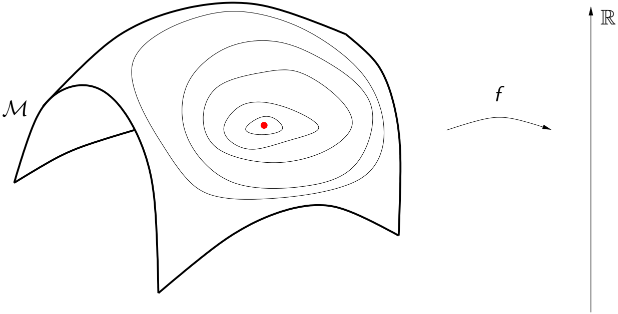
Optimization and curve fitting on manifolds

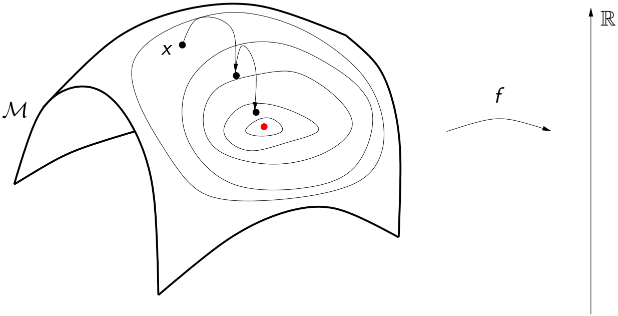
Pierre-Antoine Absil
(Dept. of Mathematical Engineering, UCLouvain)

Journée conjointe des GDR MIA et ISIS
Optimisation Géométrique sur les Variétés
21st November 2014

Optimization on Manifolds in one picture

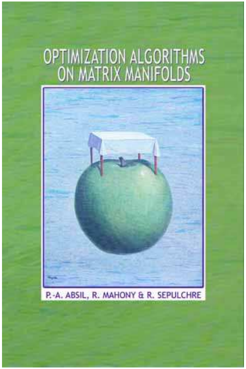


Optimization on Manifolds in one picture



A book

http://press.princeton.edu/titles/8586.html

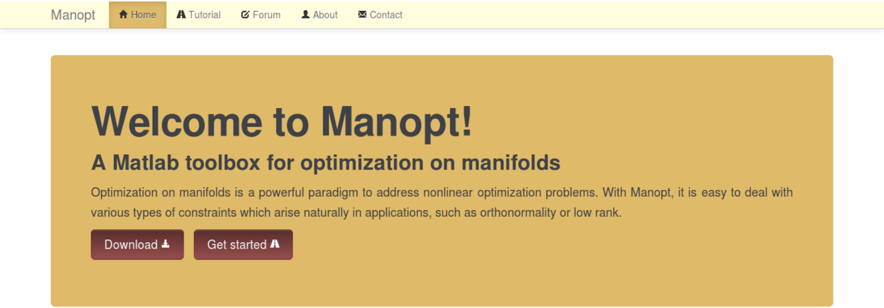


Optimization Algorithms on Matrix Manifolds
P.-A. Absil, R. Mahony, R. Sepulchre
Princeton University Press, January 2008

1. Introduction
2. Motivation and applications
3. Matrix manifolds: first-order geometry
4. Line-search algorithms
5. Matrix manifolds: second-order geometry
6. Newton's method
7. Trust-region methods
8. A constellation of superlinear algorithms

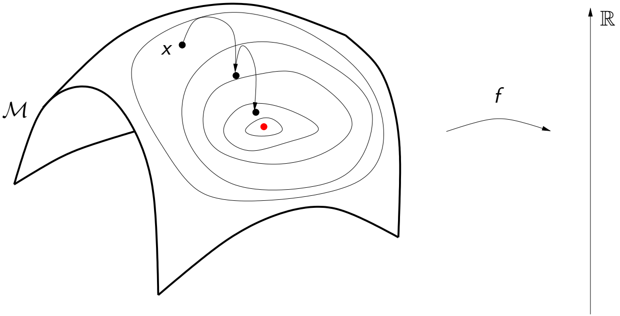
A toolbox

http://www.manopt.org/



Ref: Nicolas Boumal et al, *Manopt, a Matlab toolbox for optimization on manifolds*, JMLR 15(Apr) 1455-1459, 2014.

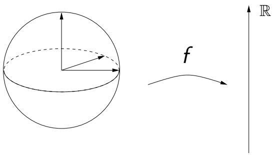
Optimization on Manifolds in one picture



Why general manifolds? – Motivating examples

Given $A = A^T \in \mathbb{R}^{n\times n}$
and $N = \mathrm{diag}(p, p-1, \ldots, 1)$,

$$\min f(X) = -\operatorname{trace}(X^T A X N)$$
subj. to $X \in \mathbb{R}^{n\times p} : X^T X = I$

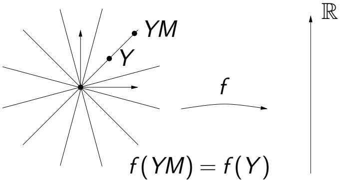


Feasible set: $\mathrm{St}(p, n)$
$= \{X \in \mathbb{R}^{n\times p} : X^T X = I\}$

Embedded submanifold

Given $A = A^T \in \mathbb{R}^{n\times n}$,

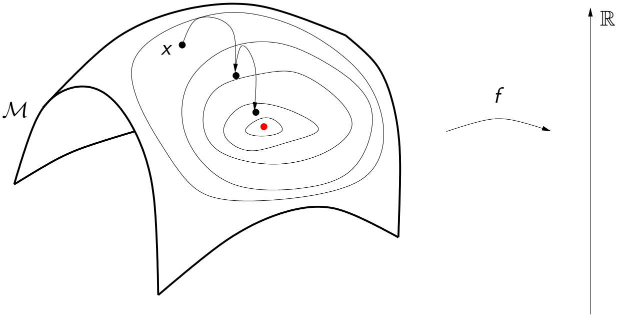
$$\min f(Y) = -\operatorname{trace}\left((Y^T Y)^{-1}(Y^T A Y)\right)$$
subj. to $Y \in \mathbb{R}^{n\times p}_*$ (i.e., Y full rank)



Feasible set: $\mathrm{Gr}(p, n)$
$= \left\{\{YM : M \in \mathbb{R}^{p\times p}_*\} : Y \in \mathrm{R}^{n\times p}_*\right\}$

Quotient manifold

Optimization on Manifolds in one picture



Specific manifolds, and where they appear

- Stiefel manifold $\mathrm{St}(p, n)$ and orthogonal group $O_p = \mathrm{St}(n, n)$

$$\mathrm{St}(p, n) = \{X \in \mathbb{R}^{n \times p} : X^T X = I_p\}$$

  Applications: computer vision; principal component analysis; independent component analysis...

- Grassmann manifold $\mathrm{Gr}(p, n)$

  Set of all p-dimensional subspaces of $\mathbb{R}^n$

  Applications: various dimension reduction problems...

- Set of fixed-rank PSD matrices $\mathrm{S}_+(p, n)$. A quotient representation:

$$X \sim Y \Leftrightarrow \exists Q \in O_p : Y = XQ$$

  Applications: Low-rank approximation of symmetric matrices; algorithms for (large-scale) semidefinite programming...

Specific manifolds, and where they appear

- Low-rank manifold $\mathbb{R}^{m\times n}_{\mathrm{rk}p}$

$$\mathbb{R}^{m\times n}_{\mathrm{rk}p} = \{M \in \mathbb{R}^{m\times n} : \mathrm{rk}(M) = p\}$$

Applications: dimensionality reduction; model for matrix completion...

- Shape manifold $O_n \backslash \mathbb{R}^{n\times p}_*$

$$Y \sim X \Leftrightarrow \exists U \in O_n : Y = UX$$

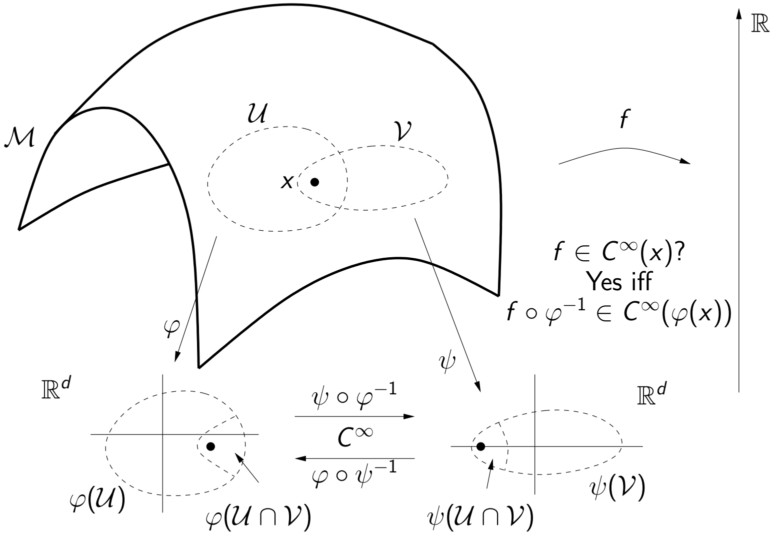
Applications: shape analysis

- Oblique manifold $\mathbb{R}^{n\times p}_* / \mathcal{S}_{\mathrm{diag}+}$

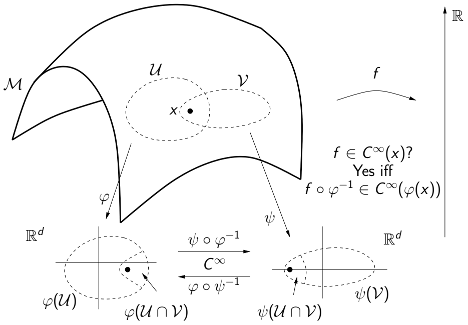
$$\mathbb{R}^{n\times p}_* / \mathcal{S}_{\mathrm{diag}+} \simeq \{Y \in \mathbb{R}^{n\times p}_* : \mathrm{diag}(Y^T Y) = I_p\}$$

Applications: blind source separation; factor analysis (oblique Procrustes problem)...

Smooth optimization problems on general manifolds



Optimization on manifolds in its most abstract formulation



Given:

- A set $\mathcal{M}$ endowed (explicitly or implicitly) with a manifold structure (i.e., a collection of compatible charts).
- A function $f : \mathcal{M} \to \mathbb{R}$, smooth in the sense of the manifold structure.

Task: Compute a local minimizer of f.

Algorithms formulated on abstract manifolds

- Steepest-descent
  Needs: Riemannian structure and retraction
- Newton
  Needs: affine connection and retraction
- Conjugate Gradients
  Needs: Riemannian structure, retraction, and vector transport
- BFGS
  Needs: needs Riemannian structre, retraction, and vector transport
- Trust Region
  Needs: Riemannian structure and retraction

Steepest descent on abstract manifolds

Required: Riemannian manifold $\mathcal{M}$; retraction R on $\mathcal{M}$.
Iteration $x_k \in \mathcal{M} \mapsto x_{k+1} \in \mathcal{M}$ defined by

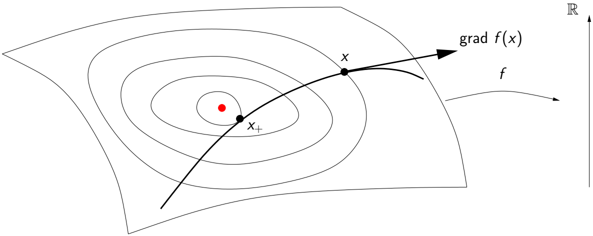
1. Compute steepest-descent direction in $T_{x_k}\mathcal{M}$:

$$\eta_k = -\operatorname{grad} f(x_k).$$

2. Set

$$x_{k+1} := R_{x_k}(t_k \eta_k)$$

where t_k is chosen using a line-search method.



Newton on abstract manifolds

Required: Riemannian manifold $\mathcal{M}$; retraction R on $\mathcal{M}$; affine connection ∇ on $\mathcal{M}$; real-valued function f on $\mathcal{M}$.
Iteration $x_k \in \mathcal{M} \mapsto x_{k+1} \in \mathcal{M}$ defined by

1. Solve the Newton equation

$$\text{Hess } f(x_k)\eta_k = -\text{grad } f(x_k)$$

   for the unknown $\eta_k \in T_{x_k}\mathcal{M}$, where

$$\text{Hess } f(x_k)\eta_k := \nabla_{\eta_k} \text{grad } f.$$

2. Set

$$x_{k+1} := R_{x_k}(\eta_k).$$

Newton on submanifolds of $\mathbb{R}^n$

Required: Riemannian submanifold $\mathcal{M}$ of $\mathbb{R}^n$; retraction R on $\mathcal{M}$; real-valued function f on $\mathcal{M}$.

Iteration $x_k \in \mathcal{M} \mapsto x_{k+1} \in \mathcal{M}$ defined by

1. Solve the Newton equation

$$\text{Hess } f(x_k)\eta_k = -\text{grad } f(x_k)$$

   for the unknown $\eta_k \in T_{x_k}\mathcal{M}$, where

$$\text{Hess } f(x_k)\eta_k := \mathrm{P}_{T_{x_k}\mathcal{M}} \mathrm{D} \, \text{grad } f(x_k)[\eta_k].$$

2. Set

$$x_{k+1} := R_{x_k}(\eta_k).$$

Newton on the unit sphere S^{n-1}

Required: real-valued function f on S^{n-1}.
Iteration $x_k \in S^{n-1} \mapsto x_{k+1} \in S^{n-1}$ defined by

1. Solve the Newton equation

$$\begin{cases} \mathrm{P}_{x_k} \mathrm{D}(\operatorname{grad} f)(x_k)[\eta_k] = -\operatorname{grad} f(x_k) \\ x^T \eta_k = 0, \end{cases}$$

   for the unknown $\eta_k \in \mathbb{R}^n$, where

$$\mathrm{P}_{x_k} = (I - x_k x_k^T).$$

2. Set

$$x_{k+1} := \frac{x_k + \eta_k}{\|x_k + \eta_k\|}.$$

Newton for Rayleigh quotient optimization on unit sphere

Iteration $x_k \in S^{n-1} \mapsto x_{k+1} \in S^{n-1}$ defined by

1. Solve the Newton equation

$$\begin{cases} \mathrm{P}_{x_k} A \, \mathrm{P}_{x_k} \eta_k - \eta_k x_k^T A x_k = -\mathrm{P}_{x_k} A x_k, \\ x_k^T \eta_k = 0, \end{cases}$$

   for the unknown $\eta_k \in \mathbb{R}^n$, where

$$\mathrm{P}_{x_k} = (I - x_k x_k^T).$$

2. Set

$$x_{k+1} := \frac{x_k + \eta_k}{\|x_k + \eta_k\|}.$$

Conjugate Gradients on abstract manifolds

Require: Riemannian manifold $\mathcal{M}$; vector transport $\mathcal{T}$ on $\mathcal{M}$ with associated retraction R; real-valued function f on $\mathcal{M}$; initial iterate $x_0 \in \mathcal{M}$.

1: Set $\eta_0 = -\operatorname{grad} f(x_0)$.

2: **for** $k = 0, 1, 2, \ldots$ **do**

Compute a step size α_k and set

$$x_{k+1} = R_{x_k}(\alpha_k \eta_k). \tag{1}$$

Compute β_{k+1} and set

$$\eta_{k+1} = -\operatorname{grad} f(x_{k+1}) + \beta_{k+1} \mathcal{T}_{\alpha_k \eta_k}(\eta_k). \tag{2}$$

5: **end for**

Fletcher-Reeves: $\beta_{k+1} = \frac{\langle \operatorname{grad} f(x_{k+1}), \operatorname{grad} f(x_{k+1}) \rangle}{\langle \operatorname{grad} f(x_k), \operatorname{grad} f(x_k) \rangle}$.

Polak-Ribière: $\beta_{k+1} = \frac{\langle \operatorname{grad} f(x_{k+1}), \operatorname{grad} f(x_{k+1}) - \mathcal{T}_{\alpha_k \eta_k}(\operatorname{grad} f(x_k)) \rangle}{\langle \operatorname{grad} f(x_k), \operatorname{grad} f(x_k) \rangle}$.

Ref: PAA et al [AMS08, §8.3], Sato & Iwai [SI13].

BFGS on abstract manifolds

Given: Riemannian manifold M with Riemannian metric g; vector transport $\mathcal{T}$ on M with associated retraction R; smooth real-valued function f on M; initial iterate $\mathbf{x}_0 \in M$; initial Hessian approximation $\mathcal{B}_0$.

for k = 0, 1, 2, . . . **do**

  Obtain $\eta_k \in T_{\mathbf{x}_k} M$ by solving $\mathcal{B}_k \eta_k = -\operatorname{grad} f(\mathbf{x}_k)$.

  Compute step size α_k and set $\mathbf{x}_{k+1} = R_{\mathbf{x}_k}(\alpha_k \eta_k)$.

  Define $s_k = \mathcal{T}_{\alpha\eta_k}(\alpha\eta_k)$ and $y_k = \operatorname{grad} f(\mathbf{x}_{k+1}) - \mathcal{T}_{\alpha\eta_k}(\operatorname{grad} f(\mathbf{x}_k))$.

  Define the linear operator $\mathcal{B}_{k+1} : T_{\mathbf{x}_{k+1}} M \to T_{\mathbf{x}_{k+1}} M$ by

$$\mathcal{B}_{k+1} p = \tilde{\mathcal{B}}_k p - \frac{g(s_k, \tilde{\mathcal{B}}_k p)}{g(s_k, \tilde{\mathcal{B}}_k s_k)} \tilde{\mathcal{B}}_k s_k + \frac{g(y_k, p)}{g(y_k, s_k)} y_k \quad \text{for all } p \in T_{\mathbf{x}_{k+1}} M, \tag{3}$$

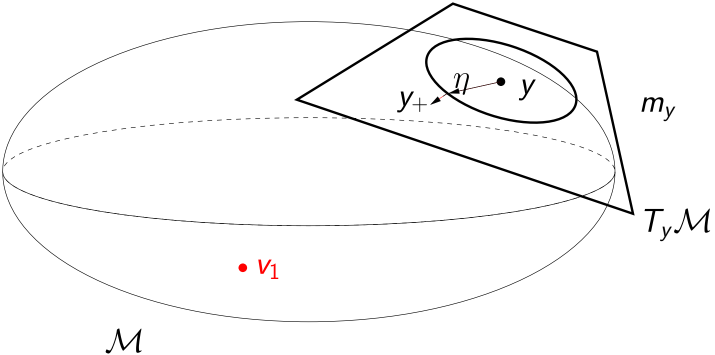
  with

$$\tilde{\mathcal{B}}_k = \mathcal{T}_{\alpha\eta_k} \circ \mathcal{B}_k \circ (\mathcal{T}_{\alpha\eta_k})^{-1}. \tag{4}$$

end for

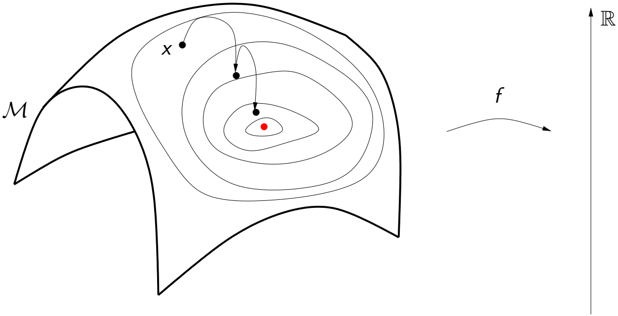
Ref: Qi et al [QGA10], Ring & Wirth [RW12].

Trust region on abstract manifolds

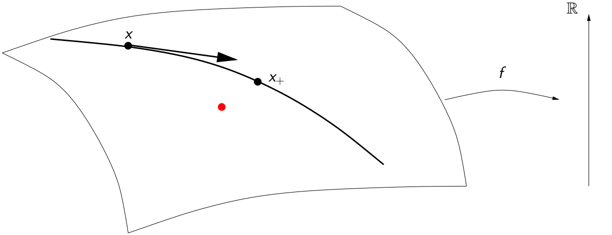


Refs: PAA et al [ABG07], Huang et al [HAG14].

Optimization on Manifolds in one picture



Some classics on Optimization On Manifolds (I)



Luenberger (1973), *Introduction to linear and nonlinear programming*. Luenberger mentions the idea of performing line search along geodesics, "which we would use if it were computationally feasible (which it definitely is not)".

Some classics on Optimization On Manifolds (II)

Gabay (1982), *Minimizing a differentiable function over a differential manifold*. Stepest descent along geodesics; Newton's method along geodesics; Quasi-Newton methods along geodesics.

Smith (1994), *Optimization techniques on Riemannian manifolds*. Levi-Civita connection ∇; Riemannian exponential; parallel translation. But Remark 4.9: If Algorithm 4.7 (Newton's iteration on the sphere for the Rayleigh quotient) is simplified by replacing the exponential update with the update

$$x_{k+1} = \frac{x_k + \eta_k}{\|x_k + \eta_k\|}$$

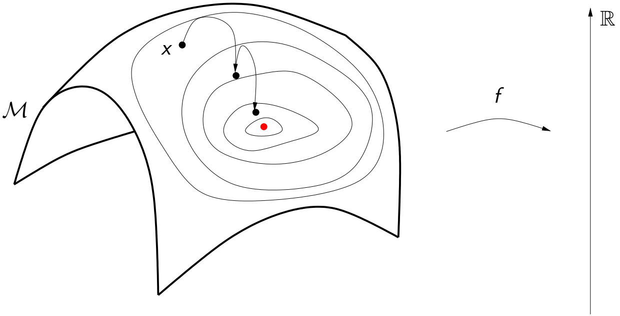
then we obtain the Rayleigh quotient iteration.

Some classics on Optimization On Manifolds (III)

Manton (2002), *Optimization algorithms exploiting unitary constraints* "The present paper breaks with tradition by not moving along geodesics". The geodesic update $\text{Exp}_x \eta$ is replaced by a projective update $\pi(x + \eta)$, the *projection* of the point $x + \eta$ onto the manifold.

Adler, Dedieu, Shub, et al. (2002), *Newton's method on Riemannian manifolds and a geometric model for the human spine*. The exponential update is relaxed to the general notion of *retraction*. The geodesic can be replaced by any (smoothly prescribed) curve tangent to the search direction.

Optimization on Manifolds in one picture

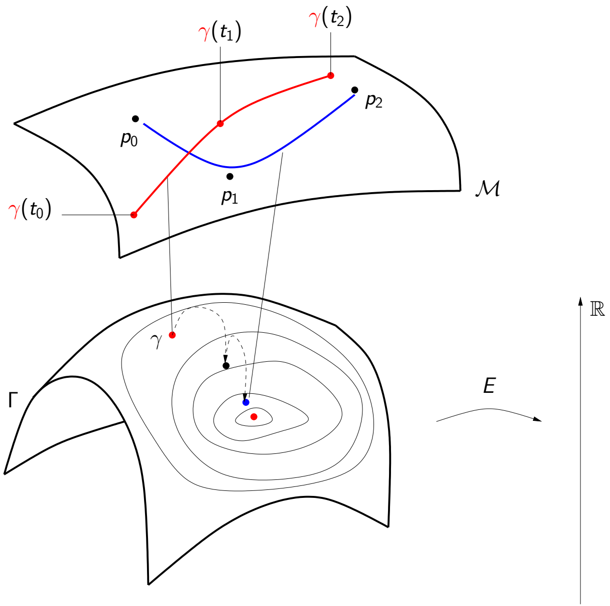


An Application: Curve Fitting

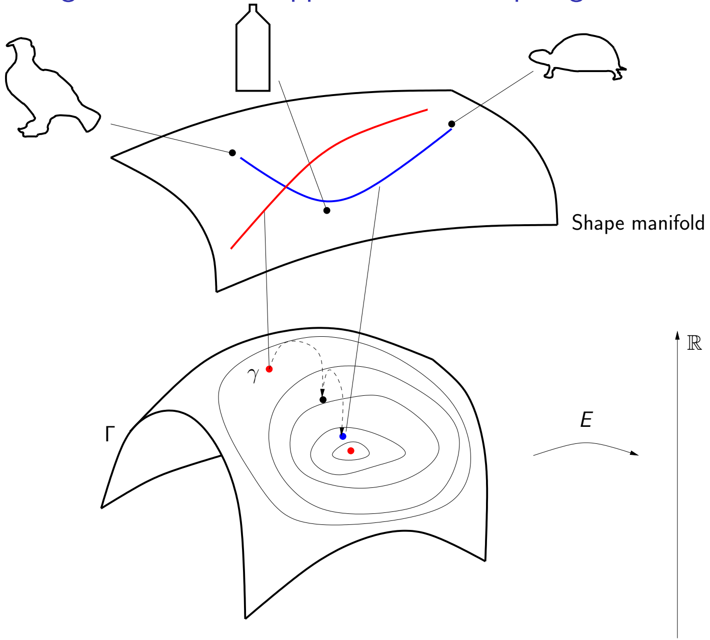
Sources

- Nonparametric curve fitting on manifolds:
  - Chafik Samir, PAA, Anuj Srivastava, Eric Klassen, *A gradient-descent method for curve fitting on Riemannian manifolds*, Foundations of Computational Mathematics, 12(1), pp. 49-73, 2012.
  - Nicolas Boumal, PAA, *Discrete regression methods on the cone of positive-definite matrices*, ICASSP 2011.
  - Nicolas Boumal, PAA, *A discrete regression method on manifolds and its application to data on SO(n)*, IFAC World Congress 2011.
- Parametric curve fitting on manifolds (see Pierre-Yves's talk):
  - C. Samir, P. Van Dooren, D. Laurent, K. A. Gallivan, PAA, *Elastic morphing of 2D and 3D objects on a shape manifold*, Lecture Notes in Computer Science, Volume 5627/2009, pp. 563-572, 2009
  - Pierre-Yves Gousenbourger, Chafik Samir, PAA, *Piecewise-Bézier C^1 interpolation on Riemannian manifolds with application to 2D shape morphing*, ICPR 2014
  - Antoine Arnould, Pierre-Yves Gousenbourger, Chafik Samir, PAA, *Fitting Smooth Paths on Riemannian manifolds: Endometrial Surface Reconstruction and Preoperative MRI-Based Navigation*, submitted.

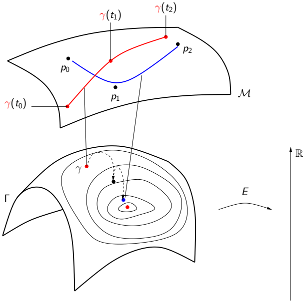
Curve fitting on manifolds



Curve fitting on manifolds: application to morphing



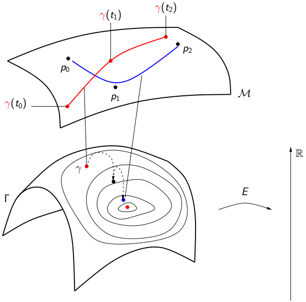
Curve fitting on manifolds: possible applications



Applications in *noise reduction*, *resampling*, and *trajectory generation*.

- Evolution of the paleomagnetic north pole, as in Jupp and Kent [JK87]: $\mathcal{M} = \mathrm{S}^2$, the sphere.
- Rigid body motion: $\mathcal{M} = \mathrm{SE}(3)$, the special Euclidean group.
- Diffusion-Tensor MRI: $\mathcal{M} = \mathcal{S}_3^{++}$, the set of all 3×3 symmetric positive-definite matrices.
- Morphing: $\mathcal{M}$ is a shape manifold.
- ...

Curve fitting on manifolds: problem considered



Given: Riemannian manifold $\mathcal{M}$; $p_0, \dots, p_N \in \mathcal{M}$;
$0 = t_0 < \cdots < t_N = 1$.
Goal: find the curve $\gamma : [0,1] \mapsto \mathcal{M}$ that minimizes

$$E_2 : \Gamma_2 \to \mathbb{R} : E_2(\gamma) = E_{\mathrm{d}}(\gamma) + \lambda E_{\mathrm{s},2}(\gamma)$$
$$= \frac{1}{2}\sum_{i=0}^{N} d^2(\gamma(t_i), p_i) + \frac{\lambda}{2}\int_0^1 \left\langle \frac{\mathrm{D}^2\gamma}{\mathrm{d}t^2}, \frac{\mathrm{D}^2\gamma}{\mathrm{d}t^2} \right\rangle \mathrm{d}t,$$

where Γ_2 is the Sobolev space $H^2([0,1], \mathcal{M})$.

Previous work

Machado and Silva Leite [ML06, Mac06] consider

$$E_2 : \Gamma_2 \to \mathbb{R} : E_2(\gamma) = \frac{1}{2}\sum_{i=0}^{N} d^2(\gamma(t_i), p_i) + \frac{\lambda}{2}\int_0^1 \langle \frac{\mathrm{D}^2\gamma}{\mathrm{d}t^2}, \frac{\mathrm{D}^2\gamma}{\mathrm{d}t^2}\rangle \,\mathrm{d}t,$$

and obtain the Euler-Lagrange equations (stationarity conditions):

On each subinterval,

$$\frac{\mathrm{D}^4\gamma}{\mathrm{d}t^4} + R\left(\frac{\mathrm{D}^2\gamma}{\mathrm{d}t^2}, \dot{\gamma}\right)\dot{\gamma} = 0,$$

and at the knot points,

$$\frac{\mathrm{D}^k\gamma}{\mathrm{d}t^k}(t_i^+) - \frac{\mathrm{D}^k\gamma}{\mathrm{d}t^k}(t_i^-) = \begin{cases} 0, & k = 0, 1, \quad (i = 1, \dots, N-1) \\ 0, & k = 2, \quad (i = 0, \dots, N) \\ \frac{1}{\lambda}\,\mathrm{Exp}^{-1}_{\gamma(t_i)}(p_i), & k = 3, \quad (i = 0, \dots, N) \end{cases},$$

with

$$\frac{\mathrm{D}^2\gamma}{\mathrm{d}t^2}(t_0^-) = \frac{\mathrm{D}^3\gamma}{\mathrm{d}t^3}(t_0^-) = \frac{\mathrm{D}^2\gamma}{\mathrm{d}t^2}(t_N^+) = \frac{\mathrm{D}^3\gamma}{\mathrm{d}t^3}(t_N^+) = 0.$$

Gradient-descent for discretized E_2

Objective: $E_2(\gamma) = \frac{1}{2}\sum_{i=0}^{N} d^2(\gamma(t_i), p_i) + \frac{\lambda}{2}\int_0^1 \langle \frac{\mathrm{D}^2\gamma}{\mathrm{d}t^2}, \frac{\mathrm{D}^2\gamma}{\mathrm{d}t^2} \rangle \,\mathrm{d}t.$

- Finite differences in $\mathbb{R}^n$:

$$\ddot{x}_0 = \frac{2}{\Delta t_{\mathrm{f}} + \Delta t_{\mathrm{b}}} \frac{1}{\Delta t_{\mathrm{f}} \Delta t_{\mathrm{b}}} \left[\Delta t_{\mathrm{b}}(x_{\mathrm{f}} - x_0) + \Delta t_{\mathrm{f}}(x_{\mathrm{b}} - x_0)\right] + \mathcal{O}(\Delta t) \tag{5}$$

Gradient-descent for discretized E_2

Objective: $E_2(\gamma) = \frac{1}{2}\sum_{i=0}^{N} d^2(\gamma(t_i), p_i) + \frac{\lambda}{2}\int_0^1 \langle \frac{D^2\gamma}{dt^2}, \frac{D^2\gamma}{dt^2} \rangle \, dt.$

- Finite differences in $\mathbb{R}^n$:

$$\ddot{x}_0 = \frac{2}{\Delta t_\mathrm{f} + \Delta t_\mathrm{b}} \frac{1}{\Delta t_\mathrm{f} \Delta t_\mathrm{b}} \left[\Delta t_\mathrm{b}(x_\mathrm{f} - x_0) + \Delta t_\mathrm{f}(x_\mathrm{b} - x_0)\right] + \mathcal{O}(\Delta t) \tag{5}$$

- Finite differences on a manifold:

$$\ddot{x}_0 \approx \frac{2}{\Delta t_\mathrm{f} + \Delta t_\mathrm{b}} \frac{1}{\Delta t_\mathrm{f} \Delta t_\mathrm{b}} \left[\Delta t_\mathrm{b} \operatorname{Log}_{x_0}(x_\mathrm{f}) + \Delta t_\mathrm{f} \operatorname{Log}_{x_0}(x_\mathrm{b})\right] \tag{6}$$

Gradient-descent for discretized E_2

Objective: $E_2(\gamma) = \frac{1}{2}\sum_{i=0}^{N} d^2(\gamma(t_i), p_i) + \frac{\lambda}{2}\int_0^1 \langle \frac{D^2\gamma}{dt^2}, \frac{D^2\gamma}{dt^2}\rangle \, dt.$

- Finite differences in $\mathbb{R}^n$:

$$\ddot{x}_0 = \frac{2}{\Delta t_\mathrm{f} + \Delta t_\mathrm{b}} \frac{1}{\Delta t_\mathrm{f} \Delta t_\mathrm{b}} \left[\Delta t_\mathrm{b}(x_\mathrm{f} - x_0) + \Delta t_\mathrm{f}(x_\mathrm{b} - x_0)\right] + \mathcal{O}(\Delta t) \tag{5}$$

- Finite differences on a manifold:

$$\ddot{x}_0 \approx \frac{2}{\Delta t_\mathrm{f} + \Delta t_\mathrm{b}} \frac{1}{\Delta t_\mathrm{f} \Delta t_\mathrm{b}} \left[\Delta t_\mathrm{b} \operatorname{Log}_{x_0}(x_\mathrm{f}) + \Delta t_\mathrm{f} \operatorname{Log}_{x_0}(x_\mathrm{b})\right] \tag{6}$$

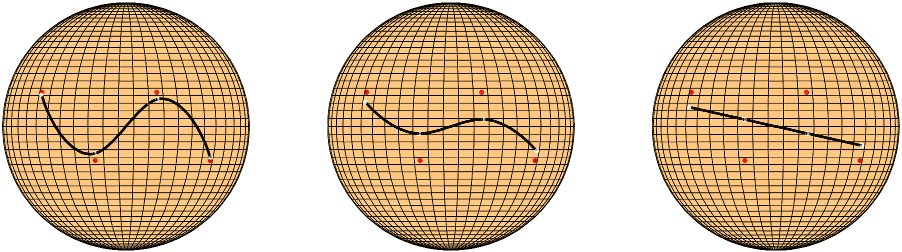
- Discretized E_2:

$$\hat{E}_2 : \mathcal{M}^{N_\mathrm{d}} \to \mathbb{R} : \hat{E}_2(\gamma) = \frac{1}{2}\sum_{i=0}^{N} d^2(p_i, \gamma_{s_i}) + \frac{\lambda}{2}\sum_{i=1}^{N_\mathrm{d}} \beta_i \|a_i\|_{\gamma_i}^2$$

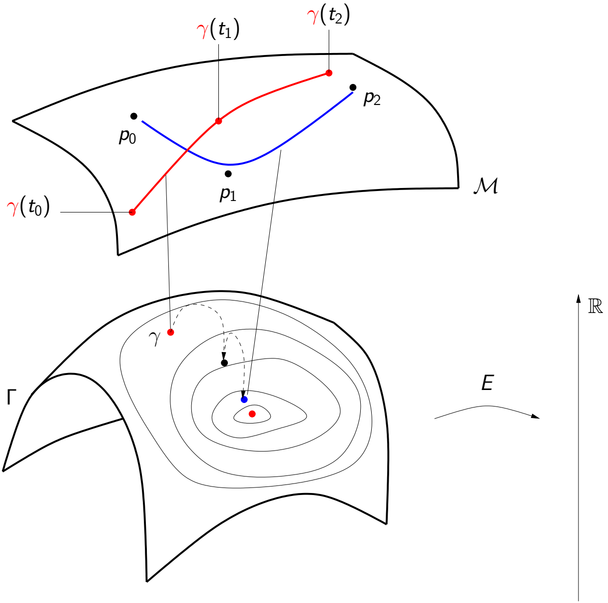
Illustrations on the sphere

Objective: $E_2(\gamma) = \frac{1}{2}\sum_{i=0}^{N} d^2(\gamma(t_i), p_i) + \frac{\lambda}{2}\int_0^1 \langle \frac{\mathrm{D}^2\gamma}{\mathrm{d}t^2}, \frac{\mathrm{D}^2\gamma}{\mathrm{d}t^2} \rangle \,\mathrm{d}t.$

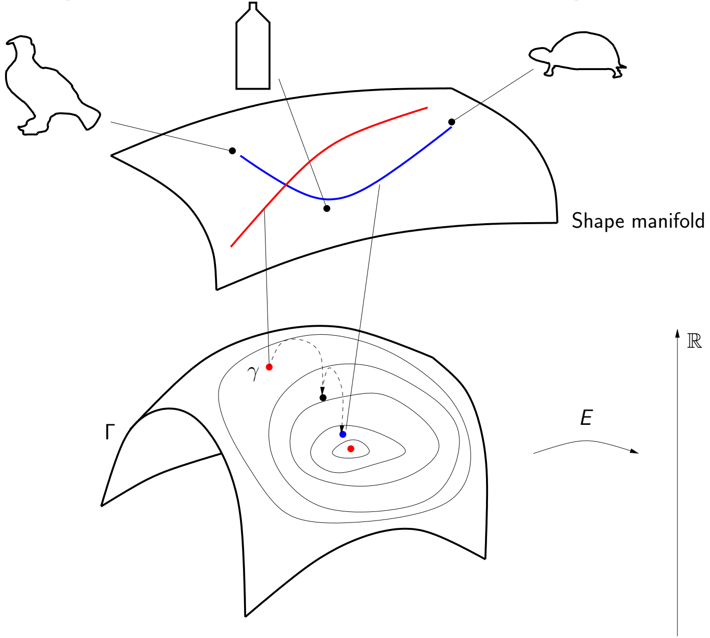
$\lambda = 10^{-4}$ $\lambda = 10^{-3}$ $\lambda = 10^{0}$



Curve fitting on manifolds



Curve fitting on manifolds: application to morphing



Polynomial interpolation on manifolds

- Polynomial interpolation reminder: Given $(t_0, x_0), \dots, (t_n, x_n)$ in $\mathbb{R}^d$, there is one and only one polynomial p_n of degree at most n such that $p(t_k) = x_k$, $k = 0, \dots, n$.

Polynomial interpolation on manifolds

- Polynomial interpolation reminder: Given $(t_0, x_0), \dots, (t_n, x_n)$ in $\mathbb{R}^d$, there is one and only one polynomial p_n of degree at most n such that $p(t_k) = x_k$, $k = 0, \dots, n$.
- $p_n(t)$ can be computed with Neville's algorithm, based on the formula

$$P_{i,j}(t) = P_{i,j-1}(t) + \frac{t - t_i}{t_j - t_i} \left(P_{i+1,j}(t) - P_{i,j-1}(t) \right), \qquad (7)$$

  where $P_{i,j}$ stands for the polynomial of degree at most $j - i$ that interpolates $(t_i, x_i), \dots, (t_j, x_j)$. We have $p_n(t) = P_{0,n}(t)$.

Polynomial interpolation on manifolds

- Polynomial interpolation reminder: Given $(t_0, x_0), \ldots, (t_n, x_n)$ in $\mathbb{R}^d$, there is one and only one polynomial p_n of degree at most n such that $p(t_k) = x_k$, $k = 0, \ldots, n$.
- $p_n(t)$ can be computed with Neville's algorithm, based on the formula

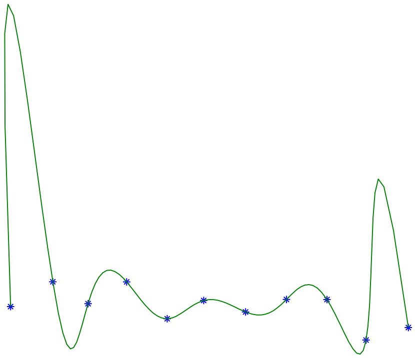
$$P_{i,j}(t) = P_{i,j-1}(t) + \frac{t - t_i}{t_j - t_i}\left(P_{i+1,j}(t) - P_{i,j-1}(t)\right), \qquad (7)$$

  where $P_{i,j}$ stands for the polynomial of degree at most $j - i$ that interpolates $(t_i, x_i), \ldots, (t_j, x_j)$. We have $p_n(t) = P_{0,n}(t)$.
- When $x_0, \ldots, x_n$ are on a manifold, (7) readily generalizes to

$$P_{i,j}(t) = \mathrm{Exp}_{P_{i,j-1}(t)}\left(\frac{t - t_i}{t_j - t_i}\,\mathrm{Log}_{P_{i,j-1}(t)}\, P_{i+1,j}(t)\right).$$

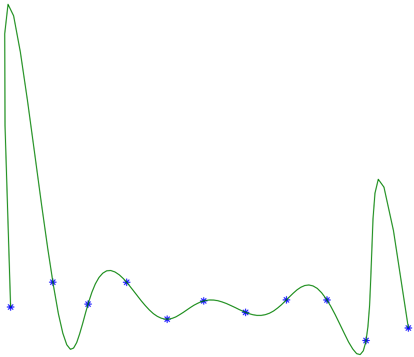
Piecewise-polynomial interpolation on manifolds

- Polynomial interpolation on manifolds is prone to the Runge phenomenon.



Piecewise-polynomial interpolation on manifolds

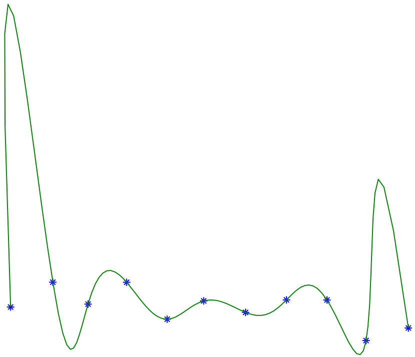
- Polynomial interpolation on manifolds is prone to the Runge phenomenon.



- Polynomial interpolation on manifolds is also prone to a Runge-like effect!

Piecewise-polynomial interpolation on manifolds

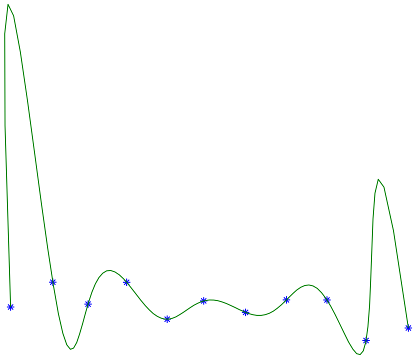
- Polynomial interpolation on manifolds is prone to the Runge phenomenon.



- Polynomial interpolation on manifolds is also prone to a Runge-like effect!
- Remedy: Piecewise-polynomial interpolation on manifolds.

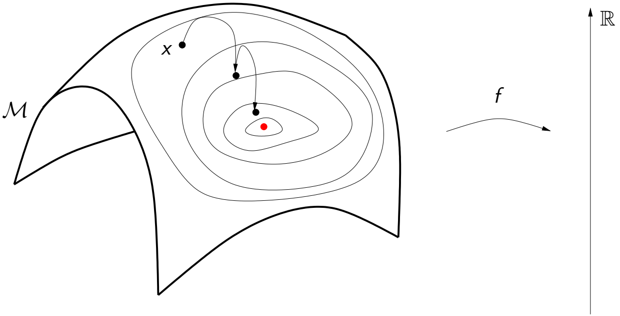
Piecewise-polynomial interpolation on manifolds

- Polynomial interpolation on manifolds is prone to the Runge phenomenon.



- Polynomial interpolation on manifolds is also prone to a Runge-like effect!
- Remedy: Piecewise-polynomial interpolation on manifolds.
- See Pierre-Yves Gousenbourger's talk later today.

Optimization on Manifolds in one picture



- P.-A. Absil, C. G. Baker, and K. A. Gallivan, *Trust-region methods on Riemannian manifolds*, Found. Comput. Math. **7** (2007), no. 3, 303–330.
- P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization algorithms on matrix manifolds*, Princeton University Press, Princeton, NJ, 2008.
- Wen Huang, P.-A. Absil, and K. A. Gallivan, *A riemannian symmetric rank-one trust-region method*, Mathematical Programming (2014), accepted for publication.
- Peter E. Jupp and John T. Kent, *Fitting smooth paths to spherical data*, J. Roy. Statist. Soc. Ser. C **36** (1987), no. 1, 34–46. MR MR887825 (88f:62066)
- Luís Miguel Faustino Machado, *Least squares problems on Riemannian manifolds*, Ph.D. thesis, Department of Mathematics, University of Coimbra, 2006.

- Luís Machado and F. Silva Leite, *Fitting smooth paths on Riemannian manifolds*, Int. J. Appl. Math. Stat. **4** (2006), no. J06, 25–53.
- Chunhong Qi, Kyle A. Gallivan, and P.-A. Absil, *An efficient bfgs algorithm for riemannian optimization*, Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2010), 2010, pp. 2221–2227.
- Wolfgang Ring and Benedikt Wirth, *Optimization methods on Riemannian manifolds and their application to shape space*, SIAM J. Optim. **22** (2012), no. 2, 596–627.
- Hiroyuki Sato and Toshihiro Iwai, *Convergence analysis for the riemannian conjugate gradient method*, 2013, arXiv:1302.0125.