Connectomics: Null Hypotheses

Horacio Gómez-Acevedo

November 1, 2022

UAMS

Hypothesis Testing

There are key components to take full advantage of th

- Likelihood
- Statistical Hypothesis
- Hypothesis Test
- Likelihood Ratio
- Asymptotic Properties of the Normal Distribution
- p-values

Likelihood

If we have $X_1, \ldots, X_n$ independent and identically distributed (iid) random variables with a common probability (mass/density) function $f(x;\theta)$ where the parameter θ is unknown ($\theta \in \Omega$). The likelihood of a sample $\vec{x} = (x_1, \ldots, x_n)$ is

$$L(\theta, \vec{x}) = \prod_{i=1}^{n} f(x_i, \theta)$$

Example: If we have $X_i \sim N(\theta, \sigma^2)$ with $\sigma^2 > 0$ known but θ unknown. Then

$$L(\theta, \vec{x}) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \overline{x})^2\right) \exp\left(-\frac{1}{2\sigma^2} n(\overline{x} - \theta)^2\right)$$

Statistical Hypothesis

A Statistical Hypothesis is a conjecture about the probability distribution of a population.

Example: We suppose that in an experiment we have a random sample from $N(\theta, 10)$.

$$H_0: \text{ The population is } N(5, 10)\text{-distributed}$$

$$H_1: \text{ The population is } N(1, 10)\text{-distributed}$$

Hypothesis Test

A Hypothesis Test is a tuple $(X_1, \ldots, X_n; H_0, H_1, G)$, where

1. $(x_1, \ldots, x_n)$ is a sample of $(X_1, \ldots, X_n)$ random variables iid.
2. H_0 and H_1 are hypothesis concerning the probability distribution of the population.
3. $G \subset \mathbb{R}^n$ is Borel set (countable unions of open sets).

The **level of significance** is defined as

$$\alpha = P^{H_0}_{X_1,\ldots,X_n}(G)$$

We will consider hypothesis such as

$$H_0 : \theta = \theta_0(\text{ or } \theta \in \Theta_0) \qquad H_1 : \theta \neq \theta_0(\text{ or } \theta \in \Theta = \Theta_1 \cup \Theta_0)$$

Maximum Likelihood Test

The **likelihood ratio** function

$$\Lambda(x_1, \ldots, x_n) = \frac{\sup_{\theta \in \Theta_0} L_\theta(x_1, \ldots, x_n)}{\sup_{\theta \in \Theta} L_\theta(x_1, \ldots, x_n)}$$

Let $\widehat{\theta}$ be the maximum likelihood estimate of θ.
If θ_0 is the true value of θ, then $L(\theta_0)$ is the maximum value of $L(\theta)$.
Since $\Lambda \leq 1$, then if H_0 is true Λ should be close to 1, whereas if H_1 is true then Λ should be smaller.
We have the decision rule

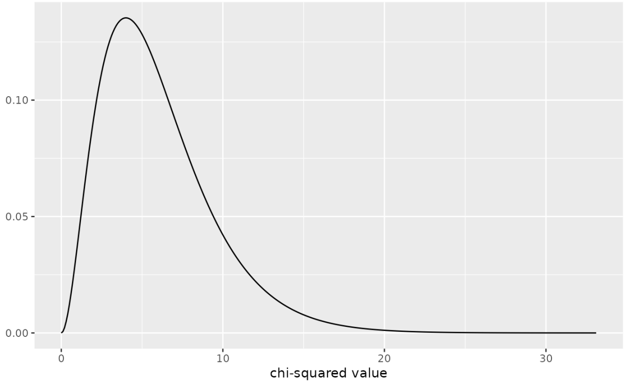
$$\text{Reject } H_0 \text{ in favor of } H_1 \text{ if } \Lambda \leq c,$$

where $\alpha = P^{\theta_0}(\Lambda \leq c)$.

Asymptotic Properties

Under some (regularity)conditions we have the following result:
If the null hypothesis $H_0 : \theta = \theta_0$.

$$-2 \log \Lambda(X_1, \ldots, X_n) \to \chi^2(1)$$



Asymptotics for the Normal distribution

When μ and σ are unknown and testing the hypothesis

$$H_0\colon \mu = \mu_0 \quad H_1\colon \mu \neq \mu_0$$

The likelihood ratio is given by

$$\Lambda(x_1, \ldots, x_n) = \left\{1 + \frac{1}{n-1}\left(\frac{\overline{x} - \mu_0}{s/\sqrt{n}}\right)\right\}^{-n/2}$$

where $s^2 = \frac{1}{n-1}\sum_i (x_i - \overline{x})^2$. The critical regions are of the form

$$G = \left\{(x_1, \ldots, x_n) \in \mathbb{R}^n : \left|\frac{\overline{x} - \mu_0}{s/\sqrt{n}}\right| \geq c\right\}$$

p-values

The test statistic $\frac{\overline{x}-\mu_0}{s/\sqrt{n}}$ is critical to reject H_0 or not. The decision procedure is as follows

$$\text{if } \left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right| \geq c \text{ then we assume } H_1$$
$$\text{if } \left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right| < c \text{ then we assume } H_0$$

Furthermore, we have the following equivalence

$$P\left(\left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right| \geq u\right) \leq \alpha \iff u \geq c$$

The p-value associated with the outcome u of the test statistic $\frac{\overline{x}-\mu_0}{S/\sqrt{n}}$

$$P\left(\left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right| \geq |u|\right)$$

p-values (cont)

Thus if an outcome u of $\frac{\overline{x}-\mu_0}{S/\sqrt{n}}$ satisfies

$$P\left(\left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right| \geq |u|\right) \leq \alpha \text{ we accept } H_1$$

Today's paper

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

ELSEVIER

Computational Neuroscience

Extracting spatial–temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition

Bingni W. Brunton [a,b,c,*], Lise A. Johnson [d,e], Jeffrey G. Ojemann [e], J. Nathan Kutz [b]

[a] Department of Biology, University of Washington, Seattle, WA 98195, USA
[b] Department of Applied Mathematics, University of Washington, Seattle, WA 98195, USA
[c] Institute for Neuroengineering, University of Washington, Seattle, WA 98195, USA
[d] Center for Sensorimotor Neural Engineering, University of Washington, Seattle, WA 98195, USA
[e] Department of Neurological Surgery, University of Washington, Seattle, WA 98195, USA

Background

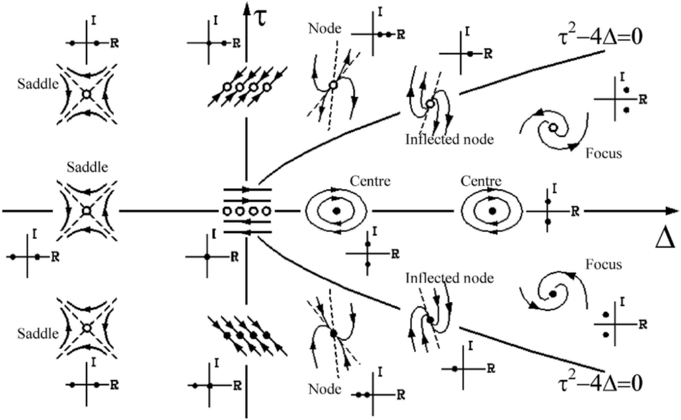
What are eigenvalues and eigenvectors?
Let A be a $n \times n$ matrix (real), a number $\lambda \in \mathbb{C}$ is an **eigenvalue** if there is a nonzero vector $x \in \mathbb{C}^n$ for which $Ax = \lambda x$. We call such a vector an **eigenvector** of A associated with λ.
The collection of all eigenvalues of A is called the **spectrum** of A.
The fundamental theorem for linear systems: If A is a $n \times n$ matrix. For any $x_0 \in \mathbb{R}^n$, the initial value problem

$$\dot{x} = Ax \qquad x(0) = x_0$$

has a unique solution

$$x(t) = \exp(A^t)x_0$$

Eigenvalues in Dynamical Systems



https://www.researchgate.net/profile/Marco-Altosole/publication/245387409/figure/fig1/AS:392836487892996@1470670925920/Types-of-phase-portrait.png

Singular Value Decomposition

If $A \in M_{m,n}$ has rank k, then it may be written in the form

$$A = V \Sigma W^*$$

where $V \in M_m$ and $W \in M_n$ are unitary $(W^* W = I_n)$. The matrix $\Sigma = diag\{\sigma_1, \ldots, \sigma_q\}$ are the non-negative square roots of the eigenvalues of AA^*. The columns of V are eigenvectors of AA^*, and the columns of W are eigenvectors of A^*A. If $m \leq n$ and if AA^* has distinct eigenvalues, then V is determined up to a right diagonal factor $D = diag(\exp(i\theta_1), \ldots, \exp(i\theta_n))$ where $\theta_j \in \mathbb{R}$.

Main Premise

The DMD algorithm seeks a best-fit linear matrix A that approximately advances the state of a system $x \in \mathbb{R}^n$ forward in time according to the linear system

$$x_{k+1} = Ax_k$$

where $x_k = x(k\Delta t)$ and Δt denotes a fixed time step that is small enough to resolve the highest frequencies in the dynamics.

DMD setup

$$\mathbf{X} = \begin{bmatrix} | & | & & | \\ \mathbf{x}(t_1) & \mathbf{x}(t_2) & \cdots & \mathbf{x}(t_m) \\ | & | & & | \end{bmatrix}, \tag{3.2a}$$

$$\mathbf{X}' = \begin{bmatrix} | & | & & | \\ \mathbf{x}(t_1') & \mathbf{x}(t_2') & \cdots & \mathbf{x}(t_m') \\ | & | & & | \end{bmatrix}. \tag{3.2b}$$

Equation (3.1) may be written in terms of these data matrices as

$$\mathbf{X}' \approx \mathbf{A}\mathbf{X}. \tag{3.3}$$

DMD trick

computing $\tilde{\mathbf{A}}$ in (3.4), we may project $\mathbf{A}$ onto the first r POD modes in $\mathbf{U}_r$ and approximate the pseudoinverse using the rank-r SVD approximation $\mathbf{X} \approx \mathbf{U}_r \mathbf{\Sigma}_r \mathbf{V}_r^*$:

$$\tilde{\mathbf{A}} = \mathbf{U}_r^* \mathbf{A} \mathbf{U}_r \tag{3.6a}$$

$$= \mathbf{U}_r^* \mathbf{X}' \mathbf{X}^\dagger \mathbf{U}_r \tag{3.6b}$$

$$= \mathbf{U}_r^* \mathbf{X}' \mathbf{V}_r \mathbf{\Sigma}_r^{-1} \mathbf{U}_r^* \mathbf{U}_r \tag{3.6c}$$

$$= \mathbf{U}_r^* \mathbf{X}' \mathbf{V}_r \mathbf{\Sigma}_r^{-1}. \tag{3.6d}$$

The leading spectral decomposition of $\mathbf{A}$ may be approximated from the spectral decomposition of the much smaller $\tilde{\mathbf{A}}$:

$$\tilde{\mathbf{A}} \mathbf{W} = \mathbf{W} \mathbf{\Lambda}. \tag{3.7}$$

The diagonal matrix $\mathbf{\Lambda}$ contains the *DMD eigenvalues*, which correspond to eigenvalues of the high-dimensional matrix $\mathbf{A}$. The columns of $\mathbf{W}$ are eigenvectors of $\tilde{\mathbf{A}}$ and provide a coordinate transformation that diagonalizes the matrix. These columns may be thought of as linear combinations of POD mode amplitudes that behave linearly with a single temporal pattern given by the corresponding eigenvalue λ.

The eigenvectors of $\mathbf{A}$ are the *DMD modes* $\mathbf{\Phi}$, and they are reconstructed using the eigenvectors $\mathbf{W}$ of the reduced system and the time-shifted data matrix $\mathbf{X}'$:

$$\mathbf{\Phi} = \mathbf{X}' \tilde{\mathbf{V}} \tilde{\mathbf{\Sigma}}^{-1} \mathbf{W}. \tag{3.8}$$

DMD expansion

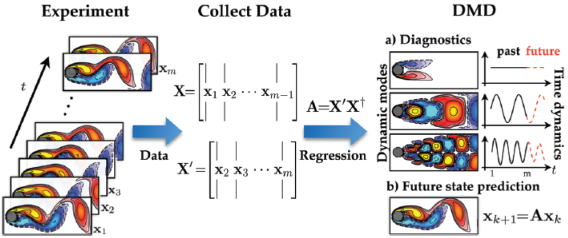


Fig. 3.1 *Overview of DMD illustrated on the fluid flow past a circular cylinder at Reynolds number* 100. *Reproduced with permission from Kutz et al.* [225].

3.1.1. Spectral Decomposition and the DMD Expansion. Once the DMD modes and eigenvalues are computed, it is possible to represent the system state in terms of the DMD expansion

$$\mathbf{x}_k = \sum_{j=1}^{r} \phi_j \lambda_j^{k-1} b_j = \mathbf{\Phi}\mathbf{\Lambda}^{k-1}\mathbf{b}, \tag{3.9}$$

where ϕ_j are eigenvectors of $\mathbf{A}$ (DMD modes), λ_j are eigenvalues of $\mathbf{A}$ (DMD eigenvalues), and b_j are the mode amplitudes. The DMD expansion (3.9) is directly analo-

DMD expansion (cont)

The spectral expansion in (3.9) may be converted to continuous time by introducing the continuous eigenvalues $\omega = \log(\lambda)/\Delta t$:

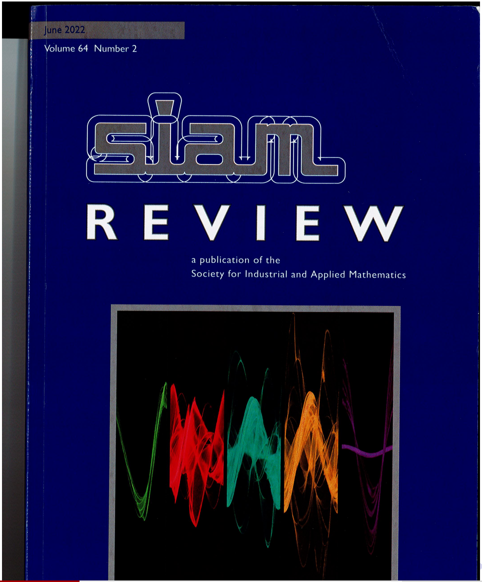
$$\mathbf{x}(t) = \sum_{j=1}^{r} \phi_j e^{\omega_j t} b_j = \mathbf{\Phi} \exp(\mathbf{\Omega} t)\mathbf{b}, \tag{3.14}$$

where $\mathbf{\Omega}$ is a diagonal matrix containing the continuous-time eigenvalues ω_j. Thus, the data matrix $\mathbf{X}$ may be represented as

(3.15)

$$\mathbf{X} \approx \begin{bmatrix} | & & | \\ \phi_1 & \cdots & \phi_r \\ | & & | \end{bmatrix} \begin{bmatrix} b_1 & & \\ & \ddots & \\ & & b_r \end{bmatrix} \begin{bmatrix} e^{\omega_1 t_1} & \cdots & e^{\omega_1 t_m} \\ \vdots & \ddots & \vdots \\ e^{\omega_r t_1} & \cdots & e^{\omega_r t_m} \end{bmatrix} = \mathbf{\Phi}\mathrm{diag}(\mathbf{b})\mathbf{T}(\boldsymbol{\omega}).$$

Colophon


June 2022
Volume 64 Number 2
siam
REVIEW
a publication of the
Society for Industrial and Applied Mathematics