大尺度变形的图像配准及公共模板的确定

Large deformation image registration and atlas building

华东师范大学数学系 沈纯理

本文主要介绍 Large Deformation Diffeomorphic Metric Mapping (LDDMM) 技术及其相关内容.

§1. 微分同胚流形中的测地线

弹性变形，即微分同胚，它是一个保持双方一一对应的可逆的光滑映射:

$$\phi: \Omega \longrightarrow \phi(\Omega)$$
$$x \longmapsto \phi(x).$$

如果微分同胚是小变形的，则 $\phi(x) \simeq x + u(x)$，其中 $u(x)$ 是在 x 处的向量，即 u 是区域 Ω 上的一个向量场. 所以微小变形的微分同胚可以用 Ω 上的一个向量场来实现.

当微分同胚 $\phi(x)$ 是大变形时，如何描述？可以让它一点一点地变过去，即大变形的微分同胚 ϕ 可通过一个连续变形过程来实现，而变形过程可用一族依赖于时间 $t\in[0,1]$ 的微分同胚 $\phi_t(x)$ 来描述，其中 $\begin{cases}\phi_0 = id\\ \phi_1 = \phi\end{cases}$.

也就是说，大变形的微分同胚可以通过一族依赖于时间 t 的向量场 v_t 来实现，其中

$$v_t(\phi_t(x)) = \frac{d\phi_t(x)}{dt}, \quad 且\ \phi_t\Big|_{t=0} = id.$$

反之，如有一族依赖于时间变化的向量场 $v_t(x)$，则从常微分方程组

$$\begin{cases} \dfrac{d\phi_t(x)}{dt} = v_t(\phi_t(x)) \\ \phi_t\big|_{t=0} = id \end{cases} \tag{1}$$

中可以解出一族依赖于时间变化的微分同胚 $\phi_t(x)$. 因此，一族依赖于 t 的微分同胚可等价地描述为一族依赖于 t 的向量场.

如果我们定义 V 为 Ω 上所有向量场所构成的向量空间，则一族依赖于 t 的微分同胚就相当于在 V 空间中的一条曲线. 更确切地说，微分同胚流形中的一条以 t 为参数的曲线的切向量场相当于 V 空间中的一条曲线.

我们可以在向量空间 V 上定义一个内积：

设 $a, b \in V$，(即 a, b 均为 Ω 上的向量场)，则令

$$\langle a, b\rangle_V = \langle La, b\rangle_{L^2} = \int_\Omega \langle La(x), b(x)\rangle dx, \tag{2}$$

其中 L 是一个正定的、自共轭微分算子

$$L: V \longrightarrow V^*,$$

这里 V^* 是 V 的对偶空间. 例如，我们可采用

$$L = -\alpha\Delta + \beta,$$

其中 Δ 是 Laplace 算子，α, β 为正常数.

设 $v \in V$ 是一个向量场，则称

$$m = Lv \in V^* \tag{3}$$

为与 v 相对偶的动量 (momentum)，并记 L 的逆算子为 $K: V^* \longrightarrow V$.

对V中的每一条曲线 v_t，$t\in[0,1]$，可定义其能量

$$E(v_t)=\int_0^1 \langle v_t, v_t\rangle_V\, dt = \int_0^1 \langle Lv_t, v_t\rangle_{L^2}\, dt. \tag{4}$$

则由这个能量的 Euler-Lagrange 方程可求出该能量的极值曲线（也称测地线）。记这条测地线 v_t 的对偶曲线为 $m_t = Lv_t \in V^*$，则测地线 v_t 应满足所谓的

EPDiff 方程

$$\frac{\partial v}{\partial t} = -K(ad_v^* m) = -K\big((Dv)^T m + Dm\, v + m\, div(v)\big), \tag{5}$$

（这里省略了 v_t, m_t 的下标 t），而其中 D 为 Jacobian 矩阵，算子 ad_v^* 是算子 $ad_v : V \longrightarrow V$ 的对偶算子，即

$$\langle u, ad_v^* m\rangle_{L^2} = \langle ad_v u, m\rangle_{L^2},$$

而

$$ad_v w = -[v, w] = Dv\, w - Dw\, v,$$

这里 $[v,w]$ 是向量场 v, w 的 Jacobi-Lie 括号，即对函数 f，有

$$[v,w](f) = v(w(f)) - w(v(f)).$$

注：EPDiff 方程的推导可在 Miller, M.I., Trouvé, A., Younes, L., 2006, Geodesic Shooting for Computational anatomy, J. Math. Imaging, Vis. 24 (2), 209-228 中查到.

* EPDiff 方程也可写成 $\dfrac{\partial m}{\partial t} = -ad_v^* m$.

* 沿着测地线，成立

$$\frac{d}{dt}\left(\langle v, v\rangle_V\right) = 0,$$

即沿测地线，$\|v\|_V^2$ 为常值.

这是因为

$$\frac{d}{dt}\|v\|_V^2 = 2<v, \frac{dv}{dt}>_V = 2<v, -K(ad_v^* m)>_V$$
$$= -2<Lv, K(ad_v^* m)>_{L^2} = -2<v, ad_v^* m>_{L^2}$$
$$= -2<ad_v v, m>_{L^2} = 0.$$

* 因为EPDiff方程是一阶的ODE，所以由初始向量场v_0就可确定出整条测地线。这就是所谓的测地射击法（geodesic shooting）。

* 沿着V中的曲线v_t，用

$$\begin{cases} \dfrac{d\phi_t(x)}{dt} = v_t(\phi_t(x)) \\ \phi_t\Big|_{t=0} = id \end{cases}$$

可积分出一族微分同胚ϕ_t，由此可得出一族变形图像，即初始图像$I_0(x)$随着微分同胚ϕ_t可得出一族变形图像

$$I_t(x) = I_0(\phi_t^{-1}(x)).$$

这族变形图像I_t满足图像迁移方程：

$$\frac{dI_t}{dt} + v\cdot\nabla I_t = 0. \qquad (6)$$

§2. 两个图像之间的配准

现在考虑图像I_0和J的配准问题。

定义能量

$$E(v, I_0, J) = \frac{1}{2}\int_0^1 \langle Lv_t, v_t \rangle_{L^2} dt + \frac{1}{2\sigma^2} \| I_0 \circ \phi_1^{-1} - J \|_{L^2}^2 , \quad (7)$$

这里 $v \in L^2([0,1], V)$，即 $v_t \in V$，$t \in [0,1]$，是一族依赖于 t 的向量场（相当于随着时间 t，从 0 到 1 的一族微分同胚。能量 E 的表达式中右边第二项表示为由 v_t 所生成的在 $t=1$ 处的微分同胚中，所产生的变形图像 I_1 与图像 J 之间的误差，而右边第一项是这个微分同胚族的正则项。配准问题就相当于要去找：

$$v^* = \underset{v \in L^2([0,1], V)}{\operatorname{argmin}} E(v, I_0, J).$$

但这个问题解决起来太复杂，因为 v_t 的变化自由度太大了，不易处理。

所以我们选用测地射击（geodesic shooting）方法设定：将能量 E 中的 v 限制为 V 中的测地线，即 $v = (v_t)$ 还应满足 EPDiff 方程。这时整条 v_t 曲线能由初始向量 v_0 所确定，于是自由度就大为减少了。因为沿着这条测地线 $\|v_t\|_V^2 = \langle Lv_t, v_t \rangle = \text{Const.}$，所以

$$\int_0^1 \langle Lv_t, v_t \rangle_{L^2} dt = \langle Lv_0, v_0 \rangle_{L^2} = \langle v_0, v_0 \rangle_V .$$

所以配准问题变成了一个具有约束条件的能量极小问题，这时能量

$$E(v, I_0, J) = \frac{1}{2} \langle v_0, v_0 \rangle_V + \frac{1}{2\sigma^2} \| I_0 \circ \phi_1^{-1} - J \|_{L^2}^2 , \quad (8)$$

其中微分同胚 ϕ_1 是用测地线的初始向量 v_0 生成的测地线 v_t 所形成的微分同胚族 ϕ_t 在 $t=1$ 处的取值。所以 $E(v, I_0, J)$ 中的约束条件应为

$$\begin{cases} \dot{v} + K \operatorname{ad}_v^* m = 0 & \text{(EPDiff 方程)} \\ \dot{I} + \nabla I \cdot v = 0 & \text{(图像迁移方程)} \\ m - Lv = 0 & \text{(动量的定义)} \end{cases} \quad (9)$$

也可以采用 Bayesian 模型来观察我们的能量 $E(v, I_0, J)$. 设我们的噪声模型在图像的每个 voxel 处是 i.i.d. 高斯噪声，则或然率 (Likelihood) 可表为

$$p(J \mid v, I_0, \sigma) = \frac{1}{(2\pi)^{M/2}\sigma^M} \exp\left(-\frac{\| I_0 \circ \phi_1^{-1} - J \|^2}{2\sigma^2}\right),$$

这里 M 是 voxel 的总数，σ^2 是噪声方差.

v 的先验概率是多变元高斯型，即

$$p(v) = \frac{1}{(2\pi)^{M/2} |L^{-1}|^{1/2}} \exp\left(-\frac{\langle Lv, v\rangle}{2}\right),$$

这里 $|L|$ 是算子 L 的行列式. 于是

$$\log p(v \mid I_0, J, \sigma) \propto -\frac{1}{2}\langle Lv, v\rangle - M \log \sigma - \frac{1}{2\sigma^2} \| I_0 \circ \phi_1^{-1} - J \|^2$$

因为求最大或然率相当于求负对数或然率的极小问题，所以图像配准问题就相当于求

$$E(v, I, \sigma) = \frac{1}{2}\langle Lv, v\rangle + M \log \sigma + \frac{1}{2\sigma^2} \| I_1 - J \|^2 \tag{10}$$

的极小值问题，其中 I_1 是变形图像 I_t 在 $t=1$ 时的取值。这里比前面的模型稍复杂一些，即方差值 σ^2 也作为可调整的参数.

对这个能量极小问题可采用 Lagrange 乘子法，将其转化为无约束的变分问题. 为此，我们定义增强能量 (augmented energy):

$$\begin{aligned}\tilde{E}(v, I, \sigma, \hat{v}, \hat{I}, \hat{m}) = E(v, I, \sigma) + \int_0^1 [&\langle \hat{v}, \dot{v} + K\, ad_v^* m\rangle_{L^2} \\ &+ \langle \hat{I}, \dot{I} + \nabla I \cdot v\rangle_{L^2} \\ &+ \langle \hat{m}, m - Lv\rangle_{L^2}]\, dt\end{aligned} \tag{11}$$

这里 $\hat{v}_t, \hat{I}_t, \hat{m}_t$ 分别是相应于三个约束方程（EPDiff 方程、图像迁移方程和动量定义）的 Lagrange 乘子.

我们采用负梯度下降法去求解上述变分问题.

经计算，$\tilde{E}$ 关于初始速度 v_0 的梯度为

$$\nabla_{v_0} \tilde{E} = v_0 - K\hat{v}_0 . \tag{12}$$

（具体计算过程详见下面的 §3.）

所以初始速度 v_0 的迭代格式为

$$v_0^{(k+1)} = v_0^{(k)} - \tau\left(v_0^{(k)} - K\hat{v}_0^{(k)}\right), \tag{13}$$

这里 τ 是迭代步长.

将初始速度的迭代收敛值 v_0 作为 EPDiff 方程的初值可解出测地线 v_t，再用

$$\begin{cases} \dfrac{d\phi_t(x)}{dt} = v_t\left(\phi_t(x)\right) \\ \phi_t\big|_{t=0} = id \end{cases}$$

可生成微分同胚族 ϕ_t，则 $\phi_1(x)$ 为最终所需求出的大尺度变形的微分同胚.

这时，不难看出

$$\sigma^2 = \frac{1}{M}\| I_0\circ\phi_1^{-1} - J\|_{L^2}^2 . \tag{14}$$

§3. $\tilde{E}$ 关于 I, v, m 的一阶变分的推导过程.

3.1. 考虑 $\tilde{E}$ 关于依赖于时间 t 的变量 I 的变分计算.

$$
\begin{aligned}
\partial_I \tilde{E} &= \frac{\partial}{\partial \varepsilon}\Big|_{\varepsilon=0} \int_0^1 \langle \hat{I}, (\dot{I} + \varepsilon \delta \dot{I}) + \nabla (I + \varepsilon \delta I)\cdot v \rangle_{L^2} \\
&\quad + \frac{1}{\sigma^2} \langle \delta I_1, I_1 - J \rangle_{L^2} \\
&= \frac{1}{\sigma^2} \langle \delta I_1, I_1 - J \rangle_{L^2} + \int_0^1 \langle \hat{I}, \delta \dot{I} + \nabla \delta I \cdot v \rangle_{L^2} \, dt \\
&= \frac{1}{\sigma^2} \langle \delta I_1, I_1 - J \rangle_{L^2} + \langle \hat{I}, \delta I \rangle_{L^2} \Big|_{t=0}^{t=1} - \int_0^1 \langle \dot{\hat{I}}, \delta I \rangle_{L^2} \, dt \\
&\quad + \int_0^1 \langle \hat{I}, \nabla \delta I \cdot v \rangle_{L^2} \\
&= \frac{1}{\sigma^2} \langle \delta I_1, I_1 - J \rangle_{L^2} + \langle \hat{I}_1, \delta I_1 \rangle_{L^2} - \langle \hat{I}_0, \delta I_0 \rangle_{L^2} \\
&\quad - \int_0^1 \langle \dot{\hat{I}} + \nabla\cdot(\hat{I} v), \delta I \rangle_{L^2} .
\end{aligned} \tag{15}
$$

3.2 考虑 $\tilde{E}$ 关于 v 的变分计算.

$$
\begin{aligned}
\partial_v \tilde{E} &= \langle L v_0, v_0 \rangle_{L^2} + \frac{\partial}{\partial \varepsilon}\Big|_{\varepsilon=0} \int_0^1 \langle \hat{v}, \varepsilon \delta \dot{v}, + K \mathrm{ad}^*_{v+\varepsilon \delta v} m \rangle_{L^2} \\
&\qquad + \langle \hat{I}, \dot{I} + \nabla I \cdot (v + \varepsilon \delta v) \rangle_{L^2} \\
&\qquad + \langle \hat{m}, m - L(v + \varepsilon \delta v) \rangle_{L^2} \\
&= \langle L v_0, \delta v_0 \rangle_{L^2} + \int_0^1 \langle v, \delta \dot{v} \rangle dt + \frac{\partial}{\partial \varepsilon}\Big|_{\varepsilon=0} \int_0^1 \langle \mathrm{ad}_{v+\varepsilon\delta v} K \hat{v}, m \rangle_{L^2} \, dt \\
&\quad + \int_0^1 \langle \hat{I}, \nabla I \cdot \delta v \rangle_{L^2} dt - \int_0^1 \langle \hat{m}, m \delta v \rangle_{L^2} \, dt \\
&= \langle L v_0, \delta v_0 \rangle_{L^2} + \langle \hat{v}, \delta v \rangle_{L^2} \Big|_{t=0}^{t=1} - \int_0^1 \langle \dot{\hat{v}}, \delta v \rangle_{L^2} \, dt \\
&\quad + \frac{\partial}{\partial \varepsilon}\Big|_{\varepsilon=0} \int_0^1 \langle -\mathrm{ad}_{K\hat{v}} (v + \varepsilon \delta v), m \rangle_{L^2} \, dt \\
&\quad + \int_0^1 \langle \hat{I}, \nabla I \cdot \delta v \rangle_{L^2} \, dt - \int_0^1 \langle \hat{m}, m \delta v \rangle_{L^2} \, dt
\end{aligned}
$$

$$= \langle L v_0, \delta v_0 \rangle_{L^2} + \langle \hat{v}_1, \delta v_1 \rangle_{L^2} - \langle \hat{v}_0, \delta v_0 \rangle_{L^2}$$
$$+ \int_0^1 \langle -ad^*_{K\hat{v}} m + \hat{I} \nabla I - \dot{\hat{v}} - L\hat{m}, \delta v \rangle_{L^2} \, dt \qquad (16)$$

3.3.

$$\partial_m \tilde{E} = \frac{\partial}{\partial \varepsilon}\Big|_{\varepsilon=0} \int_0^1 \left(\langle \hat{v}, K ad^*_v (m + \varepsilon \delta m) \rangle_{L^2} + \langle \hat{m}, (m + \varepsilon \delta m) \rangle_{L^2} \right) dt$$
$$= \int_0^1 \langle ad_v K\hat{v} + \hat{m}, \delta m \rangle_{L^2} \, dt .$$

3.4. 因为在最优解处，上述三项变分均应为 0，即 $\partial_I \tilde{E} = 0$，$\partial_v \tilde{E} = 0$，$\partial_m \tilde{E} = 0$，而且在最优解 (I, v) 处应有 $\begin{cases} \delta I_0 = 0 \\ \delta v_0 = 0 \end{cases}$，于是我们可将 $\hat{v}, \hat{I}, \hat{m}$ 设定为

$$\begin{cases} -\dot{\hat{I}} - \nabla \cdot (\hat{I} v) = 0 \\ -ad^*_{K\hat{v}} m + \hat{I} \nabla I - \dot{\hat{v}} - L\hat{m} = 0 \\ ad_v K\hat{v} + \hat{m} = 0 \end{cases} , \qquad (18)$$

而且令

$$\begin{cases} \hat{v}_1 = 0 \\ \hat{I}_1 = \frac{1}{\sigma^2}(I_1 - J) \end{cases} , \qquad (19)$$

这样就确保了在最优解处，增强能量 $\tilde{E}$ 关于 I, v, m 的一阶变分为 0. 我们将上述关于 $\hat{I}$, $\hat{v}$ 和 $\hat{m}$ 的約束方程视为一组 ODE，且带有初始条件 (19)，并称其为伴随方程 (adjoint equations). 这样，在计及了伴随方程后，$\partial_v \tilde{E}$ 就

简化成

$$\partial_v \tilde{E} = \langle L v_0, \delta v_0 \rangle_{L^2} - \langle \hat{v}_0, \delta v_0 \rangle_{L^2} = \langle L v_0 - \hat{v}_0, \delta v_0 \rangle_{L^2}$$

$$= \langle K(L v_0 - \hat{v}_0), \delta v_0 \rangle_V = \langle v_0 - K\hat{v}_0, \delta v_0 \rangle_V,$$

所以我们就得到了增强能量关于初始向量 v_0 的一阶变分为

$$\nabla_{v_0} \tilde{E} = v_0 - K\hat{v}_0.$$

因此，在用梯度下降法计算迭代式(13)时，为了要计算 $v_0^{(k)} - K\hat{v}_0^{(k)}$，就必须先从初始向量 $v_0^{(k)}$ 出发，通过 EPDiff 方程解出 $v_t^{(k)}$，再把常微分方程组(1)让 t 从0到1解出微分同胚族 $\phi_t^{(k)}$，于是 $v_t^{(k)}$，$m_t^{(k)}$，$\phi_t^{(k)}$，$I_t^{(k)}$ 均已算出。然后再把伴随方程(18)及初始条件(19)让 t 从1反向解到 $t=0$，最终解得 $\hat{v}_0^{(k)}$.

§4. 讨论 Atlas building 前的一些预备知识

4.1. Expectation Maximization Algorithm (EM算法)

X 是随机变量，它依赖于一个参数族. 我们希望找到某个参数 θ，使得

$$P(X|\theta)$$

达到最大. 我们希望通过迭代 $\{\theta_n\}$ 得出最优解. 设已算出 θ_n，希望计算 θ_{n+1}.

令

$$L(\theta) = \ln P(X|\theta).$$

所以等价地可讨论

$$L(\theta)-L(\theta_n)=\ln P(X|\theta)-\ln P(X|\theta_n)$$

的最大化解.

设问题存在隐蔽变量集合 Z, 所以

$$P(X|\theta)=\sum_{z\in Z}P(X|z,\theta)\cdot P(z|\theta).$$

利用 Jensen 不等式

(对满足 $\sum_{i=1}^{n}\lambda_i=1$ 的常数 $\lambda_i\geq 0$, 成立 $\ln\left(\sum_{i=1}^{n}\lambda_i x_i\right)\geq\sum_{i=1}^{n}\lambda_i\ln(x_i)$.)

可推出

$$\begin{aligned}L(\theta)-L(\theta_n)&=\ln\left(\sum_z P(X|z,\theta)P(z|\theta)\right)-\ln P(X|\theta_n)\\&=\ln\left(\sum_z P(z|X,\theta_n)\cdot\frac{P(X|z,\theta)\cdot P(z|\theta)}{P(z|X,\theta_n)}\right)-\ln P(X|\theta_n)\\&\geq\sum_z P(z|X,\theta_n)\ln\left(\frac{P(X|z,\theta)P(z|\theta)}{P(z|X,\theta_n)}\right)-\ln P(X|\theta_n)\\&=\sum_z P(z|X,\theta_n)\ln\left(\frac{P(X|z,\theta)\,P(z|\theta)}{P(z|X,\theta_n)P(X|\theta_n)}\right)\\&\stackrel{\text{记}}{=}\Delta(\theta|\theta_n)\end{aligned}$$

令

$$l(\theta|\theta_n)=L(\theta_n)+\Delta(\theta|\theta_n),$$

所以

$$L(\theta)\geq l(\theta|\theta_n),$$

而且对 $\theta=\theta_n$ 处

$$\begin{aligned}l(\theta|\theta_n)&=L(\theta_n)+\Delta(\theta_n|\theta_n)=L(\theta_n)+\sum_z P(z|X,\theta_n)\ln\frac{P(X|z,\theta_n)P(z|\theta_n)}{P(z|X,\theta_n)P(X|\theta_n)}\\&=L(\theta_n)+\sum_z P(z|X,\theta_n)\ln\frac{P(X,z|\theta_n)}{P(X,z|\theta_n)}=L(\theta_n).\end{aligned}$$

所以EM算法是去找 θ，使得 $l(\theta|\theta_n)$ 达到最大：

$$
\begin{aligned}
\theta_{n+1} &= \arg\max_{\theta}\left\{ l(\theta|\theta_n)\right\} \\
&= \arg\max_{\theta}\left\{ L(\theta_n) + \sum_{z} P(z|X,\theta_n)\ln\frac{P(X|z,\theta)P(z|\theta)}{P(X|\theta_n)P(z|X,\theta_n)}\right\} \\
&= \arg\max_{\theta}\left\{ \sum_{z} P(z|X,\theta_n)\ln\big(P(X|z,\theta)P(z|\theta)\big)\right\} \\
&= \arg\max_{\theta}\left\{ \sum_{z} P(z|X,\theta_n)\ln\left(\frac{P(X,z,\theta)}{P(z,\theta)}\cdot\frac{P(z,\theta)}{P(\theta)}\right)\right\} \\
&= \arg\max_{\theta}\left\{ \sum_{z} P(z|X,\theta_n)\ln P(X,z|\theta)\right\} \\
&= \arg\max_{\theta}\left\{ E_{Z|X,\theta_n}\left\{\ln P(X,z|\theta)\right\}\right\}.
\end{aligned}
$$

因此EM算法分两步：

1. E-Step：确定出条件期望值 $Q(\theta)=E_{Z|X,\theta_n}\{\ln P(X,z|\theta)\}$
2. M-Step：关于 θ 求 $Q(\theta)$ 的最大值.

于是我们有必要研究下述问题：

设随机变量 Z 的条件概率分布密度函数（pdf）为 $p(Z|X,\theta_n)$，我们需要根据这个pdf函数选出 Z 的 S 个样本 $\{z_1,\cdots,z_S\}$，于是

$$Q(\theta)\approx\frac{1}{S}\sum_{j=1}^{S}\ln P(X,z_j|\theta).$$

4.2 Hamiltonian Monte Carlo 取样方法（HMC）

为了根据pdf函数 $f(x)$ 去取样，我们定义一个Hamiltonian：

$$H(x,\mu) = \underset{\text{势能}}{U(x)} + \underset{\text{动能}}{V(\mu)}$$ ，这里 x 是位置，μ 是动量.

可取 $U(x) = -\log f(x)$

$V(\mu) = -\log g(\mu)$，这里 $g(\mu)$ 不妨取为各向同性的高斯分布密度.
（例如 $g(\mu) = e^{-\frac{\mu^2}{2}}$）

然后按照 Hamiltonian 系统去求解

$$\begin{cases} \dot{x} = \dfrac{\partial H}{\partial \mu} \\ \dot{\mu} = -\dfrac{\partial H}{\partial x} \end{cases} \qquad (*)$$

从 (x,μ) 解出下一时刻的 $(\tilde{x},\tilde{\mu})$，再按照

$$P(\text{接受}) = \min\left(1, \exp\left(-U(\tilde{x}) - V(\tilde{\mu}) + U(x) + V(\mu)\right)\right)$$

的概率去决定是否接受 $\tilde{x}$ 为一个新的样本.

我们采用青蛙跳（leapfrog）方法去求解 Hamiltonian 体系 $(*)$.

我们把从时刻 τ 经步长 ε 到达时刻 $\tau+\varepsilon$ 的计算过程分成三步:

第一步 $\quad \mu(\tau+\varepsilon/2) = \mu(\tau) + \frac{\varepsilon}{2}\dfrac{\partial \log f}{\partial x}(x(\tau))$

第二步 $\quad x(\tau+\varepsilon) = x(\tau) + \varepsilon\,\mu(\tau+\varepsilon/2)$

第三步 $\quad \mu(\tau+\varepsilon) = \mu(\tau+\varepsilon/2) + \frac{\varepsilon}{2}\dfrac{\partial \log f}{\partial x}(x(\tau+\varepsilon))$.

§5. Atlas building in Diffeomorphic Image Registration

给定一组图像 $J_1,\cdots,J_N$.

希望找出一个公共模板图像 I^* 及微分同胚流形中 K条测地线的切向量场 v_t^K，使得

$$I^* = \arg\min_I \frac{1}{N}\sum_{k=1}^{N}\left(\int_0^1 \langle Lv_t^K, v_t^K\rangle dt + \frac{1}{2\sigma^2}\| I\circ(\phi_1^K)^{-1} - J_K\|^2\right)$$

如前相仿，我们可以定义能量

$$E(v_0, I) = \sum_{k=1}^{N} \langle v_0^K, v_0^K\rangle_V + \frac{1}{2\sigma^2}\| I\circ(\phi_1^K)^{-1} - J_K\|_{L^2}^2$$

这里我们运用了测地射击（geodesic Shooting）技术，由初始切向量场 v_0^K 所确定的微分同胚族为 ϕ_t^K，并记变形图像为 $I_t^k(x) = I\left((\phi_t^K)^{-1}(x)\right)$.

如果从 Bayesian 模型的角度来考虑，似然率（Likelihood）可表为

$$p(J_K \mid v_0^K, I, \sigma) = \frac{1}{(2\pi)^{M/2}\sigma^M}\exp\left(-\frac{\| I\circ(\phi^K)^{-1} - J_K\|^2}{2\sigma^2}\right)$$

而 v_0^K 的先验概率

$$p(v_0^K) = \frac{1}{(2\pi)^{M/2}|L^{-1}|^{1/2}}\exp\left(-\frac{\langle Lv_0^K, v_0^K\rangle_{L^2}}{2}\right)$$

所以

$$\begin{aligned}\log\prod_{k=1}^{N} p(J_K, v_0^K \mid I, \sigma) \;\propto\; & -\frac{1}{2}\sum_{k=1}^{N}\langle Lv_0^K, v_0^K\rangle_{L^2} \\ & - MN\log\sigma - \frac{1}{2\sigma^2}\sum_{k=1}^{N}\| I\circ(\phi^K)^{-1} - J_K\|^2 \\ & \overset{\text{def}}{=} E(v_0^K, I, \sigma)\end{aligned}$$

这里 σ^2 也成为整个模型的可供调节的参数.

这时能量带有的本条件为

$$
\begin{cases}
\dot{v}^k + K\, ad^*_{v^k} m^k = 0 & \text{(EPDiff方程)} \\
\dot{I}^k + \nabla I^k \cdot v^k = 0 & \text{(图像迁移方程)} \\
m^k - L v^k = 0 & \text{(动量定义)}
\end{cases}
$$

这时增强能量为

$$
\begin{aligned}
\tilde{E}(v_0^k, I, \sigma, \hat{v}^k, \hat{I}^k, \hat{m}^k) = E(v_0^k, I, \sigma) \\
+ \int_0^1 \Big[\langle \hat{v}^k, \dot{v}^k + K\, ad^*_{v^k} m^k \rangle_{L^2} \\
+ \langle \hat{I}^k, \dot{I}^k + \nabla I^k \cdot v^k \rangle_{L^2} \\
+ \langle \hat{m}^k, m^k - L v^k \rangle_{L^2} \Big] dt
\end{aligned}
$$

相应的伴随方程为

$$
\begin{cases}
\dot{\hat{I}}^k + \nabla \cdot (\hat{I}^k v^k) = 0 \\
-ad^*_{K\hat{v}^k} + \hat{I}^k \nabla I^k - \dot{\hat{v}}^k - L\hat{m}^k = 0 \\
ad_{v^k} K\hat{v}^k + \hat{m}^k = 0
\end{cases}
$$

附带初始条件

$$
\begin{cases}
\hat{v}_1^k = 0 \\
\hat{I}_1^k = \frac{1}{\sigma^2}(I_1^k - J_k)
\end{cases}
$$

我们利用EM算法去求解增强能量 $\tilde{E}$ 的优化问题. 这时参数集为 $\theta = (I, \sigma)$, 隐变量为 v_0^k, $k = 1, \cdots, N$.

1. E-Step: 计算

$$
\begin{aligned}
Q(\theta \mid \theta^{(i)}) &= E_{v_0^k \mid J_k; \theta^{(i)}} \left[\tilde{E}(v_0^k, I, \sigma, \hat{v}^k, \hat{I}^k, \hat{m}^k) \right] \\
&\approx \frac{1}{S} \sum_{j=1}^{S} \sum_{k=1}^{N} \tilde{E}(v_0^{kj}, I, \sigma, \hat{v}^k, \hat{I}^k, \hat{m}^k),
\end{aligned}
$$

这里 v_0^{kj}, $(j = 1, \cdots, S)$ 为 v_0^k 的采样, 这时 pdf 函数为 $p(v_0^k \mid J_k; \theta^{(i)})$

2. M-Step：求解优化问题

$$\theta^{(i+1)} = \arg\min_{\theta} Q(\theta \mid \theta^{(i)})$$

对E-步而言，剩下来的问题是如何用 HMC 取样法求出 v_0^K 的 S个样本

$$v_0^{Kj}, \quad j=1, \cdots, S.$$

关键是要写出这时的 Hamiltonian 系统. 令 Hamiltonian 为

$$H(v_0^K, \mu^K) = U(v_0^K) + V(\mu^K)$$

其中

$$\begin{cases} U(v_0^K) = -\log p(v_0^K \mid J_K; \theta) \\ V(\mu^K) = \langle L\mu^K, \mu^K \rangle \end{cases}$$

所以 Hamiltonian 系统为

$$\begin{cases} \dfrac{dv_0^K}{dt} = \dfrac{\partial H}{\partial \mu^K} = L\mu^K \\ \dfrac{d\mu^K}{dt} = -\dfrac{\partial H}{\partial v_0^K} = -\nabla_{v_0^K} \tilde{E} = -(v_0^K - K\hat{v}_0^K). \end{cases}$$

为了计算 $\hat{v}_0^K$，我们必须运用 EPDiff 方程，将初始向量 v_0^K 从 $t=0$ 解到 $t=1$，然后再从 $t=1$ 出发反向求解伴随方程，得到 $t=0$ 时的 $\hat{v}_0^K$.

对 M-步而言，迭代更新是简单的，它们有 closed form.

$$\sigma^2 = \frac{1}{MNS} \sum_{j=1}^{S} \sum_{K=1}^{N} \| I \circ (\phi_1^{Kj})^{-1} - J^K \|^2$$

$$I = \frac{\sum_{j=1}^{S} \sum_{K=1}^{N} J^K \circ \phi_1^{Kj} \, |D\phi_1^{Kj}|}{\sum_{j=1}^{S} \sum_{K=1}^{N} |D\phi_1^{Kj}|},$$

其中 ϕ^{Kj} 是由 v_0^{Kj} 所决定的微分同胚.

本文以设内容可通过下面两篇论文及其参考文献来获得:

[1] Zhang, M., Singh, N., Fletcher, P.T., 2013, Bayesian estimation of regularization and atlas building in diffeomorphic image registration, In: Information Processing in Medical Imaging, Springer, p.37-48.

[2] Miaomiao Zhang, P. Thomas Fletcher, Bayesian principal geodesic analysis for estimating intrinsic diffeomorphic image variability, Medical Image Analysis 25 (2015) 37-44.