

For the exercise 1 and 3 download the following dataset http://lcvwww.epfl.ch/teaching/modelling_dna/protected_files/codes_exercises/palin_md_data.mat. The dataset consists an oligomer-based statistics of a 24 basepairs long Palindrome. The matlab structure contains the following fields:

- seq : sequence,
- nbp : number of base pair,
- nsnap : total number of accepted snapshots from the MD simulation (= M).
- shape : ensemble mean ($\bar{\mathbf{w}} = \frac{1}{M} \sum_{j=1}^M \mathbf{w}^{[j]}$),
- c1b : ensemble covariance, ($C = \frac{1}{M} \sum_{j=1}^M (\mathbf{w}^{[j]} - \bar{\mathbf{w}}) \otimes (\mathbf{w}^{[j]} - \bar{\mathbf{w}})$)
- stiff_me : maximum entropy fit to c1b.

1 Banded matrices and their inverses

Consider the following symmetrically partitioned matrix:

$$C = \begin{bmatrix} a & e & \mathbf{x} \\ e^T & b & d \\ \mathbf{x}^T & d^T & c \end{bmatrix}, \quad a = a^T, \quad b = b^T, \quad c = c^T, \quad (1)$$

where for simplicity we assume $C > 0$. We want to show that if $\mathbf{x} = eb^{-1}d$, then $K := C^{-1}$ has zeros blocks in the (1,3) and (3,1) entries. For that prove the following statements:

1. Show that the matrix C can be decomposed as follow

$$C = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & \Psi & I \end{bmatrix} \begin{bmatrix} a & e & 0 \\ e^T & b & 0 \\ 0 & 0 & H \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ 0 & I & \Omega \\ 0 & 0 & I \end{bmatrix}, \quad (2)$$

where $\Omega = b^{-1}d$, $\Psi = d^T b^{-1}$, and $H = c - d^T b^{-1}d$.

2. Using the decomposition obtained in the previous point, compute the inverse of C .
3. Conclude that the blocks in the (1,3) and (3,1) entries of $K = C^{-1}$ are zero.
4. From the latter point derive an algorithm that will allow you to compute the matrix K that involve only the blocks a, b, c, d and e (but not \mathbf{x}). Code your algorithm and test it on the two first blocks of the c1b matrix.
5. Can you capture the five blocks a, b, c, d and e of C from the five non zero blocks of K more simply than just inverting K ?

2 Entropy and Relative entropy formulas for Gaussians

1. The entropy of a probability density function $p : \mathbb{R}^N \mapsto \mathbb{R}$ is defined as

$$h(p) = - \int_{\mathbb{R}^N} p(x) \ln(p(x)) dx, \quad (3)$$

(sometimes without the minus signs).

Prove that when p is Gaussian, i.e,

$$p(x) = \frac{1}{(2\pi)^{N/2} |K^{-1}|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (x - \hat{x}) \cdot K (x - \hat{x}) \right\}, \quad x \in \mathbb{R}^N,$$

the entropy can be evaluated as

$$h(p) = -\frac{1}{2} \ln \left((2\pi e)^N |K^{-1}| \right) = \frac{1}{2} \ln(|K|) - \frac{N}{2} (\ln(2\pi) + 1) \quad (4)$$

Note: The latter computation can be used to compute the expectation of a quadratic energy $U(x) = \frac{\beta}{2} (x - \hat{x}_1) \cdot K (x - \hat{x}_1)$ with respect to p . In fact one can show that

$$\langle U(x) \rangle_p = \int_{\mathbb{R}^N} U(x) p(x) dx = \frac{\beta}{2} N. \quad (5)$$

In the DNA context $\beta = \frac{1}{k_B T}$, where k_B is the Boltzmann constant and T is the kinetic temperature of the bath where the DNA is immersed in.

2. The relative entropy (or Kullback-Liebler divergence) between two probability density functions $p : \mathbb{R}^N \mapsto \mathbb{R}$ and $q : \mathbb{R}^N \mapsto \mathbb{R}$ is defined by

$$D(p, q) := \int_{\mathbb{R}^N} p(x) \ln \left[\frac{p(x)}{q(x)} \right] dx. \quad (6)$$

Consider now the particular case when p and q are Gaussian probability density functions, i.e,

$$\begin{aligned} p(x) &= \frac{1}{(2\pi)^{N/2} |K_1^{-1}|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (x - \hat{x}_1) \cdot K_1 (x - \hat{x}_1) \right\}, \quad x \in \mathbb{R}^N, \\ q(x) &= \frac{1}{(2\pi)^{N/2} |K_2^{-1}|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (x - \hat{x}_2) \cdot K_2 (x - \hat{x}_2) \right\}, \quad x \in \mathbb{R}^N, \end{aligned}$$

Prove that the relative entropy of p and q can be written as

$$\begin{aligned} D(p, q) &= \frac{1}{2} \left[\text{tr} (K_2 K_1^{-1}) - \ln \frac{|K_2|}{|K_1|} - N \right] + \frac{1}{2} (\hat{x}_1 - \hat{x}_2) \cdot K_2 (\hat{x}_1 - \hat{x}_2) \\ &= D^\dagger(K_1, K_2) + \frac{1}{2} (\hat{x}_1 - \hat{x}_2) \cdot K_2 (\hat{x}_1 - \hat{x}_2), \end{aligned}$$

where alternative form for $D^\dagger(K_1, K_2)$ is

$$D^\dagger(K_1, K_2) = \frac{1}{2} \left[K_2 : K_1^{-1} - \ln \frac{|K_2|}{|K_1|} - \text{I} : \text{I} \right].$$

3 Kullback-Liebler divergence between : $\rho_{obs}(S)$, $\rho_{band}(S)$, $\rho_{cgDNAp}(S, \mathcal{P})$

In this exercise we want to study the Kullback-Leibler divergence between the different steps of the approximation of the stiffness and the mean for a given sequence. We will use the following notation: $\rho_{obs}(S)$ is the Gaussian distribution which parameters are the ensemble mean and the ensemble covariance, $\rho_{band}(S)$ is the Maximum Entropy Gaussian distribution (which follows the cgDNA+ sparsity pattern ie the sparsity pattern of the stiffness matrix is 42×42 blocks with 18×18 overlaps in the interior of the sequence and 36×36 blocks with 18×18 overlaps for ends) fitted to the ensemble mean and covariance, and $\rho_{cgDNAp}(S, \mathcal{P})$ is the cgDNA+ reconstruction of S using the parameter set \mathcal{P} (use `cgDNA+ps1.mat`).

1. Write a matlab script implementing the Kullback-Leibler divergence for Gaussian distributions (see exercise 2.2). Allow your script to output the value of the Kullback-Leibler divergence, the value of the stiffness part, and the value of the mean part (Mahalanobis distance).
2. Using your script do the following computations:

- i) $D(\rho_{band}(S), \rho_{obs}(S))$,
- ii) $D(\rho_{cgDNAp}(S, \mathcal{P}), \rho_{band}(S))$,

Compare the values of the Kullback-Leibler divergence per degree of freedom as well as the values coming from the stiffness part and the mean part. What can you say?

4 Jensen's inequality - Part 1

Let $\Omega \subset \mathbb{R}^n$ be an nonempty open set, such that $\mu(\Omega) := |\Omega| < \infty$ with respect to a measure μ . Let $u \in L^1(\Omega)$ with respect to μ , with range $R(u(\Omega))$, and $\phi : R(u(\Omega)) \rightarrow \mathbb{R}$ be convex. Then a version of Jensen's inequality is

$$\phi\left(\frac{1}{|\Omega|} \int_{\Omega} u(x) d\mu(x)\right) \leq \frac{1}{|\Omega|} \int_{\Omega} \phi(u(x)) d\mu(x). \quad (7)$$

To prove this inequality we know that because ϕ is convex $\forall y^* \in R(u(\Omega))$, $\exists a, b, \in \mathbb{R}$ such that $\phi(y) \geq ay + b$ for all $y \in R(u(\Omega))$, and $\phi(y^*) = ay^* + b$. By making a choice for y^* , obtain the inequality for any Ω with $|\Omega|$ finite.

[Note on the regularity of ϕ & $u(x)$: the case $\phi \in C^2$ with $\phi'' > 0$ and $u(x) \in C(\Omega)$ with Ω connected is sufficient for us.]