Curvature-aware Regularization on Riemannian Submanifolds

Kwang In Kim,[1,2] James Tompkin,[1,3] and Christian Theobalt[1]
Max-Planck-Institut für Informatik,[1] Lancaster University,[2] Intel Visual Computing Institute[3]

Abstract
One fundamental assumption in pattern classification problems is that the data generation process lies on a manifold. This holds true for several algorithms for diffusion and regularization, e.g., in graph-Laplacian-based algorithms. Existing algorithms can be improved if we additionally account for how the manifold is embedded within the ambient space — if we consider the extrinsic geometry of the manifold. We characterize the extrinsic curvature of a manifold, and use this in anisotropic diffusion and regularization. The resulting re-weighted graph Laplacian demonstrates superior performance over classical graph Laplacian in semi-supervised learning and spectral clustering.

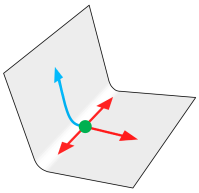

Figure 1: Controlling diffusivity depending on curvature: diffusivity is large along *flat* paths (red); small along the *curved* path (blue).

Anisotropic diffusion on manifolds:

$$\frac{\partial f}{\partial t} = -\Delta_D f := \operatorname{div} D \operatorname{grad} f,$$

D: a positive definite (p.d.) operator that controls the strength and direction of diffusion.

Characterizing curvature on a sub-manifold $M \subset \mathbb{R}^n$ — the second fundamental form:

$$II = \sum_{r,s=1}^{m} \sum_{i=m+1}^{n} \left[\left(\frac{\partial^2 y^i}{\partial x^r \partial x^s} \right) dx^r dx^s \right] Y_i,$$

$\{x^1, \ldots, x^m\}$ and $\{y^1, \ldots, y^n\}$: local coordinates in M and $\mathbb{R}^n$, respectively. Embedding:

$$y^i = y^i(x^1, \ldots, x^m) \text{ for } i = 1, \ldots, n,$$

and $\{\frac{\partial}{\partial y^1}, \ldots, \frac{\partial}{\partial y^m}\} = \{Y_1, \ldots, Y_m\}$.

Constructing D from II — the shape operator:

$$s = \sum_{r,s,\delta=1}^{m} \sum_{i=m+1}^{n} [|H^i|_P]_{rs} \, g^{r\delta} \partial_\delta dx^s,$$

H^i: Hessian in $\{x^i\}$; $|A|_P$: a p.d. version of a matrix A.
s expands the input vector into direction of high curvature.
Our vector-valued diffusivity operator D_p at point p:

$$D_p = (S_p + I)^{-1},$$

S_p is a matrix representation of s.
Scalar-valued diffusivity operator d_p:

$$d_p(Z_p) = \|D_p(Z_p)\| / \|Z_p\|.$$

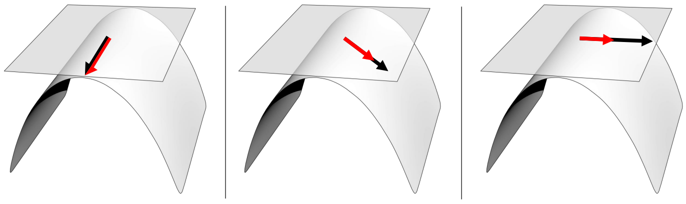

Figure 2: Applying the diffusivity operator D_p: (Left) The black input vector is orthogonal to the direction where M has no curvature, and so the red output vector is identical; (Right) Input parallel to the maximally curved direction of M causes maximum output shrinkage. (Middle) In general, the input vector is shrunk depending on how M is curved.

Practical algorithm — re-weighted graph Laplacian:
Hessian H^i estimated based on locally fitting quadratic polynomials. Scalar-valued diffusivity operator d_p scales the weight (adjacency) matrix W in the standard graph Laplacian:

$$L = G - W,$$

G: column sum of W.

Results:
Theorem: estimated II converges to analytic version.

Table 1: Classification performance (error rate) of graph Laplacian (Lap) and re-weighted graph Laplacian (r-Lap).

Algorithm	USPS	COIL2	BCI	Text	C-PASCAL
Lap	6.72	0.47	37.19	22.3	10.63
r-Lap	**5.78**	**0.41**	**35.67**	**20.8**	**9.83**
Improvement (%)	14.00	12.77	4.09	6.73	6.02
Lap (GT)	5.92	**0**	32.60	20.9	8.89
r-Lap (GT)	**4.94**	**0**	**25.94**	**19.9**	**8.20**
Improvement (%)	15.55	0	20.43	4.79	7.40

Table 2: Clustering performance of Lap and r-Lap; m: manifold dimensionality.

Algorithm	Lap	r-Lap		
		m	Error rate	Improvement (%)
USPS	0.22	2	0.23	-4.54
		3	0.28	-27.27
		4	**0.15**	31.82
		5	0.21	4.54
		6	0.24	-9.09
MNIST	0.31	2	**0.19**	38.71
		3	0.21	32.26
		4	0.32	-3.23
		5	0.25	19.35
		6	0.32	-3.23

max planck institut informatik
LANCASTER UNIVERSITY
VISUAL COMPUTING INSTITUTE