Mathematical Modelling of DNA	
Prof. John Maddocks	
Lecture Notes	M.Mordig

The following notes has been written by Maximilian Mordig and are based on the lecture about cgDNAeq given by Prof. John Maddocks as part of the course "Mathematical Modelling of DNA", spring semester 2016.

1 Notes on Lagrange multipliers and cgDNAeq

We want to compute the most probably DNA configuration subject to constraints, i.e. both ends of the DNA are fixed in space. We call $\mathbf{w}^*$ the most probable configuration of the cgDNA PDF. It coincides with the minimum μ of the cgDNA energy because the distribution is Boltzmann, i.e. $\mathbf{w}^* = \mu$. This does no longer hold if there are constraints to satisfy, e.g. if one end of the DNA is fixed. In the minicircles experiment, the DNA is circular, that is, both ends are joined and in the magnetic tweezer experiment both ends of the DNA are fixed. The goal of the following notes is to find the most probable DNA configuration under these constraints. More precisely, the goal is to solve

$$\min_{\mathbf{w}} f(\mathbf{w}) \qquad \text{subject to } \mathbf{G}(\mathbf{w}) = \mathbf{g}. \tag{1}$$

In the cgDNA case, $f(\mathbf{w})$ is a quadratic form, $\mathbf{G}(\mathbf{w})$ is some function and $\mathbf{g}$ is a constant.

1.1 Lagrange multipliers

We will recall the general framework of Lagrange multipliers. The goal is to solve the constrained minimisation problem

$$\min_{\mathbf{x}} f(\mathbf{x}) \qquad \text{subject to } \mathbf{g}(\mathbf{x}) = \mathbf{g}_0. \tag{2}$$

$f : \mathbb{R}^n \to \mathbb{R}$, $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$ are given functions, $\mathbf{g}_0 \in \mathbb{R}^m$, $m \leq n$. Usually, $m \ll n$, i.e. there are only a few constraints. We need to understand the tangent space to the constraints at a point $\mathbf{x}_0$. Because there are n components and m constraints, we expect it to be of dimension $n - m$.

Consider all curves $\mathbf{x}(\varepsilon)$ starting at $\mathbf{x}_0$ and staying in the constrained set, i.e.

$$\mathbf{g}(\mathbf{x}(\varepsilon)) = \mathbf{g}_0 \quad \forall \varepsilon, \qquad \mathbf{x}(0) = \mathbf{x}_0. \tag{3}$$

$\forall \varepsilon$ can be replaced by $\forall \varepsilon$ in a neighbourhood of $\varepsilon = 0$. Differentiating the expression gives

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon} \mathbf{g}(\mathbf{x}(\varepsilon)) \bigg|_{\varepsilon=0} = \mathbf{0} \in \mathbb{R}^m \iff \nabla \mathbf{g}(\mathbf{x}_0) \mathbf{x}'(0) = \mathbf{0} \in \mathbb{R}^m. \tag{4}$$

$\nabla \mathbf{g}(\mathbf{x}_0) \in \mathbb{R}^{m \times n}$ is the Jacobian matrix[1], $\mathbf{x}_0 \in \mathbb{R}^n$ a column vector and $\mathbf{x}'$ is the derivative with respect to ε. The tangents $\mathbf{x}'(0)$ of $\mathbf{x}(\varepsilon)$ satisfying (3) form a basis of the tangent space $T_{\mathbf{g}}(\mathbf{x}_0)$ that contains all $\delta \mathbf{x}$ such that

$$\nabla \mathbf{g}(\mathbf{x}_0) \delta \mathbf{x} = \mathbf{0} \in \mathbb{R}^m. \tag{5}$$

[1] Be careful about the dimensions of the matrix. We take the convention that for a function $f : \mathbb{R}^n \to \mathbb{R}^m$, the Jacobian matrix is in $\mathbb{R}^{m \times n}$, i.e. row j corresponds to the gradient of f_j as a row vector.

That is, $T_{\mathbf{g}}(\mathbf{x}_0) = \text{Nul}(\nabla \mathbf{g}(\mathbf{x}_0))$. Using that $\text{Nul}(A)^\perp = \text{Range}(A^T)$, we obtain the vector space $N_{\mathbf{g}}(\mathbf{x}_0)$ normal to the tangent space

$$N_{\mathbf{g}}(\mathbf{x}_0) = T_{\mathbf{g}}(\mathbf{x}_0)^\perp = \text{Range}(\nabla \mathbf{g}(\mathbf{x}_0)^T). \tag{6}$$

In other words, we have

$$\mathbf{x} \in N_{\mathbf{g}}(\mathbf{x}_0) \iff \mathbf{x} = \nabla \mathbf{g}(\mathbf{x}_0)^T \lambda \quad \text{for some } \lambda \in \mathbb{R}^m \text{ (column vector)}. \tag{7}$$

We have characterised all valid curves $\mathbf{x}(\varepsilon)$. Now, let $\mathbf{x}_0$ be the solution of (2) that minimises f under the given constraints. The first-order necessary conditions read as

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(\mathbf{x}(\varepsilon))\bigg|_{\varepsilon=0} = \mathbf{0} \in \mathbb{R}^m \iff \nabla f(\mathbf{x}_0)\mathbf{x}'(0) = \mathbf{0} \in \mathbb{R}^m \tag{8}$$

for any path that satisfies (3), i.e. for any path such that $\mathbf{x}'(0) \in T_{\mathbf{g}}(\mathbf{x}_0)$. Hence, $\nabla f(\mathbf{x}_0) \in T_{\mathbf{g}}(\mathbf{x}_0)^\perp = N_{\mathbf{g}}(\mathbf{x}_0)$. Using (7), we find that there must exist $\lambda \in \mathbb{R}^m$ such that

$$\nabla f(\mathbf{x}_0) = \nabla \mathbf{g}(\mathbf{x}_0)^T \lambda. \tag{9}$$

Taking the transpose and redefining λ gives

$$\nabla f(\mathbf{x}_0)^T + \lambda^T \nabla \mathbf{g}(\mathbf{x}_0) = 0. \tag{10}$$

Hence, the minimisation problem (2) can be rewritten as: Find $\mathbf{x}_0 \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^m$ s.t.

$$\begin{cases} \nabla f(\mathbf{x}_0)^T + \lambda^T \nabla \mathbf{g}(\mathbf{x}_0) = 0, \\ \mathbf{g}(\mathbf{x}_0) = \mathbf{g}_0. \end{cases} \tag{11}$$

Observe $\nabla f(\mathbf{x}_0)^T \in \mathbb{R}^{n\times 1}$ is a column vector adopting the definition of the Jacobian matrix from before, it is the usual gradient.

1.2 Particular constraints in cgDNA

Consider the coordinates $\mathbf{z}_j, j = 1, \ldots, N$, where each $\mathbf{z}_m$ is a Cayley vector for a matrix $Q_m(\mathbf{z}_m) \in SO(3)$ and define $\mathbf{x} = (\mathbf{z}_1, \ldots, \mathbf{z}_N) \in \mathbb{R}^{3N}$. The goal is again to minimise a function under constraints as in (2). f is quadratic in the coordinates in the cgDNA model. We fix the orientation of the last base frame with respect to the first base frame, i.e. we define

$$\mathbf{g}(\mathbf{x}) = Q(\mathbf{z}_1) \cdots Q(\mathbf{z}_N) \tag{12}$$

and minimise under the constraint $\mathbf{g}(\mathbf{x}) = \mathbf{g}_0$ for some $\mathbf{g}_0 \in SO(3)$. The $\mathbf{z}_i$ are the Cayley vectors of the inter coordinates in the configuration vector $\mathbf{w}$. To solve the minimisation problem (2) with these choices of f and $\mathbf{g}$, we need to compute the normal space to this particular constraint and derive a multiplier rule. It boils down to finding $\nabla \mathbf{g}(\mathbf{x}_0)$. In one approach, one uses the injection $SO(3) \hookrightarrow \mathbb{R}^9$ and one grinds. A more clever approach exploits the geometry of the problem and is exposed now.

Consider a curve $\mathbf{x}(\varepsilon) \in \mathbb{R}^{3N}$ satisfying the constraints. The first-order necessary conditions are

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(\mathbf{x}(\varepsilon))\bigg|_{\varepsilon=0} = \mathbf{0} \in \mathbb{R}^m. \tag{13}$$

For a quadratic form f, this yields $K(\mathbf{x}(0)-\mathbf{x})\cdot\mathbf{x}'(0)=\mathbf{0}$ for all $\mathbf{x}'(0)$ in the tangent space. In the cgDNA context, the whole problem consists in characterising the tangent space. We use a clever chain rule for differentiation to do so.

Assume a path $R(\varepsilon)\in SO(3)^N$ defined by

$$R(\varepsilon)=(Q_1(\varepsilon),\ldots,Q_N(\varepsilon))\in SO(3)^N. \tag{14}$$

For instance, $Q_1(\varepsilon)=Q_1(\mathbf{x}(\varepsilon))$, but Q_1 actually only depends on the first Cayley vector in $\mathbf{x}(\varepsilon)$. Differentiating the constraints yields

$$\mathbb{R}^{3\times 3}\ni\mathbf{0}=\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathbf{g}(\mathbf{x}(\varepsilon))\bigg|_{\varepsilon=0}=\sum_{j=1}^{N}\left(\prod_{i=1}^{j-1}Q_i\,\frac{\mathrm{d}}{\mathrm{d}\varepsilon}Q_j\bigg|_{\varepsilon=0}\prod_{k=j+1}^{N}Q_k\right). \tag{15}$$

How to differentiate a matrix $Q(\varepsilon)\in SO(3)$? From Session 2 and 3, there exist vectors $\delta\varphi,\delta\psi\in\mathbb{R}^3$ such that

$$Q'(\varepsilon)=\delta\psi^{\times}Q=Q\delta\varphi^{\times}. \tag{16}$$

We have defined

$$u^{\times}=\begin{pmatrix}0 & u_3 & u_2\\ u_3 & 0 & -u_1\\ -u_2 & u_1 & 0\end{pmatrix}. \tag{17}$$

We can then rewrite (16) to find

$$Q\delta\varphi^{\times}=Q\delta\varphi^{\times}Q^TQ=\delta\psi^{\times}Q, \tag{18}$$

which is equivalent to

$$\delta\psi^{\times}=Q\delta\varphi^{\times}Q^T=(Q\delta\varphi)^{\times}, \tag{19}$$

where the last equality also comes from Exercise 1 Session 2.

We introduce the notation $R_j=\prod_{i=1}^{j}Q_i$ and continue from (15)

$$\begin{aligned}\sum_{j=1}^{N}\left(\prod_{i=1}^{j-1}Q_i\cdot Q_j\delta\varphi_j^{\times}\cdot\prod_{k=j+1}^{N}Q_k\right)\Bigg|_{\varepsilon=0}&=\sum_{j=1}^{N}R_j\delta\varphi_j^{\times}R_j^TR_N\Bigg|_{\varepsilon=0}=\sum_{j=1}^{N}(R_j\delta\varphi_j)^{\times}R_N\Bigg|_{\varepsilon=0}\\&=\left(\sum_{j=1}^{N}R_j\delta\varphi_j\right)^{\times}R_N\Bigg|_{\varepsilon=0}=\mathbf{0}\in\mathbb{R}^{3\times 3}.\end{aligned} \tag{20}$$

R_N is invertible and we obtain conditions on $\delta\varphi_j$ (evaluated at $\varepsilon=0$)

$$\left(\sum_{j=1}^{N}R_j(0)\delta\varphi_j\right)^{\times}=\mathbf{0}\in\mathbb{R}^{3\times 3}\iff R\delta\varphi=\mathbf{0}\in\mathbb{R}^3, \tag{21}$$

where $R=\begin{pmatrix}R_1 & \cdots & R_N\end{pmatrix}\in\mathbb{R}^{3\times 3N}$, $\delta\varphi=(\delta\varphi_1,\ldots,\delta\varphi_N)\in\mathbb{R}^{3N}$.

Similarly to before, the Lagrange multipliers $\lambda\in\mathbb{R}^3$ describe the tangent space and we can write something similar to λR to characterise the normal space. We got a condition involving $\delta\varphi_j$, but we would like to get an expression in terms of $\delta\mathbf{x}$ or $\delta\mathbf{z}_j$, i.e. variations in the $\mathbf{x}$ coordinates, because we minimise with respect to these coordinates. By a second chain rule, we find that for any coordinates, there exists a linear operator $L(\mathbf{z}_i)$ such that

$$\delta\psi_i = L(\mathbf{z}_i)\delta\mathbf{z}_i. \tag{22}$$

Note that $L(\mathbf{z}_i)$ is non-linear with respect to $\mathbf{z}_i$. For Cayley vectors, the operator $L(\mathbf{z}_i)$ is given by

$$\frac{2\gamma}{1+\gamma^2\left\|z_i\right\|^2}\left(I_3+\gamma\mathbf{z}_i^\times\right). \tag{23}$$

The first term is related to Jacobian corrections in the non-Gaussian PDF part. $\gamma = \frac{1}{10}$ is a scale factor, a factor of 2 to be consistent with the definition of Cayley vectors (choice of θ or $\frac{\theta}{2}$), a factor of 5 to account for the rescaling of the rotational coordinates (see two lectures before). Finally, the constrained problem is equivalent to finding $\mathbf{x} = (\mathbf{z}_1, \dots, \mathbf{z}_N)$ such that

$$\begin{cases} \nabla_x f(\mathbf{x}) + \lambda RL &= \mathbf{0}, \\ g(\mathbf{x}) &= \mathbf{g}_0. \end{cases} \tag{24}$$

To obtain it, we have used the analogy with the previous Lagrange multiplier calculations. The matrix RL is defined as

$$RL = \begin{pmatrix} R_1L(\mathbf{z}_1) & \cdots & R_NL(\mathbf{z}_N) \end{pmatrix} \in \mathbb{R}^{3\times 3N}. \tag{25}$$

From a simulation point of view, $\nabla_x f(\mathbf{x}) + \lambda RL$ will not be exactly zero and this can be interpreted as the force exerted on each part of the DNA. A tangential force contracts the DNA because it decreases the radius of the helical structure of the DNA.

In the cgDNA context, we are interested in $SE(3)$ coordinates and this will also allow fixing, e.g., the translation of the last base frame with respect to the first base frame as well as the rotation between both frames. Analogous formulae exists for differentiating $SE(3)$ matrices, see the notes about cgDNAeq in Session 13..