Matrix Exponential

Moo K. Chung

University of Wisconsin-Madison, USA

mkchung@wisc.edu

Abstract. The computation of the exponent of symmetric matrices, often called the exponential map, is often needed in medical imaging for advanced geometric data analysis. Given a point in the manifold (space of positive definite symmetric matrices), we show how to estimate the point using basis in its tangent space (space of symmetric matrices) through the exponential map.

Problem Statement

Consider $p \times p$ symmetric matrix A such as correlation and covariance matrices obtained from functional magnetic resonance imaging (fMRI). The matrix A is theoretically assumed to be positive definite symmetric (PDS) but for various reasons, the observed matrix A may be nonnegative definite with multiple zero eigenvalues. The question is how to make it PDS. The approach here is based on the exponential map technique that project the scatter points in the tangent space to scatter points to the underlying manifold (Huang et al. 2020).

Matrix Exponential

Given a symmetric matrix A, the exponential of A, denoted as e^A, is defined analogously to the scalar exponential function, using its Taylor series expansion:

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!},$$

where A^k is the k-th power of A. The matrix exponential plays a key role in solving systems of linear ordinary differential equations, where solutions are often expressed as $\mathbf{x}(t) = e^{At}\mathbf{x}_0$ for initial condition $\mathbf{x}_0$. Computing e^A efficiently for large matrices is a challenging but often needed in medical imaging. For small matrices, one can directly use the series expansion, but for larger matrices, we proceed as follows.

Symmetric matrix A can be decomposed as $A = PDP^{-1}$, where D is a diagonal matrix of eigenvalues of A and P is the matrix of eigenvectors, which forms an orthogonal matrix, then the exponential of A can be computed as

$$e^A = Pe^D P^{-1}.$$

Here, e^D is simply a diagonal matrix with entries $e^{\lambda_1}, e^{\lambda_2}, \ldots, e^{\lambda_n}$, where λ_i are the eigenvalues of A. This method is computationally efficient for diagonalizable matrices, as the diagonal form simplifies the computation of the exponential.

Nonnegative Least Squares

Nonnegative Least Squares (NNLS) is a specialized type of least squares optimization problem where the solution is constrained to be nonnegative. Given a system of linear equations $A\mathbf{x} = \mathbf{b}$, where A is an $m \times n$ matrix, $\mathbf{b}$ is an $m \times 1$ vector of observations, and $\mathbf{x}$ is the $n \times 1$ vector of unknowns, the goal of NNLS is to find $\mathbf{x}$ that minimizes the residual error

$$\min_{\mathbf{x} \geq 0} \|A\mathbf{x} - \mathbf{b}\|^2,$$

subject to the constraint that $\mathbf{x} \geq 0$, meaning every component of $\mathbf{x}$ must be nonnegative.

The NNLS problem is particularly useful in applications where negative values for the solution are physically or practically meaningless. For example, in medical image reconstruction, variables such as intensities, concentrations, or costs are inherently nonnegative.

The method for solving NNLS often involves iterative algorithms, as closed-form solutions do not generally exist due to the inequality constraints. One popular algorithm for NNLS is the active set method, which iteratively adjusts a set of active constraints (i.e., constraints that are treated as equalities) and solves the resulting unconstrained least squares problem in each step. Another commonly used method is the projected gradient descent, where the solution is updated iteratively while ensuring that the nonnegativity constraint is enforced at every step by projecting negative values of $\mathbf{x}$ to zero. In Maltab, it is computed using `lsqnonneg.m`.

Space of Positive Definite Symmetric Matrices

Let Sym_p be the space of all $p \times p$ symmetric matrices with inner product $\langle A, B \rangle = \mathrm{tr}(AB)$. Such space forms a vector space. The space of $p \times p$ symmetric positive-definite (SPD) matrices, denoted by Sym_p^+, is a **subset** of Sym_p but no longer forms a vector space (Arsigny et al. 2007). Given $X, Y \in Sym_p^+$, any positive sum $\alpha X + \beta Y \in Sym_p^+$ for $\alpha, \beta > 0$. Thus, Sym_p^+ is a curved convex manifold. **For $C_1, C_2 \in Sym_p^+$, we have different tangent planes $T_{C_1}(Sym_p^+)$ and $\mathbf{T}_{C_2}(Sym_p^+)$, which are all subspaces of Sym_p. Figure 1 provides a schematic of relationship between Sym_p and Sym_p^+.**

The matrix exponential of a symmetric matrix always results in a SPD matrix. Thus, the exponential map serves as a transformation from the tangent space to the manifold of SPD matrices. Moreover, the exponential map is one-to-one between Sym_p and Sym_p^+. We can further define the logarithm (inverse

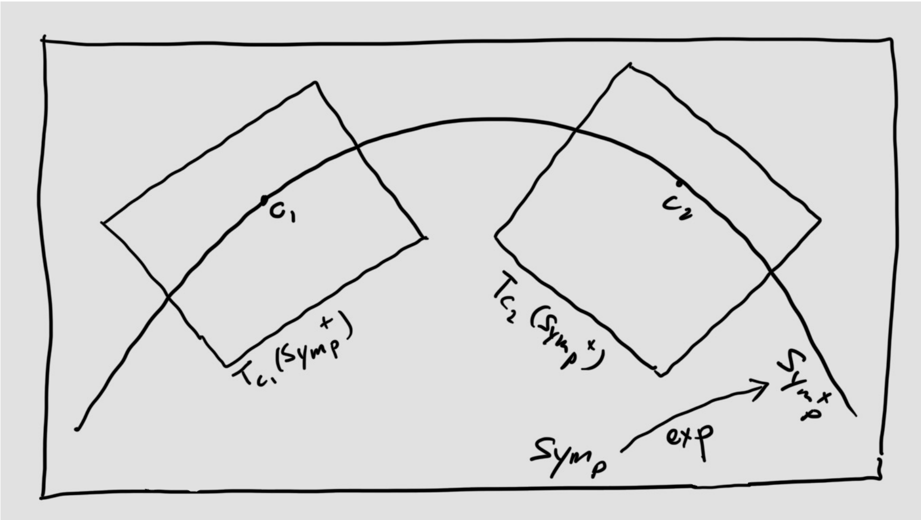

Fig. 1: Sym_p^+ **is a smooth manifold embedded in the larger vector space of symmetric matrices,** Sym_p**. For any** $C_i \in Sym^+p$**, its tangent space** $T_{C_i}(Sym_p^+)$ **is a subspace of** Sym_p**. The matrix exponential defines a smooth, one-to-one mapping** $\exp : Sym_p \to Sym_p^+$**, ensuring that the image of any symmetric matrix under this mapping remains in** Sym_p^+**.**

map) of an SPD matrix is a symmetric matrix. Given $X \in Sym_p$, its exponential map $X \to e^X$ is defined through matrix exponential

$$e^X = \sum_{n=0}^{\infty} \frac{X^n}{n!}.$$

The exponential map may introduce too much error between X and e^X:

$$X - e^X = I - \frac{X^2}{2} - \frac{X^3}{6} - \cdots.$$

Thus, we use the basis expansion to increase the overall fit as follows. The goal is to find basis $I_{ij} \in Sym_p$ such that its expansion

$$\|X - \sum_{i \geq j} c_{ij} e^{I_{ij}}\|$$

is minimized.

Let I_{ij} be the $p \times p$ matrix whose (i,j)-th and (j,i)-th entries are $1/\sqrt{2}$ if $i \neq j$ and all other entries are 0. Let I_{ii} be the $p \times p$ diagonal matrix whose

(i,i)-th entry is 1 and all other entries are 0. For instance, for $p=3$,

$$I_{31} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \text{ and } I_{22} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{1}$$

Then, we have

$$\langle I_{ij}, I_{kl} \rangle = \delta_{ik}\delta_{jl}.$$

and $\{I_{ij}, i \geq j\}$ forms an orthonormal basis in Sym_p (Huang et al. 2020). Then we expand any $X \in Sym_p^+$ as

$$X = \sum_{i \geq j} c_{ij} exp(I_{ij}),$$

where $c_{ij} \geq 0$ can be estimated using the nonnegative least squares method (Esser et al. 2013).

Numerical implementation

The matrix exponential is computed more efficiently through singular value decomposition:

$$X = UDU^\top$$

where D is the diagonal matrix with diagonal entries d_i. Then, the matrix exponential is computed as

$$e^X = Ue^DU^\top,$$

where e^D is the diagonal matrix with diagonal entries given by e^{d_i}. For instance, the matrix exponentials of I_{31} and I_{22} in (1) are

$$e^{I_{31}} = \begin{pmatrix} 1.2606 & 0 & 0.7675 \\ 0 & 1 & 0 \\ 0.7675 & 0 & 1.2606 \end{pmatrix} \text{ and } e^{I_{22}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In MATLAB, the nonnegative least squares estimation is performed using `lsqnonneg.m` by vectorizing the upper triangle matrix of symmetric matrices. For instance, consider the following nonnegative definite matrix: $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 2 \\ 3 & 2 & 1 \end{pmatrix}$ It has eigenvalues $(0, 2, 6)$. This is estimated in Sym_3^+ as $\begin{pmatrix} 1.0205 & 1.8973 & 3.0000 \\ 1.8973 & 1.8973 & 2.0000 \\ 3.0000 & 2.0000 & 1.2865 \end{pmatrix}$. It has eigenvalues $(0.0477, 1.8495, 6.0061)$ making it positive definite.

Bibliography

Arsigny, V., Fillard, P., Pennec, X. & Ayache, N. (2007), 'Geometric means in a novel vector space structure on symmetric positive-definite matrices', *SIAM journal on matrix analysis and applications* **29**, 328–347.

Esser, E., Lou, Y. & Xin, J. (2013), 'A method for finding structured sparse solutions to nonnegative least squares problems with applications', *SIAM Journal on Imaging Sciences* **6**, 2010–2046.

Huang, S.-G., Samdin, S.-T., Ting, C., Ombao, H. & Chung, M. (2020), 'Statistical model for dynamically-changing correlation matrices with application to brain connectivity', *Journal of Neuroscience Methods* **331**, 108480.