only work in finite dimensional vector spaces, and since our emphasis is on computation, we adopt the following conventions.

- A vector *is* a matrix with one column.
- A row-vector *is* a matrix with one row.
- Derivatives *are* matrices.

Notations

The length of vectors and the dimension of matrices will generally be clear from the context.

- e_i unit vector (element i is one, other elements zero).
- e vector with all elements one.
- E matrix with all elements one.
- 0 real number zero, also vector or matrix with all elements 0.
- I identity matrix.

- $J = I - \frac{ee'}{e'e}$ centering matrix.
- $A \otimes B$ Kronecker product of matrices A and B.
- $X \oplus Y$ direct sum of matrices X and Y.
- $X \times Y$ elementwise (Hadamard) product of matrices X and Y.
- vec(X) matrix X to vector (columns on top of each other).
- vecr(X) elements below diagonal of matrix X to vector (columns on top of each other).
- X' transpose of matrix X.
- X^+ Moore-Penrose inverse of matrix X.
- X^{-T} inverse of the transpose X' (and transpose of the inverse).
- $X \gtrsim Y$ Loewner order of symmetric matrices ($X - Y$ is positive semi-definite).
- $X \lesssim Y$ Loewner order of symmetric matrices ($Y - X$ is positive semi-definite).
- $:=$ definition.
- $X \times Y$ Cartesian product of sets X and Y.
- (x, y) is an ordered pair, i.e. an element of $X \times Y$.
- x_{is} or $\{X\}_{is}$ element (i, s) of matrix X.
- $a_{\bullet s}$ column s of matrix A.
- $a_{i\bullet}$ row i of matrix A.
- $[A]_{is}$ submatrix (i, s) of block-matrix A.
- $A^{[p]}$ direct sum of p copies of matrix A.
- $a^{(k)}$ the k^{th} element of the sequence $\{a\} = a^{(1)}, \cdots, a^{(k)}, \cdots$.
- $\mathbb{R}^n$ space of all vectors of length n.
- $\overline{\mathbb{R}}^n$ space of all centered vectors of length n (i.e. $x'e = 0$).
- $\mathbb{R}^{n\times p}$ space of all $n \times p$ matrices.
- $\overline{\mathbb{R}}^{n\times p}$ space of all column-centered $n \times p$ matrices (i.e. with $X'e = 0$).
- $f : X \Rightarrow Y$ function with arguments in X and values in Y.
- If $f : X \times Y \Rightarrow Z$ then $f(\bullet, y) : X \Rightarrow Z$ and $f(x, \bullet) : Y \Rightarrow Z$.
- $x'y$ inner product in $\mathbb{R}^n$.
- tr $X'Y$ inner product in $\mathbb{R}^{n\times p}$.
- $\|x\| = \sqrt{x'x}$ Euclidean norm of $x \in \mathbb{R}^n$.
- $\|X\| = \sqrt{\text{tr } X'X}$ Euclidean norm of $X \in \mathbb{R}^{n\times p}$.
- $\mathcal{D}f(x)$ derivative of f at x.
- $\mathcal{D}^2 f(x)$ second derivative of f at x.
- $\mathcal{D}_s f(x) = \{\mathcal{D}f(x)\}_s$ partial derivative with respect to x_s at x.
- $\mathcal{D}_{st} f(x) = \{\mathcal{D}^2 f(x)\}_{st}$ second partial with respect to x_s and x_t at x.

1 Introduction

In this paper we study techniques to speed up the basic smacof iteration

$$X^{(k+1)} = \Gamma(X^{(k)})$$

with Γ the Guttman transform, by using updates of the form

$$X^{(k+1)} = \sum_{r=1}^{s} \alpha_r \Gamma^r(X^{(k)})$$

with $\Gamma^0(X) = X$ and $\Gamma^r(X) = \Gamma(\Gamma^{r-1}(X))$

1.1 Matrix Basis

An important special case of DCDD imposes the constraint

$$X = \sum_{v=1}^{r} \theta_v G_v, \tag{1}$$

where the G_s are $n \times p$ matrices. To see that this is indeed a special case of DCDD define Y_s as the matrix collecting the s^{th} columns of all G_v. Thus there are r of these $n \times r$ matrices Y_s. Now $\vec{(}X) = Y\theta$, with Y the direct sum of the Y_s, as usual, and

$$\theta = \left.\begin{bmatrix}\theta\\ \vdots \\ \theta\end{bmatrix}\right\} r \text{ times.} \tag{2}$$

The distinguishing DCDD characteristic in using the *matrix basis* (1) is that all r pieces of θ in (2) must be equal.

Better in configuration space

No $V = I$

`One important application of the matrix basis is finding the optimal step size in`

iterative procedures, or, more generally, finding optimal weights in multistep procedures. For the steepest descent method, for example, we choose $G_1 = X$ and $G_2 = \nabla\sigma(X)$.

$$\{C_{ij}\}_{vw} = \text{tr } G_v' A_{ij} G_w$$

Thus

$$\{V_s\}_{vw} = \text{tr } G_v' V_0 G_w$$

$$\{B_s(\theta)\}_{vw} = \text{tr } G_v' B_0(\theta) G_w$$

$$\sigma(\theta) := \frac{1}{2}\{1 - 2\theta' B(\theta)\theta + \theta' V\theta\}$$

1.2 Matrix Based

$$X^{(k+1)} = \theta_1 X^{(k)} + \theta_2 \Gamma(X^{(k)}) + \theta_3 \Gamma^2(X^{(k)}) + \cdots + \theta_r \Gamma^{r-1}(X^{(k)})$$

Problem: near a fixed point B and V become almost singular (of rank one)

Chebyshev:

$$\min_\theta \max_s |\theta_1 + \theta_2\lambda_s + \cdots + \theta_r \lambda_s^{r-1}|$$

`Sidi`

Generalizes relax.