For questions 1 and 2 there are no solutions provided. Next week we will use cgDNA+ matlab package for computations.

3 Optimisation in SE(3)

i) Is just a review of standard material.

ii) This part is really just a matter of notation as a general matrix can be regarded as a vector in $\mathbb{R}^{n^2}$ and the statement of results in part i) can be used. But as a matter of notation it is simpler to write

$$M(\varepsilon) = M(0) + \varepsilon H + O(\varepsilon^2) \tag{1}$$

where $H = M'(0)$ is the tangent "matrix/vector" which can take any value in $\mathbb{R}^{n\times n}$. Then,

$$\frac{d}{d\varepsilon} I\left(M(\varepsilon)\right)\Big|_{\varepsilon=0} = 0 \in \mathbb{R} \Leftrightarrow H\colon \nabla I(M) = 0, \forall H \Leftrightarrow \nabla I(M) = 0 \in \mathbb{R}^{n\times n} \tag{2}$$

with the stated meaning of : & ∇I.

iii) The point of the entire exercise is that the tangent space of curves $R(\varepsilon)$ of matrices in $SO(n)$ is restricted because of the condition

$$R(\varepsilon)R^T(\varepsilon) = I_d, \ \forall|\varepsilon| << 1. \tag{3}$$

Differentiating w.r.t. ε implies $R'R^T + RR'^T = 0$ so that $(R'R^T)$ is a skew-symmetric matrix $S = -S^T \in \mathbb{R}^{n\times n}$. In other words $R' = SR = R\mathbf{S}$ with $\mathbf{S} = R^T SR = -\mathbf{S}^T$. The Taylor expansion is

$$R(\varepsilon) = R(0) + \varepsilon R(0)\mathbf{S} + O(\varepsilon^2), \tag{4}$$

and the tangent "matrix/vector" is of the form $R\mathbf{S}$ for a skew-symmetric $\mathbf{S}$. Then the first-order necessary conditions are

$$\begin{aligned} \frac{d}{d\varepsilon} I\left(R(\varepsilon)\right)\Big|_{\varepsilon=0} = 0 \quad &\Leftrightarrow \quad R\mathbf{S}\colon \nabla I(x) = 0, \forall \mathbf{S} = -\mathbf{S}^T \\ &\Leftrightarrow \quad \mathbf{S}\colon R^T\nabla I(R) = 0, \forall \mathbf{S} = -\mathbf{S}^T \\ &\Leftrightarrow \quad R^T\nabla I(R) \in \mathbb{R}^{n\times n} \ \text{ is symmetric.} \end{aligned} \tag{5}$$

Last two equivalences follow from the properties of the matrix inner product, ($CA\colon B = A\colon C^T B, \forall A, B, C \in \mathbb{R}^{n\times n}$ and $\mathbf{S}\colon A = 0, \forall \mathbf{S} = -\mathbf{S}^T \Leftrightarrow A = A^T$).

iv) First order necessary conditions w.r.t $\mathbf{r}$, according to part i) are

$$\mathbf{h}\cdot\nabla_r I = \mathbf{h}\cdot\sum_{i=1}^{M}(\mathbf{r} + R\boldsymbol{\alpha}_i - \mathbf{p}_i) = 0, \forall \mathbf{h} \in \mathbb{R}^3 \tag{6}$$

or

$$\mathbf{r} = \frac{1}{M}\sum_i(\mathbf{p}_i - R\boldsymbol{\alpha}_i) = \frac{1}{M}\sum_i \mathbf{p}_i - \frac{R}{M}\sum_i \boldsymbol{\alpha}_i \tag{7}$$

First-order necessary conditions w.r.t R, according to part iii) are $\mathbf{S} : R^T \nabla_R I = 0$, $\forall \mathbf{S} = -\mathbf{S}^T \in \mathbb{R}^{3\times 3}$. Only issue is how to compute the gradient $\nabla_R I$. Simplest is to rewrite (11), using $R^T R = Id$, in the form

$$I(\mathbf{r}, R) = 2R : \sum_i (\mathbf{r} - \mathbf{p}_i) \otimes \boldsymbol{\alpha}_i + \sum_i \{||\boldsymbol{\alpha}_i||^2 + ||\mathbf{r} - \mathbf{p}_i||^2\} \tag{8}$$

which implies that

$$\begin{aligned} \nabla_R I &= 2 \sum_i (\mathbf{r} - \mathbf{p}_i) \otimes \boldsymbol{\alpha}_i \\ &= 2\mathbf{r} \otimes \sum_i \boldsymbol{\alpha}_i - 2 \sum_i \mathbf{p}_i \otimes \boldsymbol{\alpha}_i \\ &= \frac{2}{M} \sum \mathbf{p}_i \otimes \sum \boldsymbol{\alpha}_i - 2\frac{R}{M} \sum_i \boldsymbol{\alpha}_i \otimes \sum_i \boldsymbol{\alpha}_i - 2 \sum_i \mathbf{p}_i \otimes \boldsymbol{\alpha}_i. \end{aligned} \tag{9}$$

Then

$$R^T \nabla_R I = \frac{2}{M} R^T \Big((\sum_i \mathbf{p}_i) \otimes (\sum_i \boldsymbol{\alpha}_i) - M \sum_i \mathbf{p}_i \otimes \boldsymbol{\alpha}_i \Big) - \frac{2}{M} (\sum \boldsymbol{\alpha}_i) \otimes (\sum \boldsymbol{\alpha}_i) \tag{10}$$

The last term is symmetric, so in order that $R^T \nabla_R I$ is symmetric we need to take R to be the polar factor of the (known) matrix

$$\Big[(\sum_i \mathbf{p}_i) \otimes (\sum_i \boldsymbol{\alpha}_i) - M \sum \mathbf{p}_i \otimes \boldsymbol{\alpha}_i \Big] \tag{11}$$

(see notes on matrix factorisation). Such a matrix $R \in SO(3)$ is unique, and can be computed for example from the SVD. Then for this R, (7) is an explicit expression for $\mathbf{r}$.

4 Connection between the Cayley transform and the matrix exponential

1. Let $N \in \mathbb{R}^{n\times n}$ and $(I - N)$ invertible. We recall the Neumann expansion

$$(I - N)^{-1} = I + N + N^2 + N^3 + \dots \tag{12}$$

The Cayley transformation of N with a factor $\frac{1}{2}$ can be written as

$$\begin{aligned} CayTra\left(\frac{1}{2}N\right) &= \left(I + \frac{1}{2}N\right)\left(I - \frac{1}{2}N\right)^{-1} \\ &= \left(I + \frac{1}{2}N\right)\left(I + \frac{1}{2}N + \frac{1}{4}N^2 + \frac{1}{8}N^3 + \dots\right) \\ &= I + N + \frac{1}{2}N^2 + \frac{1}{4}N^3 + \dots \end{aligned} \tag{13}$$

The definition of the matrix exponential of N is given by the following Taylor expansion

$$\exp(N) = I + N + \frac{1}{2}N^2 + \frac{1}{6}N^3 + \dots \tag{14}$$

Using (13) and (14) we obtain that

$$\exp(N) - CayTra\left(\frac{1}{2}N\right) = -\frac{1}{12}N^3 + \mathcal{O}(\|O\|^4). \tag{15}$$

Thus for a matrix N with $\|N\| << 1$ we obtain that

$$\exp(N) \approx CayTra\left(\frac{1}{2}N\right) \tag{16}$$

2. Let $\mathbf{v} \in \mathbb{R}^3$ with $\|\mathbf{v}\| = \theta < \pi$. We have that $\exp([\mathbf{v}\times]) = Q(\theta, \mathbf{n}) \equiv Q$ where $\mathbf{n} = \frac{\mathbf{v}}{\theta}$. By direct computation (using equation (2) of Serie 2) we have

$$\begin{aligned}(Q - Q^T) &= 2\sin(\theta)[\mathbf{n}\times] = \frac{2\sin(\theta)}{\theta}[\mathbf{v}\times] \\ \frac{\theta}{2\sin(\theta)}(Q - Q^T) &= [\mathbf{v}\times].\end{aligned}$$

Remark: One can compute rotation matrix Q from given Cayley vector $\mathbf{u}$ as

$$Q = I + \frac{2}{1 + \|\mathbf{u}\|^2}\left([\mathbf{u}\times] + [\mathbf{u}\times]^2\right) \tag{17}$$

and Cayley vector $\mathbf{u}$ can be computed from given rotation matrix Q using the following equation

$$[\mathbf{u}\times] = \frac{1}{1 + \mathrm{tr}Q}\left(Q - Q^T\right). \tag{18}$$

Similarly, rotation matrix Q corresponding to exponential logarithmic coordinates $\mathbf{v}$ can be computed as

$$Q = \mathrm{I} + \sin\theta\,[\mathbf{n}\times] + (1 - \cos\theta)\,[\mathbf{n}\times]^2 \tag{19}$$

where $\mathbf{n} = \frac{\mathbf{v}}{\theta}$ and $\pi > \theta = \|\mathbf{v}\|$ and to compute $\mathbf{v}$ from Q one can use

$$[\mathbf{v}\times] = \frac{\theta}{2\sin(\theta)}\left(Q - Q^T\right). \tag{20}$$

Notice that (19) is analogues to (17) and (20) is analogues to (18). Basically these equations relate the rotation matrix with its Cayley vector or with its exponential logarithmic coordinates and inverse.

Note: By the way in all of these case where Q is rotation through π (i.e. its eigen value -1) is a special case. You can also show that Q has the eigen value -1, a proper rotation matrix has the eigen value -1 if and only if its symmetric.

5 Completing the square in vector quadratic forms

Compare powers of $\mathbf{x}$, i.e. equate coefficient of quadratic, linear & constant terms.
Quadratic term:

$$\mathbf{x}\cdot\left(\sum_{i=1}^{N} K_i\right)\mathbf{x} = \mathbf{x}\cdot K\mathbf{x},\ \forall\ \mathbf{x}\ \Leftrightarrow K = \sum_{i=1}^{N} K_i. \tag{21}$$

Linear term: $\mathbf{x} \cdot \sum K_i \mu_i = \mathbf{x} \cdot K\boldsymbol{\mu}$, $\forall \mathbf{x} \Leftrightarrow$ linear system $K\boldsymbol{\mu} = \sum K_i \mu_i$ is solvable for given K_i & μ_i with $\boldsymbol{\mu}$ &K as in (21). Simplest case (& only case where $\boldsymbol{\mu}$ is unique) is K invertible (which does not require the individual K_i to be invertible). Then

$$\boldsymbol{\mu} = K^{-1}\Big(\sum_{i=1}^{N} K_i \mu_i\Big) \tag{22}$$

Constant term:

$$\begin{aligned} \sum_{i=1}^{N}\Big(\boldsymbol{\mu}_i K_i \boldsymbol{\mu} + c_i\Big) &= c + \boldsymbol{\mu} \cdot K\boldsymbol{\mu} = c + \Big(\sum_{i=1}^{N} K_i \mu_i\Big) C^{-1} \Big(\sum_{i=1}^{N} K_i \mu_i\Big) \\ \Rightarrow c &= \sum \{c_i + \boldsymbol{\mu}_i \cdot K\boldsymbol{\mu}_i\} - (\sum_i K_i \boldsymbol{\mu}_i) \cdot K^{-1}(\sum K_i \boldsymbol{\mu}_i) \end{aligned} \tag{23}$$

And we note formulas are slightly simpler if we introduce $\boldsymbol{\sigma}_i := K_i \boldsymbol{\mu}_i$ & $\boldsymbol{\sigma} = K\boldsymbol{\mu}$. Also we note that there is no simple formula for K^{-1} with K of the form (21) in terms of the K_i and K_i^{-1} even in the case each K_i invertible.