Local Tangent Space Alignment Symbol Glossary

Thomas Cohn

June 14, 2019

$\mathcal{F}$	The data manifold. $\mathcal{F}$ is a parameterizable d-manifold in $\mathbb{R}^m$, with $d < m$.
d	The underlying dimension of the data manifold $\mathcal{F}$, i.e., the "feature" space.
m	The dimension of the higher dimensional "input" space.
f	$f : C \subset \mathbb{R}^d \to \mathbb{R}^m$ is a smooth map which parameterizes $\mathcal{F}$.
C	$C \subset \mathbb{R}^d$ is a compact subset of $\mathbb{R}^d$.
x_i	$x_1, \ldots, x_N \in \mathbb{R}^m$ are the data points in the higher dimensional space.
τ_i	$\tau_1, \ldots, \tau_N \in \mathbb{R}^d$ are the corresponding feature space coordinates of the x_i's, where $x_i = f(\tau_i) + \epsilon_i$.
ϵ_i	$\epsilon_1, \ldots, \epsilon_N \in \mathbb{R}^d$ is the added noise to $f(\tau_i)$ to obtain x_i.
X	$X = \begin{bmatrix} \vert & & \vert \\ x_1 & \cdots & x_N \\ \vert & & \vert \end{bmatrix}$ is the matrix form of the data.
J_f	$J_f(\tau) = \begin{bmatrix} \partial f_1/\partial \tau_1 & \cdots & \partial f_1/\partial \tau_d \\ \vdots & \ddots & \vdots \\ \partial f_m/\partial \tau_1 & \cdots & \partial f_m/\partial \tau_d \end{bmatrix}$ is the Jacobi matrix of f at τ.
$\overline{\tau}$	$\overline{\tau}$ is some point near τ used for the following approximation via Taylor expansion: $f(\overline{\tau}) = f(\tau) + J_f(\tau) \cdot (\overline{\tau} - \tau) + O(\vert\vert\overline{\tau} - \tau\vert\vert^2)$.
$\mathcal{T}_\tau$	$\mathcal{T}_\tau = \text{span}(J_f(\tau))$ is the tangent space of f at τ.
Q_τ	Q_τ is a matrix forming some orthonormal basis of $\mathcal{T}_\tau$, allowing us to write $J_f(\tau)(\overline{\tau} - \tau) = Q_\tau \theta^*_\tau$.
θ^*_τ	??? Possibly defined by $Q_\tau \theta^*_\tau = f(\overline{\tau}) - f(\tau)$?
P_τ	Defined by $\theta^*_\tau = Q_\tau^T J_f(\tau)(\tau - \overline{\tau}) \equiv P_\tau(\overline{\tau} - \tau)$.
θ_τ	The mapping from τ to θ^*_τ is a local affine transformation. θ_τ is the orthogonal projection of $f(\overline{\tau}) - f(\tau)$ onto $\mathcal{T}_\tau$. $\theta_\tau \equiv Q_\tau^T(f(\overline{\tau}) - f(\tau)) = \theta^*_\tau + O(\vert\vert\overline{\tau} - \tau\vert\vert^2) \approx \theta^*_\tau$.
$\Omega(\tau)$	$\Omega(\tau)$ defines the neighborhood of τ.
X_i	$X_i = \begin{bmatrix} \vert & & \vert \\ x_{i_1} & \cdots & x_{i_k} \\ \vert & & \vert \end{bmatrix}$ is a matrix consisting of the k-nearest neighbors of x_i.
x	x is part of the minimization problem $\displaystyle\min_{x,\Theta,Q} \sum_{j=1}^{k} \vert\vert x_{i_j} - (x + Q\theta_j)\vert\vert_2^2 = \min_{x,\Theta,Q} \vert\vert X_i - (xe^T + Q\Theta)\vert\vert_2^2$. The optimal x is $\overline{x}_i$ – the mean of all the x_{i_j}'s.
Q	Q is another term in the minimzation (but what is it???). It is of d columns. Optimal Q is given by the d singular vectors of $X_i(I - ee^T/k)$ corresponding with the d largest singular values.
Θ	$\Theta = \begin{bmatrix} \vert & & \vert \\ \theta_1 & \cdots & \theta_k \\ \vert & & \vert \end{bmatrix}$ is a matrix consisting of the orthogonal projections of each $x_{i_j} - x$ onto $\mathcal{T}_{x_i}$. Θ is given by $\Theta_i = Q_i^T X_i(I - \frac{1}{k}ee^T) = \begin{bmatrix} \vert & & \vert \\ \theta_1^{(i)} & \cdots & \theta_k^{(i)} \\ \vert & & \vert \end{bmatrix}, \theta_j^{(i)} = Q_i^T(x_{i_j} - \overline{x}_i)$.
$\xi_j^{(i)}$	$\xi_j^{(i)} = (I - Q_i Q_i^T)(x_{i_j} - \overline{x}_i)$ is the global reconstruction error of x_{i_j}, with $x_{i_j} = \overline{x}_i + Q_i \theta_j^{(i)} + \xi_j^{(i)}$.
τ_{i_j}	τ_{i_j} is the feature space coordinate of x_{i_j}. We want $\tau_{i_j} = \overline{\tau_i} + L_i \theta_j^{(i)} + \epsilon_j^{(i)}$, for $j \in \{1, \ldots, k\}$ and $i \in \{1, \ldots, N\}$.
$\overline{\tau_i}$	$\overline{\tau_i}$ is the mean of the τ_{i_j}, for $j \in \{1, \ldots, k\}$.
L_i	L_i is the local affine transformation for τ_i.
E_i	$E_i = \begin{bmatrix} \vert & & \vert \\ \epsilon_1^{(i)} & \cdots & \epsilon_k^{(i)} \\ \vert & & \vert \end{bmatrix}$ is the matrix form of the local reconstruction error.
e	$e = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}^T$ is a column vector of all ones of dimension k.
T_i	$T_i = \begin{bmatrix} \vert & & \vert \\ \tau_i & \cdots & \tau_k \\ \vert & & \vert \end{bmatrix}$ is the matrix form of the feature space coordinates. It satisfies $T_i = \frac{1}{k} T_i ee^T + L_i \Theta_i + E_i$, so $E_i = T_i(I_k - \frac{1}{k}ee^T) - L_i \Theta_i$.
S_i	S_i is the 0-1 selection matrix such that $TS_i = T_i$.
S	$S = \left[\begin{array}{c:c:c} S_1 & \cdots & S_N \end{array}\right]$.
W_i	$W_i = (I_k - \frac{1}{k}ee^T)(I - \Theta_i^+ \Theta_i)$.

W	$W = \text{diag}(W_1, \ldots, W_N) = \left[\begin{array}{c:c:c:c} W_1 & 0 & \cdots & 0 \\ \hdashline 0 & W_2 & \cdots & 0 \\ \hdashline \vdots & \vdots & \ddots & \vdots \\ \hdashline 0 & 0 & \cdots & W_N \end{array}\right].$
T	We want to find T to minimize the overall reconstruction error $\sum_i \lvert\lvert E_i \rvert\rvert_F^2 = \lvert\lvert TSW \rvert\rvert_F^2$. To determine T, we require $TT^T = I_d$.
B	$B \equiv SWW^TS^T$ has eigenvector e corresponding to eigenvalue 0. Thus, optimal T is given by the d eigenvectors corresponding to the 2nd to $d+1$st smallest eigenvalues.