A User's Guide to Riemannian Newton-Type Methods on Manifolds

Felipe Álvarez

Departamento de Ingeniería Matemática
Centro de Modelamiento Matemático (CNRS UMI 2807)
Universidad de Chile

In collaboration with: J. Bolte, J. Munier, J. López

Sixièmes Journées Franco-Chiliennes d'Optimisation
Université du Sud Toulon-Var
Mai 19-21, 2008

http://www.dim.uchile.cl/~falvarez

Motivation: Nonlinear equations in a manifold

Goal: find $p^* \in M$ satisfying $\boxed{F(p^*) = 0 \in T_{p^*}M}$

- M is a connected and n-dimensional differentiable manifold.
- $T_pM \simeq \mathbb{R}^n$ is the tangent space of M at p:

  If $c(t)$ is a curve passing through p at $t = 0$ then $\dot{c}(0) \in T_pM$.
- $F : M \to TM$ is a continuously differentiable vector field:

$$M \ni p \mapsto F(p) \in T_pM$$

Motivation: Nonlinear equations in a manifold

Goal: find $p^* \in M$ satisfying $\boxed{F(p^*) = 0 \in T_{p^*}M}$

- M is a connected and n-dimensional differentiable manifold.
- $T_pM \simeq \mathbb{R}^n$ is the tangent space of M at p:

  If $c(t)$ is a curve passing through p at $t = 0$ then $\dot{c}(0) \in T_pM$.
- $F : M \to TM$ is a continuously differentiable vector field:

$$M \ni p \mapsto F(p) \in T_pM$$

Motivation: Nonlinear equations in a manifold

Goal: find $p^* \in M$ satisfying $\boxed{F(p^*) = 0 \in T_{p^*}M}$

- M is a connected and n-dimensional differentiable manifold.
- $T_pM \simeq \mathbb{R}^n$ is the tangent space of M at p:

  If $c(t)$ is a curve passing through p at $t = 0$ then $\dot{c}(0) \in T_pM$.

- $F : M \to TM$ is a continuously differentiable vector field:

$$M \ni p \mapsto F(p) \in T_pM$$

Motivation: Nonlinear equations in a manifold

Goal: find $p^* \in M$ satisfying $\boxed{F(p^*) = 0 \in T_{p^*}M}$

- M is a connected and n-dimensional differentiable manifold.
- $T_pM \simeq \mathbb{R}^n$ is the tangent space of M at p:

  If $c(t)$ is a curve passing through p at $t = 0$ then $\dot{c}(0) \in T_pM$.

- $F : M \to TM$ is a continuously differentiable vector field:

$$M \ni p \mapsto F(p) \in T_pM$$

Example 1: Rayleigh's quotient on the sphere

- $M = \mathbb{S}^n = \{x \in \mathbb{R}^{n+1} \mid x^T x = 1\}$ (unit sphere in $\mathbb{R}^{n+1}$).
- $T_x M = \{v \in \mathbb{R}^{n+1} \mid x^T v = 0\}$.
- $\boxed{F(x) = Ax - q(x)x}$ with
  - $A \in \mathbb{R}^{n \times n}$ being symmetric and positive definite.
  - $q(x) = x^T A x$.
- $x^T F(x) = 0 \Rightarrow F(x) \in T_x M$.
- $F(x^*) = 0$ iff x^* is an eigenvector of A with $q(x^*)$ the corresponding eigenvalue.

Example 2: Stiefel manifold

- $M = S_{n,k} = \{Y \in \mathbb{R}^{n\times k} \mid Y^T Y = I_k\}$.
- $T_Y S_{n,k} = \{\Delta \in \mathbb{R}^{n\times k} \mid \Delta^T Y + Y^T \Delta = 0\}$.
- If $k = 1$ then $S_{n,1} = \mathbb{S}^{n-1}$.
- If $k = n$ then $S_{n,n} = O_n$ the orthogonal group.
- $T_{I_n} O_n = \{\Delta \in \mathbb{R}^{n\times n} \mid \Delta^T = -\Delta\}$.
- $\dim S_{n,k} = nk - \frac{1}{2}k(k+1)$.
- $\boxed{F(Y) = AY - YY^T AY}$
- $F(Y^*) = 0$ iff the columns of Y^* are eigenvectors of A.

Example 2: Stiefel manifold

- $M = S_{n,k} = \{Y \in \mathbb{R}^{n \times k} \mid Y^T Y = I_k\}$.
- $T_Y S_{n,k} = \{\Delta \in \mathbb{R}^{n \times k} \mid \Delta^T Y + Y^T \Delta = 0\}$.
- If $k = 1$ then $S_{n,1} = \mathbb{S}^{n-1}$.
- If $k = n$ then $S_{n,n} = O_n$ the orthogonal group.
- $T_{I_n} O_n = \{\Delta \in \mathbb{R}^{n \times n} \mid \Delta^T = -\Delta\}$.
- $\dim S_{n,k} = nk - \frac{1}{2}k(k+1)$.
- $\boxed{F(Y) = AY - YY^T AY}$
- $F(Y^*) = 0$ iff the columns of Y^* are eigenvectors of A.

Example 2: Stiefel manifold

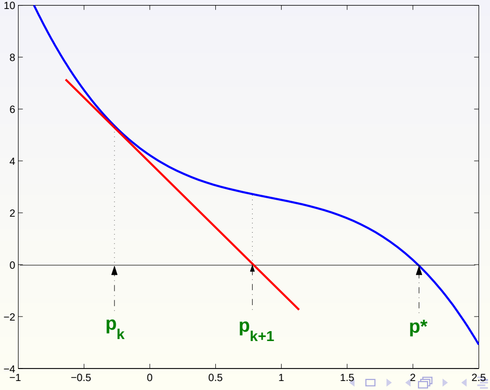
- $M = S_{n,k} = \{Y \in \mathbb{R}^{n\times k} \mid Y^T Y = I_k\}$.
- $T_Y S_{n,k} = \{\Delta \in \mathbb{R}^{n\times k} \mid \Delta^T Y + Y^T \Delta = 0\}$.
- If $k = 1$ then $S_{n,1} = \mathbb{S}^{n-1}$.
- If $k = n$ then $S_{n,n} = O_n$ the orthogonal group.
- $T_{I_n} O_n = \{\Delta \in \mathbb{R}^{n\times n} \mid \Delta^T = -\Delta\}$.
- $\dim S_{n,k} = nk - \frac{1}{2}k(k+1)$.
- $F(Y) = AY - YY^T AY$
- $F(Y^*) = 0$ iff the columns of Y^* are eigenvectors of A.

Example 2: Stiefel manifold

- $M = S_{n,k} = \{Y \in \mathbb{R}^{n\times k} \mid Y^T Y = I_k\}$.
- $T_Y S_{n,k} = \{\Delta \in \mathbb{R}^{n\times k} \mid \Delta^T Y + Y^T \Delta = 0\}$.
- If $k = 1$ then $S_{n,1} = \mathbb{S}^{n-1}$.
- If $k = n$ then $S_{n,n} = O_n$ the orthogonal group.
- $T_{I_n} O_n = \{\Delta \in \mathbb{R}^{n\times n} \mid \Delta^T = -\Delta\}$.
- $\dim S_{n,k} = nk - \frac{1}{2}k(k+1)$.
- $\boxed{F(Y) = AY - YY^T AY}$
- $F(Y^*) = 0$ iff the columns of Y^* are eigenvectors of A.

Solving nonlinear equations: Euclidean case

- Goal: find $p^* \in \Omega$ such that $\boxed{F(p^*) = 0 \in \mathbb{R}^n}$, where Ω is open and $F : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ is a C^1 vector field.
- Newton's method: $\boxed{F(p_k) + F'(p_k)(p_{k+1} - p_k) = 0.}$



Outline

1. Abstract differential geometry setting for R-Newton

2. Other explicit examples

Outline

Metric framework

- M is endowed with a Riemannian metric g:

$$\|v\|_p^2 = g(p)(v, v) \text{ for } v \in T_pM.$$

- Riemannian distance $d : M \times M \to [0, +\infty)$:

$$d(p, q) = \inf\{\int_a^b \|\dot{c}(t)\|_{c(t)} dt \mid c : [a, b] \to M,\ c(a) = p,\ c(b) = q\}$$

- Assumption:

(M, d) is a complete metric space.

- Covariant derivative:

$$F'(p)v := \nabla_v F(p) = (\nabla_Y F)(p),\ v \in T_pM,$$

where

- Y is any vector field on M satisfying $v = Y(p)$.
- ∇ is the Riemannian (or Levi-Civita) connection on (M, g).

Metric framework

- M is endowed with a Riemannian metric g:

$$\|v\|_p^2 = g(p)(v,v) \text{ for } v \in T_pM.$$

- Riemannian distance $d : M \times M \to [0, +\infty)$:

$$d(p,q) = \inf\{\int_a^b \|\dot{c}(t)\|_{c(t)} dt \mid c : [a,b] \to M,\ c(a) = p,\ c(b) = q\}$$

- Assumption:

$$(M, d) \text{ is a complete metric space.}$$

- Covariant derivative:

$$F'(p)v := \nabla_v F(p) = (\nabla_Y F)(p),\ v \in T_pM,$$

where
  - Y is any vector field on M satisfying $v = Y(p)$.
  - ∇ is the Riemannian (or Levi-Civita) connection on (M, g).

Metric framework

- M is endowed with a Riemannian metric g:

$$\|v\|_p^2 = g(p)(v, v) \text{ for } v \in T_pM.$$

- Riemannian distance $d : M \times M \to [0, +\infty)$:

$$d(p, q) = \inf\{\int_a^b \|\dot{c}(t)\|_{c(t)} dt \mid c : [a, b] \to M,\ c(a) = p,\ c(b) = q\}$$

- Assumption:

$$(M, d) \text{ is a complete metric space.}$$

- Covariant derivative:

$$F'(p)v := \nabla_v F(p) = (\nabla_Y F)(p),\ v \in T_pM,$$

where

- Y is any vector field on M satisfying $v = Y(p)$.
- ∇ is the Riemannian (or Levi-Civita) connection on (M, g).

Metric framework

- M is endowed with a Riemannian metric g:

$$\|v\|_p^2 = g(p)(v,v) \text{ for } v \in T_pM.$$

- Riemannian distance $d : M \times M \to [0, +\infty)$:

$$d(p,q) = \inf\{\int_a^b \|\dot{c}(t)\|_{c(t)} dt \mid c : [a,b] \to M,\ c(a) = p,\ c(b) = q\}$$

- Assumption:

(M, d) is a complete metric space.

- Covariant derivative:

$$F'(p)v := \nabla_v F(p) = (\nabla_Y F)(p),\ v \in T_pM,$$

where

- Y is any vector field on M satisfying $v = Y(p)$.
- ∇ is the Riemannian (or Levi-Civita) connection on (M, g).

Metric framework

- M is endowed with a Riemannian metric g:

$$\|v\|_p^2 = g(p)(v, v) \text{ for } v \in T_pM.$$

- Riemannian distance $d : M \times M \to [0, +\infty)$:

$$d(p, q) = \inf\{\int_a^b \|\dot{c}(t)\|_{c(t)} dt \mid c : [a, b] \to M,\ c(a) = p,\ c(b) = q\}$$

- Assumption:

$$(M, d) \text{ is a complete metric space.}$$

- Covariant derivative:

$$F'(p)v := \nabla_v F(p) = (\nabla_Y F)(p),\ v \in T_pM,$$

where

- Y is any vector field on M satisfying $v = Y(p)$.
- ∇ is the Riemannian (or Levi-Civita) connection on (M, g).

Exponential Map

- Geodesic: a curve $\gamma : (a, b) \to M$ with $\boxed{\nabla_{\dot\gamma}\dot\gamma = 0.}$
  If γ_i are the coordinates of γ,

$$\frac{d^2\gamma_k}{dt^2} + \sum_{i,j=1}^{n} \Gamma_{ij}^k \frac{d\gamma_i}{dt}\frac{d\gamma_j}{dt} = 0; \qquad k = 1, \dots, n,$$

  where $\Gamma_{i,j}^k$ are the Christoffel symbols.

- Exponential map: $\exp_p : T_pM \to M$ is defined by setting

$$\exp_p[v] = \gamma(1),$$

  where $\gamma : \mathbb{R} \to M$ is the geodesic with $\gamma(0) = p$ and $\dot\gamma(0) = v$.

Exponential Map

- Geodesic: a curve $\gamma : (a, b) \to M$ with $\boxed{\nabla_{\dot\gamma}\dot\gamma = 0.}$
  If γ_i are the coordinates of γ,

$$\frac{d^2\gamma_k}{dt^2} + \sum_{i,j=1}^{n} \Gamma_{ij}^k \frac{d\gamma_i}{dt}\frac{d\gamma_j}{dt} = 0; \qquad k = 1, \ldots, n,$$

  where $\Gamma_{i,j}^k$ are the Christoffel symbols.

- Exponential map: $\exp_p : T_pM \to M$ is defined by setting

$$\exp_p[v] = \gamma(1),$$

  where $\gamma : \mathbb{R} \to M$ is the geodesic with $\gamma(0) = p$ and $\dot\gamma(0) = v$.

Exponential Map

- Geodesic: a curve $\gamma : (a, b) \to M$ with $\boxed{\nabla_{\dot\gamma}\dot\gamma = 0.}$
  If γ_i are the coordinates of γ,

$$\frac{d^2\gamma_k}{dt^2} + \sum_{i,j=1}^{n} \Gamma_{ij}^k \frac{d\gamma_i}{dt}\frac{d\gamma_j}{dt} = 0; \qquad k = 1, \ldots, n,$$

  where $\Gamma_{i,j}^k$ are the Christoffel symbols.

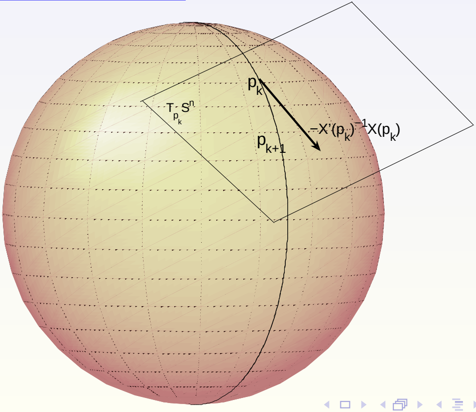
- Exponential map: $\exp_p : T_pM \to M$ is defined by setting

$$\exp_p[v] = \gamma(1),$$

  where $\gamma : \mathbb{R} \to M$ is the geodesic with $\gamma(0) = p$ and $\dot\gamma(0) = v$.

Riemannian Newton's method (Shub '86)

1. Data: Given p_k with $F(p_k) \neq 0$ and $F'(p_k)$ nondegenerate.
2. Newton's correction: Find $v_k \in T_{p_k}M$ s.t. $F'(p_k)v_k = -F(p_k)$.
3. Update: Set $\boxed{p_{k+1} = \exp_{p_k}[v_k]}$.



Rayleigh quotient (continuation)

- Manifold: $M = \mathbb{S}^n = \{x \in \mathbb{R}^{n+1} \mid x^T x = 1\}$ (unit sphere in $\mathbb{R}^{n+1}$).
- Tangent space: $T_x\mathbb{S}^n = \{v \in \mathbb{R}^{n+1} \mid x^T v = 0\}$.
- Metric: $g(v, w) = v^T w$.

Exponential map:

$$\exp_p[v] = p\cos(\|v\|) + \frac{v}{\|v\|}\sin(\|v\|)$$

Vector field: $F(x) = Ax - q(x)x$ with $q(x) = x^T Ax$.

Newton direction at p:

$$v = -p + \frac{1}{p^T w} w$$

where

$$(A - q(p)I)w = p$$

Rayleigh quotient (continuation)

- Manifold: $M = \mathbb{S}^n = \{x \in \mathbb{R}^{n+1} \mid x^T x = 1\}$ (unit sphere in $\mathbb{R}^{n+1}$).
- Tangent space: $T_x\mathbb{S}^n = \{v \in \mathbb{R}^{n+1} \mid x^T v = 0\}$.
- Metric: $g(v, w) = v^T w$.

Exponential map:

$$\boxed{\exp_p[v] = p\cos(\|v\|) + \frac{v}{\|v\|}\sin(\|v\|)}$$

Vector field: $\boxed{F(x) = Ax - q(x)x}$ with $q(x) = x^T Ax$.

Newton direction at p:

$$v = -p + \frac{1}{p^T w} w$$

where

$$(A - q(p)I)w = p$$

Rayleigh quotient (continuation)

- Manifold: $M = \mathbb{S}^n = \{x \in \mathbb{R}^{n+1} \mid x^T x = 1\}$ (unit sphere in $\mathbb{R}^{n+1}$).
- Tangent space: $T_x\mathbb{S}^n = \{v \in \mathbb{R}^{n+1} \mid x^T v = 0\}$.
- Metric: $g(v, w) = v^T w$.

Exponential map:

$$\boxed{\exp_p[v] = p\cos(\|v\|) + \frac{v}{\|v\|}\sin(\|v\|)}$$

Vector field: $\boxed{F(x) = Ax - q(x)x}$ with $q(x) = x^T Ax$.

Newton direction at p:

$$v = -p + \frac{1}{p^T w} w$$

where

$$(A - q(p)I)w = p$$

Rayleigh quotient (continuation)

- Manifold: $M = \mathbb{S}^n = \{x \in \mathbb{R}^{n+1} \mid x^T x = 1\}$ (unit sphere in $\mathbb{R}^{n+1}$).
- Tangent space: $T_x\mathbb{S}^n = \{v \in \mathbb{R}^{n+1} \mid x^T v = 0\}$.
- Metric: $g(v, w) = v^T w$.

Exponential map:

$$\boxed{\exp_p[v] = p\cos(\|v\|) + \frac{v}{\|v\|}\sin(\|v\|)}$$

Vector field: $\boxed{F(x) = Ax - q(x)x}$ with $q(x) = x^T Ax$.

Newton direction at p:

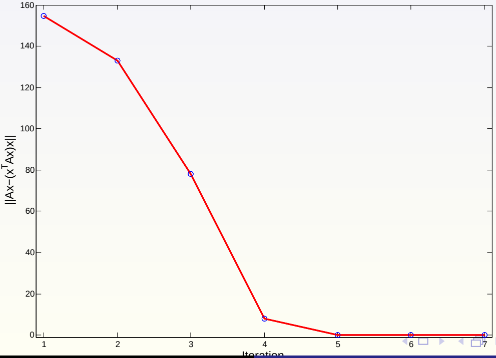
$$v = -p + \frac{1}{p^T w} w$$

where

$$(A - q(p)I)w = p$$

Rayleigh quotient (continuation)

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 3 & 6 & 10 & 15 & 21 \\ 1 & 4 & 10 & 20 & 35 & 56 \\ 1 & 5 & 15 & 35 & 70 & 126 \\ 1 & 6 & 21 & 56 & 126 & 252 \end{pmatrix}$$



Some references

- UDRISTE '94 (VOL. 297, KLUWER): Convexity and optimization on manifolds, including R-Newton.
- SMITH '94 (FIELDS INSTITUTE COMM.): Existence of a basin of attraction for quadratic convergence.
- EDELMAN, ARIAS & SMITH '98 (SIAM J. ON MATRIX ANAL. AND APPL): Matrix orthogonality constrains.
- FERREIRA & SVAITER '02 (J. OF COMPLEXITY): Kantorovich-type proximity test for quadratic convergence under a Lipschitz condition.
- DEDIEU, PRIOURET & MALAJOVICH '03 (IMA J.NUMER. ANAL.): Smale-type proximity test for analytic vector fields.
- A., BOLTE & MUNIER (FOUND. COMP. MATHEMATICS '08): General and unifying proximity test.

Outline

Positivity constraints

- $M = \mathbb{R}^n_{++}$ and $g(p)(u, v) = \sum_{k=1}^n u_k v_k / h_k^2(p_k)$ where $h_i : \mathbb{R}_{++} \to \mathbb{R}_{++}$ differentiable function.
- Isometry: $\left.\gamma = \gamma(t) \text{ geodesic} \right\} \overset{\Longleftrightarrow}{\text{solution}} \int \frac{1}{h_i(\gamma_k)} d\gamma_k = a_k t + b_k$
- For the barrier $\phi(p) = -\sum_{k=1}^n \log(p_k)$, $\nabla^2\phi(p) = \text{diag}(h_1^2(p_1), \ldots, h_n^2(p_n))$ where $h_k(p_k) = p_k$. Thus,

$$\left.\begin{array}{r}\gamma \text{ geodesics} \\ \gamma(0) = p, \dot\gamma(0) = v\end{array}\right\} \Longleftrightarrow \gamma_k(t) = p_k \exp\left(t \frac{v_k}{p_k}\right)$$

Positivity constraints

- $M = \mathbb{R}^n_{++}$ and $g(p)(u,v) = \sum_{k=1}^n u_k v_k / h_k^2(p_k)$ where $h_i : \mathbb{R}_{++} \to \mathbb{R}_{++}$ differentiable function.
- Isometry: $\gamma = \gamma(t)$ geodesic $\}$ $\overset{\text{solution}}{\Longleftrightarrow}$ $\int \frac{1}{h_i(\gamma_k)} d\gamma_k = a_k t + b_k$
- For the barrier $\phi(p) = -\sum_{k=1}^n \log(p_k)$, $\nabla^2\phi(p) = \mathrm{diag}(h_1^2(p_1), \ldots, h_n^2(p_n))$ where $h_k(p_k) = p_k$. Thus,

$$\left.\begin{array}{r}\gamma \text{ geodesics} \\ \gamma(0) = p, \dot{\gamma}(0) = v\end{array}\right\} \Longleftrightarrow \gamma_k(t) = p_k \exp\left(t \frac{v_k}{p_k}\right)$$

Positivity constraints

- $M = \mathbb{R}^n_{++}$ and $g(p)(u,v) = \sum_{k=1}^n u_k v_k / h_k^2(p_k)$ where $h_i : \mathbb{R}_{++} \to \mathbb{R}_{++}$ differentiable function.
- Isometry: $\left.\gamma = \gamma(t) \text{ geodesic} \right\} \overset{\Longleftrightarrow}{\text{solution}} \int \frac{1}{h_i(\gamma_k)} d\gamma_k = a_k t + b_k$
- For the barrier $\phi(p) = -\sum_{k=1}^n \log(p_k)$, $\nabla^2\phi(p) = \text{diag}(h_1^2(p_1), \ldots, h_n^2(p_n))$ where $h_k(p_k) = p_k$. Thus,

$$\left.\begin{array}{r}\gamma \text{ geodesics} \\ \gamma(0) = p, \dot\gamma(0) = v\end{array}\right\} \Longleftrightarrow \gamma_k(t) = p_k \exp\left(t \frac{v_k}{p_k}\right)$$

Positivity constraints

- $M = \mathbb{R}^n_{++}$ and $g(p)(u,v) = \sum_{k=1}^n u_k v_k / h_k^2(p_k)$ where $h_i : \mathbb{R}_{++} \to \mathbb{R}_{++}$ differentiable function.
- Isometry: $\left.\gamma = \gamma(t) \text{ geodesic}\right\} \underset{\text{solution}}{\Longleftrightarrow} \int \frac{1}{h_i(\gamma_k)} d\gamma_k = a_k t + b_k$
- For the barrier $\phi(p) = -\sum_{k=1}^n \log(p_k)$, $\nabla^2\phi(p) = \text{diag}(h_1^2(p_1), \ldots, h_n^2(p_n))$ where $h_k(p_k) = p_k$. Thus,

$$\left.\begin{array}{r} \gamma \text{ geodesics} \\ \gamma(0) = p, \dot\gamma(0) = v \end{array}\right\} \Longleftrightarrow \gamma_k(t) = p_k \exp\left(t \frac{v_k}{p_k}\right)$$

Relative interior of the unitary simplex

- $M = \Delta^{n-1}_{++} = \{p \in \mathbb{R}^n \mid \sum_{i=1}^{n} p_i = 1,\ p_i > 0,\ i = 1, \ldots, n\}$,

  Tangent space: $T_p\Delta^{n-1} = \{v \in \mathbb{R}^n \mid \sum_{i=1}^{n} v_i = 0\}$

  Metric: $g(p)(u,v) = (1 - \frac{1}{n}) \sum_{k=1}^{n} \frac{u_k v_k}{h_k^2(p_k)}$.

- $\gamma = \gamma(t)$ geodesics $\} \overset{\Longleftrightarrow}{\text{solution}} \frac{d}{dt}\left(\frac{1}{h_k(\gamma_k)}\frac{d\gamma_k}{dt} - \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_i(\gamma_i)}\frac{d\gamma_i}{dt}\right) = 0$

  For $h_k(p_k) = p_k$

$$\left.\begin{array}{r}\gamma \text{ geodesics}\\ \gamma(0) = p, \dot\gamma(0) = v\end{array}\right\} \Longleftrightarrow \gamma_k(t) = \frac{p_k \exp(t\frac{v_k}{p_k})}{\sum_{i=1}^{n} p_i \exp(t\frac{v_i}{p_i})}$$

Relative interior of the unitary simplex

- $M = \Delta_{++}^{n-1} = \{p \in \mathbb{R}^n \mid \sum_{i=1}^{n} p_i = 1,\ p_i > 0,\ i = 1, \ldots, n\}$,

  Tangent space: $T_p\Delta^{n-1} = \{v \in \mathbb{R}^n \mid \sum_{i=1}^{n} v_i = 0\}$

  Metric: $g(p)(u, v) = (1 - \frac{1}{n}) \sum_{k=1}^{n} \frac{u_k v_k}{h_k^2(p_k)}$.

- $\gamma = \gamma(t)$ geodesics $\} \overset{\text{solution}}{\Longleftrightarrow} \frac{d}{dt}\left(\frac{1}{h_k(\gamma_k)}\frac{d\gamma_k}{dt} - \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_i(\gamma_i)}\frac{d\gamma_i}{dt}\right) = 0$

  For $h_k(p_k) = p_k$

$$\left.\begin{array}{c}\gamma \text{ geodesics} \\ \gamma(0) = p, \dot\gamma(0) = v\end{array}\right\} \Longleftrightarrow \gamma_k(t) = \frac{p_k \exp(t\frac{v_k}{p_k})}{\sum_{i=1}^{n} p_i \exp(t\frac{v_i}{p_i})}$$

Relative interior of the unitary simplex

- $M = \Delta^{n-1}_{++} = \{p \in \mathbb{R}^n \mid \sum_{i=1}^{n} p_i = 1,\ p_i > 0,\ i = 1, \ldots, n\}$,

  Tangent space: $T_p\Delta^{n-1} = \{v \in \mathbb{R}^n \mid \sum_{i=1}^{n} v_i = 0\}$

  Metric: $g(p)(u, v) = (1 - \frac{1}{n}) \sum_{k=1}^{n} \frac{u_k v_k}{h_k^2(p_k)}$.

- $\gamma = \gamma(t)$ geodesics $\Big\} \overset{\Longleftrightarrow}{\text{solution}} \frac{d}{dt}\left(\frac{1}{h_k(\gamma_k)}\frac{d\gamma_k}{dt} - \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_i(\gamma_i)}\frac{d\gamma_i}{dt}\right) = 0$

For $h_k(p_k) = p_k$

$$\left.\begin{array}{r}\gamma \text{ geodesics} \\ \gamma(0) = p, \dot{\gamma}(0) = v\end{array}\right\} \Longleftrightarrow \gamma_k(t) = \frac{p_k \exp(t\frac{v_k}{p_k})}{\sum_{i=1}^{n} p_i \exp(t\frac{v_i}{p_i})}$$

The Stiefel manifold (continuation)

- $S_{n,k} = \{A \in \mathbb{R}^{n\times k} : \; A^T A = I_k\}$
- Tangent space: $T_A S_{n,k} = \{\Delta \in \mathbb{R}^{n\times k} : \; \Delta^T A + A^T \Delta = 0\}$
- Euclidian Metric: $g(\Delta_1, \Delta_2) = \mathrm{trace}\,(\Delta_1^T \Delta_2)$

The equation that described the geodesics is given by

$$\ddot{Y}(t) + Y(t)(\dot{Y}(t)^T \dot{Y}(t)) = 0,$$

and the corresponding exponential map is:

$$\exp_A[\Delta] = (A \quad \Delta) \exp\left[\begin{pmatrix} A^T\Delta & -\Delta^T\Delta \\ I_p & A^T\Delta \end{pmatrix}\right] I_{2p,p} \exp(-A^T\Delta),$$

where $I_{2p,p} = \begin{pmatrix} I_{2p} \\ \mathbf{0}_p \end{pmatrix}$ and $\exp(B) = I + B + \frac{1}{2}B^2 + \ldots$.

The Stiefel manifold (continuation)

- $S_{n,k} = \{A \in \mathbb{R}^{n\times k} : \; A^T A = I_k\}$
- Tangent space: $T_A S_{n,k} = \{\Delta \in \mathbb{R}^{n\times k} : \; \Delta^T A + A^T \Delta = 0\}$
- Euclidian Metric: $g(\Delta_1, \Delta_2) = \operatorname{trace}(\Delta_1^T \Delta_2)$

The equation that described the geodesics is given by

$$\ddot{Y}(t) + Y(t)(\dot{Y}(t)^T \dot{Y}(t)) = 0,$$

and the corresponding exponential map is:

$$\exp_A[\Delta] = (A \quad \Delta) \exp\left[\begin{pmatrix} A^T\Delta & -\Delta^T\Delta \\ I_p & A^T\Delta \end{pmatrix}\right] I_{2p,p} \exp(-A^T\Delta),$$

where $I_{2p,p} = \begin{pmatrix} I_{2p} \\ \mathbf{0}_p \end{pmatrix}$ and $\exp(B) = I + B + \frac{1}{2}B^2 + \ldots$.

Stiefel manifold (continuation)

- Manifold: $S_{n,k} = \{Y \in \mathbb{R}^{n\times k} : \; Y^T Y = I_k\}$.
- Tangent space: $T_Y S_{n,k} = \{\Delta \in \mathbb{R}^{n\times k} : \; \Delta^T Y + Y^T \Delta = 0\}$.
- Euclidian Metric: $g(\Delta_1, \Delta_2) = \mathrm{trace}\,(\Delta_1^T \Delta_2)$.

Exponential map:

$$\exp_Y[\Delta] = (Y \quad \Delta) \exp\left[\begin{pmatrix} Y^T\Delta & -\Delta^T\Delta \\ I_p & Y^T\Delta \end{pmatrix}\right] I_{2p,p} \exp(-Y^T\Delta)$$

Vector field: $V(Y) = BY - Y(Y^T BY)$

Newton direction Δ at Y must satisfy:
$B\Delta - \Delta Y^T BY - \frac{1}{2} Y(Y^T B\Delta + \Delta^T BY + \Delta^T YY^T BY + Y^T BYY^T \Delta) = -V(Y)$

Stiefel manifold (continuation)

- Manifold: $S_{n,k} = \{Y \in \mathbb{R}^{n\times k} : \; Y^T Y = I_k\}$.
- Tangent space: $T_Y S_{n,k} = \{\Delta \in \mathbb{R}^{n\times k} : \; \Delta^T Y + Y^T \Delta = 0\}$.
- Euclidian Metric: $g(\Delta_1, \Delta_2) = \mathrm{trace}\,(\Delta_1^T \Delta_2)$.

Exponential map:

$$\boxed{\exp_Y[\Delta] = (Y \quad \Delta) \exp\left[\begin{pmatrix} Y^T\Delta & -\Delta^T\Delta \\ I_p & Y^T\Delta \end{pmatrix}\right] I_{2p,p} \exp(-Y^T\Delta)}$$

Vector field: $\boxed{V(Y) = BY - Y(Y^T BY)}$

Newton direction Δ at Y must satisfy:

$$B\Delta - \Delta Y^T BY - \tfrac{1}{2} Y(Y^T B\Delta + \Delta^T BY + \Delta^T YY^T BY + Y^T BYY^T\Delta) = -V(Y)$$

Stiefel manifold (continuation)

- Manifold: $S_{n,k} = \{Y \in \mathbb{R}^{n\times k} : \ Y^T Y = I_k\}$.
- Tangent space: $T_Y S_{n,k} = \{\Delta \in \mathbb{R}^{n\times k} : \ \Delta^T Y + Y^T \Delta = 0\}$.
- Euclidian Metric: $g(\Delta_1, \Delta_2) = \operatorname{trace}(\Delta_1^T \Delta_2)$.

Exponential map:

$$\boxed{\exp_Y[\Delta] = (Y \quad \Delta)\exp\left[\begin{pmatrix} Y^T\Delta & -\Delta^T\Delta \\ I_p & Y^T\Delta \end{pmatrix}\right] I_{2p,p}\exp(-Y^T\Delta)}$$

Vector field: $\boxed{V(Y) = BY - Y(Y^T BY)}$

Newton direction Δ at Y must satisfy:

$B\Delta - \Delta Y^T BY - \frac{1}{2}Y(Y^T B\Delta + \Delta^T BY + \Delta^T YY^T BY + Y^T BYY^T\Delta) = -V(Y)$

Stiefel manifold (continuation)

- Manifold: $S_{n,k} = \{Y \in \mathbb{R}^{n\times k} : \; Y^T Y = I_k\}$.
- Tangent space: $T_Y S_{n,k} = \{\Delta \in \mathbb{R}^{n\times k} : \; \Delta^T Y + Y^T \Delta = 0\}$.
- Euclidian Metric: $g(\Delta_1, \Delta_2) = \mathrm{trace}\,(\Delta_1^T \Delta_2)$.

Exponential map:

$$\exp_Y[\Delta] = (Y \quad \Delta) \exp\left[\begin{pmatrix} Y^T\Delta & -\Delta^T\Delta \\ I_p & Y^T\Delta \end{pmatrix}\right] I_{2p,p} \exp(-Y^T\Delta)$$

Vector field: $V(Y) = BY - Y(Y^T BY)$

Newton direction Δ at Y must satisfy:

$$B\Delta - \Delta Y^T BY - \frac{1}{2} Y(Y^T B\Delta + \Delta^T BY + \Delta^T YY^T BY + Y^T BYY^T \Delta) = -V(Y)$$

Stiefel Manifold (continuation)

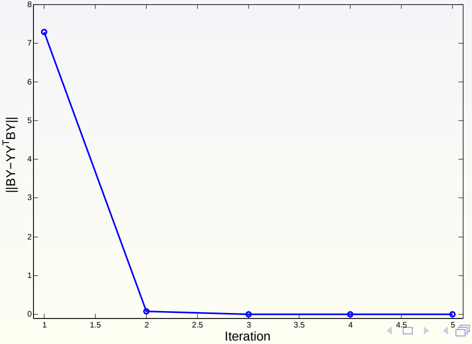
$$B = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 6 & 10 & 15 \\ 1 & 4 & 10 & 20 & 35 \\ 1 & 5 & 15 & 35 & 70 \end{pmatrix}, \quad Y_0 = \begin{pmatrix} 0.1947 & -0.6155 & -0.7448 \\ -0.4268 & 0.5660 & -0.4579 \\ 0.7236 & 0.1412 & 0.2122 \\ -0.5050 & -0.5012 & 0.3901 \\ 0.0355 & 0.1723 & -0.1958 \end{pmatrix}$$

and $S_{5,3}$.

Stiefel Manifold (continuation)

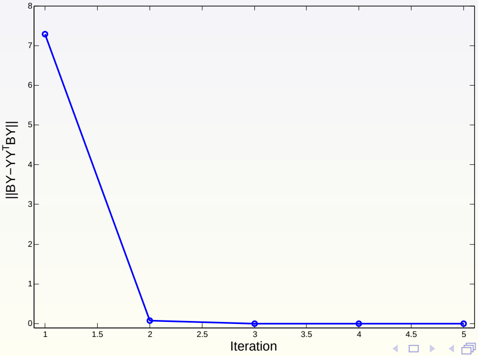
$$B = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 6 & 10 & 15 \\ 1 & 4 & 10 & 20 & 35 \\ 1 & 5 & 15 & 35 & 70 \end{pmatrix}, \quad Y_0 = \begin{pmatrix} 0.1947 & -0.6155 & -0.7448 \\ -0.4268 & 0.5660 & -0.4579 \\ 0.7236 & 0.1412 & 0.2122 \\ -0.5050 & -0.5012 & 0.3901 \\ 0.0355 & 0.1723 & -0.1958 \end{pmatrix}$$

and $S_{5,3}$.



Stiefel Manifold (continuation)

$$B = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 6 & 10 & 15 \\ 1 & 4 & 10 & 20 & 35 \\ 1 & 5 & 15 & 35 & 70 \end{pmatrix}, \quad Y^* = \begin{pmatrix} 0.1795 & -0.5882 & -0.7500 \\ -0.4682 & 0.6017 & -0.4265 \\ 0.7062 & 0.1923 & 0.2209 \\ -0.4847 & -0.4777 & 0.4030 \\ 0.1222 & 0.1635 & -0.2108 \end{pmatrix}$$



Cone of Positive Semidefinite Matrices

- Manifold: $M = \mathcal{S}^n_{++}$ cone of symmetric positive definite matrices.
- Tangent space: $T_P\mathcal{S}^n_{++} \simeq \mathcal{S}^n$ ($\mathcal{S}^n$ space symmetric matrices).
- Hessian Metric: $g(\Delta_1, \Delta_2) = \mathrm{trace}(\nabla^2\varphi(P)\Delta_1\Delta_2)$ with φ barrier on $\mathcal{S}^n_{++}$.
  Taking $\varphi(P) = -\log(\det(P))$: $g(\Delta_1, \Delta_2) = \mathrm{trace}\left(P^{-1}\Delta_1 P^{-1}\Delta_2\right)$
- Goal: minimize $f : \mathcal{S}^n \to \mathbb{R}$ over M

Exponential map:

$$\exp_P[\Delta] = P^{1/2}\exp(P^{-1/2}\Delta P^{-1/2})P^{1/2}$$

Vector field: $V(P) = P\nabla f(P)P$

Newton direction Δ at P must satisfy:

$$V'(P)\Delta = \nabla^2 f(P)\Delta + \tfrac{1}{2}(P^{-1}\Delta\nabla f(P) + \nabla f(P)\Delta P^{-1}) = -P\nabla f(P)P$$

Cone of Positive Semidefinite Matrices

- Manifold: $M = \mathcal{S}^n_{++}$ cone of symmetric positive definite matrices.
- Tangent space: $T_P\mathcal{S}^n_{++} \simeq \mathcal{S}^n$ ($\mathcal{S}^n$ space symmetric matrices).
- Hessian Metric: $g(\Delta_1, \Delta_2) = \mathrm{trace}(\nabla^2\varphi(P)\Delta_1\Delta_2)$ with φ barrier on $\mathcal{S}^n_{++}$.
  Taking $\varphi(P) = -\log(\det(P))$: $g(\Delta_1, \Delta_2) = \mathrm{trace}\left(P^{-1}\Delta_1 P^{-1}\Delta_2\right)$
- Goal: minimize $f : \mathcal{S}^n \to \mathbb{R}$ over M

Exponential map:

$$\boxed{\exp_P[\Delta] = P^{1/2}\exp(P^{-1/2}\Delta P^{-1/2})P^{1/2}}$$

Vector field: $\boxed{V(P) = P\nabla f(P)P}$

Newton direction Δ at P must satisfy:

$$V'(P)\Delta = \nabla^2 f(P)\Delta + \tfrac{1}{2}(P^{-1}\Delta\nabla f(P) + \nabla f(P)\Delta P^{-1}) = -P\nabla f(P)P$$

Cone of Positive Semidefinite Matrices

- Manifold: $M = \mathcal{S}^n_{++}$ cone of symmetric positive definite matrices.
- Tangent space: $T_P\mathcal{S}^n_{++} \simeq \mathcal{S}^n$ ($\mathcal{S}^n$ space symmetric matrices).
- Hessian Metric: $g(\Delta_1, \Delta_2) = \mathrm{trace}(\nabla^2\varphi(P)\Delta_1\Delta_2)$ with φ barrier on $\mathcal{S}^n_{++}$.
  Taking $\varphi(P) = -\log(\det(P))$: $g(\Delta_1, \Delta_2) = \mathrm{trace}\left(P^{-1}\Delta_1 P^{-1}\Delta_2\right)$
- Goal: minimize $f : \mathcal{S}^n \to \mathbb{R}$ over M

Exponential map:

$$\boxed{\exp_P[\Delta] = P^{1/2}\exp(P^{-1/2}\Delta P^{-1/2})P^{1/2}}$$

Vector field: $\boxed{V(P) = P\nabla f(P)P}$

Newton direction Δ at P must satisfy:

$$V'(P)\Delta = \nabla^2 f(P)\Delta + \tfrac{1}{2}(P^{-1}\Delta\nabla f(P) + \nabla f(P)\Delta P^{-1}) = -P\nabla f(P)P$$

Cone of Positive Semidefinite Matrices

- Manifold: $M = \mathcal{S}^n_{++}$ cone of symmetric positive definite matrices.
- Tangent space: $T_P\mathcal{S}^n_{++} \simeq \mathcal{S}^n$ ($\mathcal{S}^n$ space symmetric matrices).
- Hessian Metric: $g(\Delta_1, \Delta_2) = \mathrm{trace}(\nabla^2\varphi(P)\Delta_1\Delta_2)$ with φ barrier on $\mathcal{S}^n_{++}$.
  Taking $\varphi(P) = -\log(\det(P))$: $g(\Delta_1, \Delta_2) = \mathrm{trace}\,(P^{-1}\Delta_1 P^{-1}\Delta_2)$
- Goal: minimize $f : \mathcal{S}^n \to \mathbb{R}$ over M

Exponential map:

$$\boxed{\exp_P[\Delta] = P^{1/2}\exp(P^{-1/2}\Delta P^{-1/2})P^{1/2}}$$

Vector field: $\boxed{V(P) = P\nabla f(P)P}$

Newton direction Δ at P must satisfy:

$$V'(P)\Delta = \nabla^2 f(P)\Delta + \tfrac{1}{2}(P^{-1}\Delta\nabla f(P) + \nabla f(P)\Delta P^{-1}) = -P\nabla f(P)P$$

Unitary Generalized Simplex on $\mathcal{S}^n_{++}$

- Manifold: $M = \{P \in \mathcal{S}^n : \ \mathrm{Trace}(P) = 1, \ P \in \mathcal{S}^n_{++}\}$.
- Tangent space: $T_P M = \{\Delta \in \mathcal{S}^n : \ \mathrm{Trace}(\Delta) = \sum_{i=1}^n \lambda_i(\Delta) = 0\}$.
- $\varphi(P) = -\log(\det(P))$.

Exponential map:

$$\exp_P[\Delta] = \frac{1}{\mathrm{Trace}(P \exp(P^{-1/2}\Delta P^{-1/2}))} P^{1/2} \exp(P^{-1/2}\Delta P^{-1/2}) P^{1/2}$$

Unitary Generalized Simplex on $\mathcal{S}^n_{++}$

- Manifold: $M = \{P \in \mathcal{S}^n : \text{Trace}(P) = 1, \; P \in \mathcal{S}^n_{++}\}$.
- Tangent space: $T_P M = \{\Delta \in \mathcal{S}^n : \text{Trace}(\Delta) = \sum_{i=1}^n \lambda_i(\Delta) = 0\}$.
- $\varphi(P) = -\log(\det(P))$.

Exponential map:

$$\exp_P[\Delta] = \frac{1}{\text{Trace}(P \exp(P^{-1/2} \Delta P^{-1/2}))} P^{1/2} \exp(P^{-1/2} \Delta P^{-1/2}) P^{1/2}$$

Open problems

- Drop out completeness ?
- Manifolds with boundary.
- Globalization:

$$p_{k+1} = \exp_{p_k}[t_k v_k]$$

for some scalar parameter $t_k > 0$.