Mathematical Modelling of DNA
PROF. JOHN MADDOCKS
SESSION 9 A. PATELLI

1 Gaussian integrals II (Part I in Session 1)

Let $\beta > 0$, $n \geq 0$ and a symmetric, positive - definite matrix $K = \Sigma^{-1} \in \mathbb{R}^{n\times n}$ be given. Show that a marginal of a Gaussian distribution is also a Gaussian distribution: if $\mathbf{x} \sim N(\hat{\mathbf{x}}, \Sigma)$,

$$\begin{aligned}\mathbf{x} &= \begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{bmatrix}, \quad \hat{\mathbf{x}} = \begin{bmatrix}\hat{\mathbf{x}}_1\\ \hat{\mathbf{x}}_2\end{bmatrix}, \quad \mathbf{x}_1, \hat{\mathbf{x}}_1 \in \mathbb{R}^k, \quad \mathbf{x}_2, \hat{\mathbf{x}}_2 \in \mathbb{R}^m,\\ \Sigma &= \begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{12}^T & \Sigma_{22}\end{bmatrix}, \quad \Sigma_{11} = \Sigma_{11}^T \in \mathbb{R}^{k\times k}, \quad \Sigma_{12} \in \mathbb{R}^{k\times m},\\ &\Sigma_{22} = \Sigma_{22}^T \in \mathbb{R}^{m\times m}, \quad \text{and} \quad k + m = n,\end{aligned}$$

then $\mathbf{x}_1 \sim N(\hat{\mathbf{x}}_1, \Sigma_{11})$, i.e.

$$\frac{\left(\frac{\beta}{\pi}\right)^{\frac{n}{2}}}{\sqrt{\det[K^{-1}]}} \int_{\mathbb{R}^m} e^{-\beta(\mathbf{x}-\hat{\mathbf{x}})\cdot K(\mathbf{x}-\hat{\mathbf{x}})}\, d\mathbf{x}_2 = \frac{\left(\frac{\beta}{\pi}\right)^{\frac{k}{2}}}{\sqrt{\det \Sigma_{11}}}\, e^{-\beta(\mathbf{x}_1-\hat{\mathbf{x}_1})\cdot\Sigma_{11}^{-1}(\mathbf{x}_1-\hat{\mathbf{x}_1})}.$$

[Hint: Found explicitly the inverse of a symmetric, positive definite matrix of the following form $\begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{B}^T & \mathbf{C}\end{bmatrix}$, with $A = A^T$ and $C = C^T$ and compute its determinant.]

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

$$h(p) = -\int_{\mathbb{R}^N} p(\mathbf{x}) \ln(p(\mathbf{x})) d\mathbf{x}, \tag{1}$$

(sometimes without the minus signs).
Prove that when p is Gaussian, i.e,

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

the entropy can be evaluated as

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

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

$$D(p, q) := \int_{\mathbb{R}^N} p(\mathbf{x}) \ln\left[\frac{p(\mathbf{x})}{q(\mathbf{x})}\right] dx. \tag{3}$$

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

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

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

$$\begin{aligned} D(p,q) &= \frac{1}{2}\left[\operatorname{tr}\left(K_2K_1^{-1}\right) - \ln\frac{\det K_2}{\det K_1} - N\right] + \frac{1}{2}(\widehat{\mathbf{x}}_1-\widehat{\mathbf{x}}_2)\cdot K_2(\widehat{\mathbf{x}}_1-\widehat{\mathbf{x}}_2) \\ &= D^{\dagger}(K_1,K_2) + \frac{1}{2}(\widehat{\mathbf{x}}_1-\widehat{\mathbf{x}}_2)\cdot K_2(\widehat{\mathbf{x}}_1-\widehat{\mathbf{x}}_2). \end{aligned}$$

3 Jensen's inequality

Let $\Omega\subset\mathbb{R}^n$ be a bounded open set, $u\in L^1(\Omega)$, with range $R(u(\Omega))$ and $\phi : R(u(\Omega))\to\mathbb{R}$ be convex. Then $\forall y^*\in R(u(\Omega))\exists a,b,\in\mathbb{R}$ such that $\phi(y)\geq ay+b$ with $\phi(y^*) = ay^*+b$. A version of Jensen's inequality is

$$\phi\left(\frac{1}{|\Omega|}\int_\Omega u(x)dx\right) \leq \frac{1}{|\Omega|}\int_\Omega \phi(u(x))dx. \tag{4}$$

1. Read and understand the Proof 2 in `https://en.wikipedia.org/wiki/Jensen%27s_inequality` for the case $|\Omega| = 1$. Verify that the equality arises if and only if u =const (a.e).

2. By making a different choice for y^*, obtain the inequality for any Ω with $|\Omega|$ finite.

[Note on the regularity of ϕ: the case $\phi\in C^2$ with $\phi'' > 0$ is sufficient for us]