Introduction to Differential and Riemannian Geometry

François Lauze

1Department of Computer Science
University of Copenhagen

Ven Summer School On Manifold Learning in Image and Signal Analysis
August 19th, 2009
Outline

1 Motivation
   - Non Linearity
   - Statistics on Non Linear Data

2 Recalls
   - Geometry
   - Topology
   - Calculus on $\mathbb{R}^n$

3 Differentiable Manifolds
   - Definitions
   - Building Manifolds
   - Tangent Space

4 Riemannian Manifolds
   - Metric
   - Gradient Field
   - Length of curves
   - Geodesics
   - Covariant derivatives
Outline

1 Motivation
   - Non Linearity
     - Statistics on Non Linear Data

2 Recalls
   - Geometry
   - Topology
   - Calculus on $\mathbb{R}^n$

3 Differentiable Manifolds
   - Definitions
   - Building Manifolds
   - Tangent Space

4 Riemannian Manifolds
   - Metric
   - Gradient Field
   - Length of curves
   - Geodesics
   - Covariant derivatives
Many interesting/common objects behave non linearly.

Vector lines in $\mathbb{R}^2$

The projective line as a circle.

Right triangles with fixed hypotenuse

Right rectangles as half-circle (without endpoints).
Many interesting /common objects behave non linearly.

Vector lines in $\mathbb{R}^2$

The projective line as a circle.

Right triangles with fixed hypotenuse

Right rectangles as half-circle (without endpoints).
Non Linear data

Many interesting/common objects behave non linearly.

Vector lines in $\mathbb{R}^2$

The projective line as a circle.

Right triangles with fixed hypotenuse

Right rectangles as half-circle (without endpoints).
Outline

1 Motivation
   - Non Linearity
   - Statistics on Non Linear Data

2 Recalls
   - Geometry
   - Topology
   - Calculus on $\mathbb{R}^n$

3 Differentiable Manifolds
   - Definitions
   - Building Manifolds
   - Tangent Space

4 Riemannian Manifolds
   - Metric
   - Gradient Field
   - Length of curves
   - Geodesics
   - Covariant derivatives
Averaging

How to average points on a circle?

Hm... The linear way does not work!

A better way: use the distance on the circle!
Averaging

- How to average points on a circle?
- Hm... The linear way does not work!
- A better way: use the distance on the circle!
How to average points on a circle?
Hm... The linear way does not work!
A better way: use the distance on the circle!
Outline

1 Motivation
   - Non Linearity
   - Statistics on Non Linear Data

2 Recalls
   - Geometry
   - Topology
   - Calculus on $\mathbb{R}^n$

3 Differentiable Manifolds
   - Definitions
   - Building Manifolds
   - Tangent Space

4 Riemannian Manifolds
   - Metric
   - Gradient Field
   - Length of curves
   - Geodesics
   - Covariant derivatives
Inner Products

- **Inner Product**: product $\langle x, y \rangle$ for column vectors $x$ and $y$ in $\mathbb{R}^n$
  - linear in $x$ and $y$,
  - symmetric: $\langle x, y \rangle = \langle y, x \rangle$
  - positive and definite: $\langle x, x \rangle \geq 0$ with equality if $x = 0$.
- Simplest example: usual dot-product $x = (x_1, \ldots, x_n)^t$, $y = (y_1, \ldots, y_n)^t$,
  $$x \cdot y = \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x^t Id y$$
  $Id$ is the $n$-order identity matrix.
- Every inner product is of the form $x^t A y$, $A$ symmetric positive definite.
  Alternate notation: $\langle x, y \rangle_A$.
  Without subscript $\langle -, - \rangle$ will denote standard Euclidean dot-product (i.e. $A = Id$).
Inner Products

- **Inner Product**: product $\langle x, y \rangle$ for column vectors $x$ and $y$ in $\mathbb{R}^n$
  - linear in $x$ and $y$,
  - symmetric: $\langle x, y \rangle = \langle y, x \rangle$
  - positive and definite: $\langle x, x \rangle \geq 0$ with equality if $x = 0$.
- Simplest example: usual dot-product $x = (x_1, \ldots, x_n)^t$, $y = (y_1, \ldots, y_n)^t$,

$$x \cdot y = \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x^t I_d y$$

$I_d$ is the $n$-order identity matrix.
- Every inner product is of the form $x^t A y$, $A$ symmetric positive definite.
  Alternate notation: $\langle x, y \rangle_A$.
  Without subscript $\langle -, - \rangle$ will denote standard Euclidean dot-product (i.e. $A = I_d$).
Inner Products

- **Inner Product**: product $\langle x, y \rangle$ for column vectors $x$ and $y$ in $\mathbb{R}^n$
  - linear in $x$ and $y$,
  - symmetric: $\langle x, y \rangle = \langle y, x \rangle$
  - positive and definite: $\langle x, x \rangle \geq 0$ with equality if $x = 0$.
- Simplest example: usual dot-product $x = (x_1, \ldots, x_n)^t$, $y = (y_1, \ldots, y_n)^t$,

\[
x \cdot y = \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x^t I_d y
\]

$I_d$ is the $n$-order identity matrix.
- Every inner product is of the form $x^t A y$, $A$ symmetric positive definite. Alternate notation: $\langle x, y \rangle_A$.
  Without subscript $\langle -, - \rangle$ will denote standard Euclidean dot-product (i.e. $A = I_d$).
Inner Products

- **Inner Product**: product $\langle x, y \rangle$ for column vectors $x$ and $y$ in $\mathbb{R}^n$
  - linear in $x$ and $y$,
  - symmetric: $\langle x, y \rangle = \langle y, x \rangle$
  - positive and definite: $\langle x, x \rangle \geq 0$ with equality if $x = 0$.

- Simplest example: usual dot-product $x = (x_1, \ldots, x_n)^t$, $y = (y_1, \ldots, y_n)^t$,
  $$x \cdot y = \langle x, y \rangle = \sum_{i=1}^n x_i y_i = x^t I_d y$$

$I_d$ is the $n$-order identity matrix.

- Every inner product is of the form $x^t A y$, $A$ symmetric positive definite.
- Alternate notation: $\langle x, y \rangle_A$.
- Without subscript $\langle -, - \rangle$ will denote standard Euclidean dot-product (i.e. $A = I_d$).
**Inner Products**

- **Inner Product**: product \( \langle x, y \rangle \) for column vectors \( x \) and \( y \) in \( \mathbb{R}^n \)
  - linear in \( x \) and \( y \),
  - symmetric: \( \langle x, y \rangle = \langle y, x \rangle \)
  - positive and definite: \( \langle x, x \rangle \geq 0 \) with equality if \( x = 0 \).

- Simplest example: usual dot-product \( x = (x_1, \ldots, x_n)^t, y = (y_1, \ldots, y_n)^t \),

\[
\begin{align*}
x \cdot y &= \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x^t ld y
\end{align*}
\]

\( ld \) is the \( n \)-order identity matrix.

- Every inner product is of the form \( x^t A y \), \( A \) symmetric positive definite.
  Alternate notation: \( \langle x, y \rangle_A \).
  Without subscript \( \langle -, - \rangle \) will denote standard Euclidean dot-product (i.e. \( A = ld \)).
Inner Products

- **Inner Product**: product $\langle x, y \rangle$ for column vectors $x$ and $y$ in $\mathbb{R}^n$
  - linear in $x$ and $y$,
  - symmetric: $\langle x, y \rangle = \langle y, x \rangle$
  - positive and definite: $\langle x, x \rangle \geq 0$ with equality if $x = 0$.

- Simplest example: usual dot-product $x = (x_1, \ldots, x_n)^t$, $y = (y_1, \ldots, y_n)^t$,

$$x \cdot y = \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x^t Id y$$

$Id$ is the $n$-order identity matrix.

- Every inner product is of the form $x^t A y$, $A$ symmetric positive definite.
  Alternate notation: $\langle x, y \rangle_A$.
  Without subscript $\langle -, - \rangle$ will denote standard Euclidean dot-product (i.e. $A = Id$).
Inner Products

- **Inner Product**: product $\langle x, y \rangle$ for column vectors $x$ and $y$ in $\mathbb{R}^n$
  - linear in $x$ and $y$,
  - symmetric: $\langle x, y \rangle = \langle y, x \rangle$
  - positive and definite: $\langle x, x \rangle \geq 0$ with equality if $x = 0$.
- Simplest example: usual dot-product $x = (x_1, \ldots, x_n)^t$, $y = (y_1, \ldots, y_n)^t$,

$$x \cdot y = \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x^t \text{Id} y$$

$\text{Id}$ is the $n$-order identity matrix.

- Every inner product is of the form $x^t A y$, $A$ symmetric positive definite.
  Alternate notation: $\langle x, y \rangle_A$.
  Without subscript $\langle -, - \rangle$ will denote standard Euclidean dot-product (i.e. $A = \text{Id}$).
Inner Products

- **Inner Product**: product $\langle x, y \rangle$ for column vectors $x$ and $y$ in $\mathbb{R}^n$
  - linear in $x$ and $y$,
  - symmetric: $\langle x, y \rangle = \langle y, x \rangle$
  - positive and definite: $\langle x, x \rangle \geq 0$ with equality if $x = 0$.
- Simplest example: usual dot-product $x = (x_1, \ldots, x_n)^t$, $y = (y_1, \ldots, y_n)^t$,
  \[
x \cdot y = \langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x^t ld y
  \]
  $ld$ is the $n$-order identity matrix.
- Every inner product is of the form $x^t Ay$, $A$ symmetric positive definite.
  Alternate notation: $\langle x, y \rangle_A$.
  Without subscript $\langle -, - \rangle$ will denote standard Euclidean dot-product (i.e. $A = ld$).
Orthogonality – Norm – Distance

- A-orthogonality: \( x \perp_A y \Leftrightarrow \langle x, y \rangle_A = 0. \)
- A-norm of \( x \): \( \| x \|_A = \sqrt{\langle x, x \rangle_A} \).
- A-distance on \( \mathbb{R}^n \): \( d_A(x, y) = \| x - y \|_A \).

Standard orthogonal transform on \( \mathbb{R}^n \): \( n \times n \) matrix \( R \) satisfying \( R^t R = Id \). They form the orthogonal group \( O(n) \). Matrices \( R \) with \( \det = 1 \) form the special orthogonal group \( SO(n) \).

- for general inner product \( \langle -, - \rangle_A \): \( R \) is A-orthogonal if \( R^t A R = A \).
Orthogonality – Norm – Distance

- **A-orthogonality**: $\mathbf{x} \perp_A \mathbf{y} \iff \langle \mathbf{x}, \mathbf{y} \rangle_A = 0$.

- **A-norm of $\mathbf{x}$**: $\|\mathbf{x}\|_A = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_A}$.

- **A-distance on $\mathbb{R}^n$**: $d_A(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_A$.

- Standard orthogonal transform on $\mathbb{R}^n$: $n \times n$ matrix $R$ satisfying $R^t R = \text{Id}$. They form the orthogonal group $O(n)$. Matrices $R$ with $\det = 1$ form the special orthogonal group $SO(n)$.

- for general inner product $\langle - , - \rangle_A$: $R$ is $A$-orthogonal if $R^t A R = A$. 
Orthogonality – Norm – Distance

- **A-orthogonality:** \( \mathbf{x} \perp_{A} \mathbf{y} \iff \langle \mathbf{x}, \mathbf{y} \rangle_{A} = 0 \).
- **A-norm of \( \mathbf{x} \):** \( \| \mathbf{x} \|_{A} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_{A}} \).
- **A-distance on \( \mathbb{R}^{n} \):** \( d_{A}(\mathbf{x}, \mathbf{y}) = \| \mathbf{x} - \mathbf{y} \|_{A} \).

Standard orthogonal transform on \( \mathbb{R}^{n} \): \( n \times n \) matrix \( R \) satisfying \( R^{t}R = \text{Id} \). They form the orthogonal group \( O(n) \). Matrices \( R \) with \( \det = 1 \) form the special orthogonal group \( SO(n) \).

- for general inner product \( \langle - , - \rangle_{A} \): \( R \) is A-orthogonal if \( R^{t}AR = A \).
Orthogonality – Norm – Distance

- **A-orthogonality**: $\mathbf{x} \perp_A \mathbf{y} \Leftrightarrow \langle \mathbf{x}, \mathbf{y} \rangle_A = 0$.
- **A-norm of $\mathbf{x}$**: $\| \mathbf{x} \|_A = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_A}$.
- **A-distance on $\mathbb{R}^n$**: $d_A(\mathbf{x}, \mathbf{y}) = \| \mathbf{x} - \mathbf{y} \|_A$.

- Standard orthogonal transform on $\mathbb{R}^n$: $n \times n$ matrix $R$ satisfying $R^t R = \text{Id}$. They form the orthogonal group $O(n)$. Matrices $R$ with $\det = 1$ form the special orthogonal group $SO(n)$.
- for general inner product $\langle -, - \rangle_A$: $R$ is $A$-orthogonal if $R^t A R = A$. 
Orthogonality – Norm – Distance

- $A$-orthogonality: $\mathbf{x} \perp_A \mathbf{y} \iff \langle \mathbf{x}, \mathbf{y} \rangle_A = 0$.
- $A$-norm of $\mathbf{x}$: $\| \mathbf{x} \|_A = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_A}$.
- $A$-distance on $\mathbb{R}^n$: $d_A(\mathbf{x}, \mathbf{y}) = \| \mathbf{x} - \mathbf{y} \|_A$.

- Standard orthogonal transform on $\mathbb{R}^n$: $n \times n$ matrix $R$ satisfying $R^t R = \text{Id}$. They form the orthogonal group $O(n)$. Matrices $R$ with $\det = 1$ form the special orthogonal group $SO(n)$.

- For general inner product $\langle -, - \rangle_A$: $R$ is $A$-orthogonal if $R^t A R = A$. 
Orthogonality – Norm – Distance

- **A-orthogonality**: $\mathbf{x} \perp_A \mathbf{y} \iff \langle \mathbf{x}, \mathbf{y} \rangle_A = 0$.
- **A-norm of $\mathbf{x}$**: $\| \mathbf{x} \|_A = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle_A}$.
- **A-distance on $\mathbb{R}^n$**: $d_A(\mathbf{x}, \mathbf{y}) = \| \mathbf{x} - \mathbf{y} \|_A$.

- Standard orthogonal transform on $\mathbb{R}^n$: $n \times n$ matrix $R$ satisfying $R^t R = I_d$.
  They form the **orthogonal group** $O(n)$. Matrices $R$ with $\det = 1$ form the **special orthogonal group** $SO(n)$.
- for general inner product $\langle -, - \rangle_A$: $R$ is $A$-orthogonal if $R^t A R = A$. 

François Lauze (University of Copenhagen)
Duality

- Linear form $h : \mathbb{R}^n \to \mathbb{R}$: $h(x) = \sum_{i=1}^{n} h_i x_i$.
- Given an inner product $\langle - , - \rangle_A$ on $\mathbb{R}^n$, $h$ represented by a unique vector $h_A$ s.t
  \[ h(x) = \langle h_A, x \rangle_A \]
- $h_A$ is the dual of $h$ (w.r.t $\langle - , - \rangle_A$).
- For standard dot product:
  \[ h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \]
- For general inner product $\langle - , - \rangle_A$
  \[ h_A = A^{-1} h. \]
Duality

- Linear form $h : \mathbb{R}^n \to \mathbb{R}$: $h(x) = \sum_{i=1}^{n} h_i x_i$.

- Given an inner product $\langle -, - \rangle_A$ on $\mathbb{R}^n$, $h$ represented by a unique vector $h_A$ s.t

  $$h(x) = \langle h_A, x \rangle_A$$

  $h_A$ is the dual of $h$ (w.r.t $\langle -, - \rangle_A$).

- For standard dot product:

  $$h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$$

- For general inner product $\langle -, - \rangle_A$

  $$h_A = A^{-1} h.$$
**Duality**

- Linear form $h : \mathbb{R}^n \rightarrow \mathbb{R}$: $h(x) = \sum_{i=1}^{n} h_i x_i$.
- Given an inner product $\langle -, - \rangle_A$ on $\mathbb{R}^n$, $h$ represented by a unique vector $h_A$ s.t
  $$h(x) = \langle h_A, x \rangle_A$$

  $h_A$ is the **dual** of $h$ (w.r.t $\langle -, - \rangle_A$).
- for standard dot product:
  $$h = \begin{pmatrix}
  h_1 \\
  \vdots \\
  h_n
  \end{pmatrix}$$

- for general inner product $\langle -, - \rangle_A$
  $$h_A = A^{-1}h.$$
Duality

- Linear form \( h : \mathbb{R}^n \rightarrow \mathbb{R} : h(x) = \sum_{i=1}^{n} h_i x_i. \)
- Given an inner product \( \langle -, - \rangle_A \) on \( \mathbb{R}^n \), \( h \) represented by a unique vector \( h_A \) s.t
  \[
  h(x) = \langle h_A, x \rangle_A
  \]
- \( h_A \) is the dual of \( h \) (w.r.t \( \langle -, - \rangle_A \)).
- for standard dot product:
  \[
  h = \begin{pmatrix}
  h_1 \\
  \vdots \\
  h_n
  \end{pmatrix}
  \]
- for general inner product \( \langle -, - \rangle_A \)
  \[
  h_A = A^{-1} h.
  \]
Duality

- Linear form $h : \mathbb{R}^n \to \mathbb{R}$: $h(x) = \sum_{i=1}^{n} h_i x_i$.
- Given an inner product $\langle -, - \rangle_A$ on $\mathbb{R}^n$, $h$ represented by a unique vector $h_A$ s.t
  $$h(x) = \langle h_A, x \rangle_A$$
- $h_A$ is the dual of $h$ (w.r.t $\langle -, - \rangle_A$).
- For standard dot product:
  $$h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$$
- For general inner product $\langle -, - \rangle_A$
  $$h_A = A^{-1} h.$$
<table>
<thead>
<tr>
<th>1</th>
<th>Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Non Linearity</td>
</tr>
<tr>
<td>1.2</td>
<td>Statistics on Non Linear Data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2</th>
<th>Recalls</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Geometry</td>
</tr>
<tr>
<td>2.2</td>
<td>Topology</td>
</tr>
<tr>
<td>2.3</td>
<td>Calculus on $\mathbb{R}^n$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3</th>
<th>Differentiable Manifolds</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Definitions</td>
</tr>
<tr>
<td>3.2</td>
<td>Building Manifolds</td>
</tr>
<tr>
<td>3.3</td>
<td>Tangent Space</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4</th>
<th>Riemannian Manifolds</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Metric</td>
</tr>
<tr>
<td>4.2</td>
<td>Gradient Field</td>
</tr>
<tr>
<td>4.3</td>
<td>Length of curves</td>
</tr>
<tr>
<td>4.4</td>
<td>Geodesics</td>
</tr>
<tr>
<td>4.5</td>
<td>Covariant derivatives</td>
</tr>
</tbody>
</table>
Open sets, Continuity...

- Topology on $\mathbb{R}^n$. An open set of $\mathbb{R}^n$ is a union (not necessarily finite) of open balls. $\mathbb{R}^N$ and the empty set $\emptyset$ are open.

- A map $f : X \rightarrow Y$ between topological spaces is continuous if

$$V \subset Y \text{ open } \Rightarrow f^{-1}(V) \subset X \text{ open}.$$ 

- A map $h : X \rightarrow Y$ between topological spaces is a homeomorphism if it is continuous, one-to-one and $h^{-1}$ is continuous.
Open sets, Continuity...

- **Topology on** \( \mathbb{R}^n \). A open set of \( \mathbb{R}^n \) is a union (not necessarily finite) of open balls. \( \mathbb{R}^N \) and the empty set \( \emptyset \) are open.

- A map \( f : X \rightarrow Y \) between topological spaces is **continuous** if

  \[ V \subset Y \text{ open} \Rightarrow f^{-1}(V) \subset X \text{ open}. \]

- A map \( h : X \rightarrow Y \) between topological spaces is a **homeomorphism** if it is continuous, one-to-one and \( h^{-1} \) is continuous.
Topology on $\mathbb{R}^n$. A open set of $\mathbb{R}^n$ is a union (not necessarily finite) of open balls. $\mathbb{R}^N$ and the empty set $\emptyset$ are open.

A map $f : X \rightarrow Y$ between topological spaces is continuous if

$$V \subset Y \text{ open} \implies f^{-1}(V) \subset X \text{ open}.$$ 

A map $h : X \rightarrow Y$ between topological spaces is a homeomorphism if it is continuous, one-to-one and $h^{-1}$ is continuous.
Open sets, Continuity...

- Topology on $\mathbb{R}^n$. A open set of $\mathbb{R}^n$ is a union (not necessarily finite) of open balls. $\mathbb{R}^N$ and the empty set $\emptyset$ are open.

- A map $f : X \rightarrow Y$ between topological spaces is continuous if

  $$V \subset Y \text{ open} \Rightarrow f^{-1}(V) \subset X \text{ open}.$$  

- A map $h : X \rightarrow Y$ between topological spaces is a homeomorphism if it is continuous, one-to-one and $h^{-1}$ is continuous.
Outline

1 Motivation
   - Non Linearity
   - Statistics on Non Linear Data

2 Recalls
   - Geometry
   - Topology
   - Calculus on $\mathbb{R}^n$

3 Differentiable Manifolds
   - Definitions
   - Building Manifolds
   - Tangent Space

4 Riemannian Manifolds
   - Metric
   - Gradient Field
   - Length of curves
   - Geodesics
   - Covariant derivatives
Differentiable and smooth functions

- $f : U \text{ open} \subset \mathbb{R}^n \to \mathbb{R}^q$ continuous: write
  $$ (y_1, \ldots, y_1) = f(x_1, \ldots, x_n) $$

- $f$ is of class $C^r$ if $f$ has continuous partial derivatives
  $$ \frac{\partial^{r_1 + \cdots + r_n} y_k}{\partial x_1^{r_1} \cdots \partial x_n^{r_n}} $$
  $$ k = 1 \ldots q, r_1 + \ldots r_n \leq r. $$

- When $r = \infty$, I say that $f$ is smooth. This is the main situation of interest.
Differentiable and smooth functions

- \( f : U \text{ open} \subset \mathbb{R}^n \rightarrow \mathbb{R}^q \) continuous: write
  \[
  (y_1, \ldots, y_1) = f(x_1, \ldots, x_n)
  \]

- \( f \) is of class \( C^r \) if \( f \) has continuous partial derivatives
  \[
  \frac{\partial^{r_1+\ldots+r_n} y_k}{\partial x_1^{r_1} \ldots \partial x_n^{r_n}}
  \]

  \( k = 1 \ldots q, \ r_1 + \ldots r_n \leq r \).

- When \( r = \infty \), I say that \( f \) is smooth. This is the main situation of interest.
Differentiable and smooth functions

- \( f : U \text{ open } \subset \mathbb{R}^n \rightarrow \mathbb{R}^q \) continuous: write
  \[ (y_1, \ldots, y_1) = f(x_1, \ldots, x_n) \]

- \( f \) is of class \( C^r \) if \( f \) has continuous partial derivatives
  \[ \frac{\partial^{r_1 + \cdots + r_n} y_k}{\partial x_1^{r_1} \cdots \partial x_n^{r_n}} \]

- \( k = 1 \ldots q, r_1 + \cdots r_n \leq r \).

- When \( r = \infty \), I say that \( f \) is smooth. This is the main situation of interest.
Differentiable and smooth functions

- \( f : U \text{ open} \subset \mathbb{R}^n \to \mathbb{R}^q \) continuous: write
  \[
  (y_1, \ldots, y_q) = f(x_1, \ldots, x_n)
  \]

- \( f \) is of class \( C^r \) if \( f \) has continuous partial derivatives
  \[
  \frac{\partial^{r_1 + \cdots + r_n} y_k}{\partial x_1^{r_1} \cdots \partial x_n^{r_n}}
  \]
  \( k = 1 \ldots q, \ r_1 + \cdots r_n \leq r \).

- When \( r = \infty \), I say that \( f \) is smooth. This is the main situation of interest.
Differential, Jacobian Matrix

- **Differential of $f$ in $x$:** unique linear map (if exists) $d_x f : \mathbb{R}^n \rightarrow \mathbb{R}^q$ s.t.

  \[ f(x + h) = f(x) + d_x f(h) + o(h). \]

- **Jacobian matrix of $f$:** matrix $q \times n$ of partial derivatives of $f$:

  \[
  J_x f = \begin{pmatrix}
  \frac{\partial y_1}{\partial x_1}(x) & \cdots & \frac{\partial y_1}{\partial x_q}(x) \\
  \vdots & \ddots & \vdots \\
  \frac{\partial y_q}{\partial x_1}(x) & \cdots & \frac{\partial y_q}{\partial x_n}(x)
  \end{pmatrix}
  \]

  - if $n \geq q$ and rank($J_x f$) = $q$, $f$ is a submersion at $x$.
  - if $n \leq q$ and rank($J_x f$) = $n$, $f$ is an immersion at $x$. 
Differential, Jacobian Matrix

- **Differential of $f$ in $x$:** unique linear map (if exists) $d_x f : \mathbb{R}^n \rightarrow \mathbb{R}^q$ s.t.
  $$f(x + h) = f(x) + d_x f(h) + o(h).$$

- **Jacobian matrix of $f$:** matrix $q \times n$ of partial derivatives of $f$:
  $$J_x f = \begin{pmatrix}
  \frac{\partial y_1}{\partial x_1}(x) & \cdots & \frac{\partial y_1}{\partial x_n}(x) \\
  \vdots & \ddots & \vdots \\
  \frac{\partial y_q}{\partial x_1}(x) & \cdots & \frac{\partial y_q}{\partial x_n}(x)
  \end{pmatrix}
  $$

  - if $n \geq q$ and $\text{rank}(J_x f) = q$, $f$ is a submersion at $x$.
  - if $n \leq q$ and $\text{rank}(J_x f) = n$, $f$ is an immersion at $x$. 
Differential, Jacobian Matrix

- Differential of \( f \) in \( x \): unique linear map (if exists) \( d_x f : \mathbb{R}^n \rightarrow \mathbb{R}^q \) s.t.

\[
f(x + h) = f(x) + d_x f(h) + o(h).
\]

- Jacobian matrix of \( f \): matrix \( q \times n \) of partial derivatives of \( f \):

\[
J_x f = \begin{pmatrix}
\frac{\partial y_1}{\partial x_1}(x) & \cdots & \frac{\partial y_1}{\partial x_n}(x) \\
\vdots & \ddots & \vdots \\
\frac{\partial y_q}{\partial x_1}(x) & \cdots & \frac{\partial y_q}{\partial x_n}(x)
\end{pmatrix}
\]

- if \( n \geq q \) and \( \text{rank}(J_x f) = q \), \( f \) is a submersion at \( x \).
- if \( n \leq q \) and \( \text{rank}(J_x f) = n \), \( f \) is an immersion at \( x \).
Differential, Jacobian Matrix

- **Differential of $f$ in $x$:** unique linear map (if exists) $d_x f : \mathbb{R}^n \rightarrow \mathbb{R}^q$ s.t.
  \[
  f(x + h) = f(x) + d_x f(h) + o(h).
  \]

- **Jacobian matrix of $f$:** matrix $q \times n$ of partial derivatives of $f$:
  \[
  J_x f = 
  \begin{pmatrix}
  \frac{\partial y_1}{\partial x_1}(x) & \cdots & \frac{\partial y_1}{\partial x_n}(x) \\
  \vdots & \ddots & \vdots \\
  \frac{\partial y_q}{\partial x_1}(x) & \cdots & \frac{\partial y_q}{\partial x_n}(x)
  \end{pmatrix}
  \]

- If $n \geq q$ and $\text{rank}(J_x f) = q$, $f$ is a submersion at $x$.
- If $n \leq q$ and $\text{rank}(J_x f) = n$, $f$ is an immersion at $x$. 
Differential, Jacobian Matrix

- **Differential of** $f$ **in** $x$: unique linear map (if exists) $d_x f : \mathbb{R}^n \to \mathbb{R}^q$ s.t.
  \[
  f(x + h) = f(x) + d_x f(h) + o(h).
  \]

- **Jacobian matrix of** $f$: matrix $q \times n$ of partial derivatives of $f$:
  \[
  J_x f = \begin{pmatrix}
  \frac{\partial y_1}{\partial x_1}(x) & \cdots & \frac{\partial y_1}{\partial x_n}(x) \\
  \vdots & \ddots & \vdots \\
  \frac{\partial y_q}{\partial x_1}(x) & \cdots & \frac{\partial y_q}{\partial x_n}(x)
  \end{pmatrix}
  \]

  - if $n \geq q$ and rank($J_x f$) = $q$, $f$ is a submersion at $x$.
  - if $n \leq q$ and rank($J_x f$) = $n$, $f$ is an immersion at $x$. 
Diffeomorphism

- When \( n = q \): if \( f \) is 1-1 \( C^r \) and its inverse is also \( C^r \), \( f \) is a \( C^r \)-diffeomorphism. A smooth diffeomorphism is simply referred to as a diffeomorphism.

- If \( f \) is a diffeomorphism, \( \det(J_x f) \neq 0 \). Conversely, if \( \det(J_x f) \neq 0 \), by the Inverse Function Theorem, \( f \) is a local diffeomorphism in a neighborhood of \( x \).

- \( f \) may be a local diffeomorphism everywhere but fail to be a global diffeomorphism. Example:

  \[
  f : \mathbb{R}^2 \setminus 0 \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto (e^x \cos(y), e^x \sin(y)).
  \]

- If \( f \) is 1-1 and a local diffeomorphism everywhere, it is a global diffeomorphism.
Diffeomorphism

- when $n = q$: if $f$ is 1-1 $C^r$ and its inverse is also $C^r$, $f$ is a $C^r$-diffeomorphism. A smooth diffeomorphism is simply referred to as a diffeomorphism.

- If $f$ is a diffeomorphism, $\det(J_x f) \neq 0$. Conversely, if $\det(J_x f) \neq 0$, by the Inverse Function Theorem, $f$ is a local diffeomorphism in a neighborhood of $x$.

- $f$ may be a local diffeomorphism everywhere but fail to be a global diffeomorphism. Example:

  $$f : \mathbb{R}^2 \setminus 0 \to \mathbb{R}^2, \quad (x, y) \to (e^x \cos(y), e^x \sin(y))$$

- if $f$ is 1-1 and a local diffeomorphism everywhere, it is a global diffeomorphism.
Diffeomorphism

- when \( n = q \): if \( f \) is 1-1 \( C^r \) and its inverse is also \( C^r \), \( f \) is a \( C^r \)-diffeomorphism. A smooth diffeomorphism is simply referred to as a diffeomorphism.

- If \( f \) is a diffeomorphism, \( \det(J_x f) \neq 0 \). Conversely, if \( \det(J_x f) \neq 0 \), by the Inverse Function Theorem, \( f \) is a local diffeomorphism in a neighborhood of \( x \).

- \( f \) may be a local diffeomorphism everywhere but fail to be a global diffeomorphism. Example:

\[
f : \mathbb{R}^2 \setminus 0 \rightarrow \mathbb{R}^2, \quad (x, y) \rightarrow (e^x \cos(y), e^x \sin(y)).
\]

- if \( f \) is 1-1 and a local diffeomorphism everywhere, it is a global diffeomorphism.
Diffeomorphism

- when \( n = q \): if \( f \) is 1-1 \( C^r \) and its inverse is also \( C^r \), \( f \) is a \( C^r \)-diffeomorphism. A smooth diffeomorphism is simply referred to as a diffeomorphism.
- If \( f \) is a diffeomorphism, \( \det(J_x f) \neq 0 \). Conversely, if \( \det(J_x f) \neq 0 \), by the Inverse Function Theorem, \( f \) is a local diffeomorphism in a neighborhood of \( x \).
- \( f \) may be a local diffeomorphism everywhere but fail to be a global diffeomorphism. Example:

\[
f : \mathbb{R}^2 \setminus 0 \rightarrow \mathbb{R}^2, \quad (x, y) \rightarrow (e^x \cos(y), e^x \sin(y)).
\]

- if \( f \) is 1-1 and a local diffeomorphism everywhere, it is a global diffeomorphism.
Diffeomorphism

- when \( n = q \): if \( f \) is 1-1 \( C^r \) and its inverse is also \( C^r \), \( f \) is a \( C^r \)-diffeomorphism. A smooth diffeomorphism is simply referred to as a diffeomorphism.

- If \( f \) is a diffeomorphism, \( \det(J_x f) \neq 0 \). Conversely, if \( \det(J_x f) \neq 0 \), by the Inverse Function Theorem, \( f \) is a local diffeomorphism in a neighborhood of \( x \).

- \( f \) may be a local diffeomorphism everywhere but fail to be a global diffeomorphism. Example:

  \[
  f : \mathbb{R}^2 \setminus 0 \rightarrow \mathbb{R}^2, \quad (x, y) \rightarrow (e^x \cos(y), e^x \sin(y)).
  \]

- if \( f \) is 1-1 and a local diffeomorphism everywhere, it is a global diffeomorphism.
Gradient of a function

- \( f : U \subset \mathbb{R}^n \to \mathbb{R} \), \( d_x f \) its differential at \( x \in U \).
- \( d_x f \) is represented by a unique vector, the gradient of \( f \) for the standard inner product:

\[
d_x f(h) = \nabla f_x \cdot h
\]

- If one changes the inner product, the gradient changes too, but not the differential.
- The gradient indicates the direction of largest change (by Cauchy-Schwarz).
Gradient of a function

- \( f : U \subset \mathbb{R}^n \to \mathbb{R} \), \( d_x f \) it differential at \( x \in U \).
- \( d_x f \) is represented by a unique vector, the gradient of \( f \) for the standard inner product:
  \[
  d_x f(h) = \nabla f_x \cdot h
  \]
- If one changes the inner product, the gradient changes too, but not the differential.
- The gradient indicates the direction of largest change (by Cauchy-Schwarz).
Gradient of a function

- $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $d_x f$ it differential at $x \in U$.
- $d_x f$ is represented by a unique vector, the gradient of $f$ for the standard inner product:
  \[ d_x f(h) = \nabla f_x \cdot h \]

- If one changes the inner product, the gradient changes too, but not the differential.
- The gradient indicates the direction of largest change (by Cauchy-Schwarz).
Gradient of a function

\[ f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}, \quad d_x f \text{ its differential at } x \in U. \]

\[ d_x f \text{ is represented by a unique vector, the gradient of } f \text{ for the standard inner product:} \]

\[ d_x f(h) = \nabla f_x \cdot h \]

- If one changes the inner product, the gradient changes too, but not the differential.
- The gradient indicates the direction of largest change (by Cauchy-Schwarz).
Gradient of a function

- \( f : U \subset \mathbb{R}^n \rightarrow \mathbb{R} \), \( d_x f \) its differential at \( x \in U \).
- \( d_x f \) is represented by a unique vector, the gradient of \( f \) for the standard inner product:
  \[
d_x f(h) = \nabla f_x \cdot h
  \]

- If one changes the inner product, the gradient changes too, but not the differential.
- The gradient indicates the direction of largest change (by Cauchy-Schwarz).
Outline

1 Motivation
   - Non Linearity
   - Statistics on Non Linear Data

2 Recalls
   - Geometry
   - Topology
   - Calculus on $\mathbb{R}^n$

3 Differentiable Manifolds
   - Definitions
   - Building Manifolds
   - Tangent Space

4 Riemannian Manifolds
   - Metric
   - Gradient Field
   - Length of curves
   - Geodesics
   - Covariant derivatives
Differentiable Manifold

Definitions

Differentiable Manifold $M$ of dim $n$:
- smoothly glued open pieces of Euclidean space $\mathbb{R}^n$ via $M = \bigcup_i V_i$, homeomorphisms $\varphi_i : V_i \to W_i \subset \mathbb{R}^n$,
- $\varphi_i(P) = (x_1(P), \ldots, x_n(P))$

Charts or local coordinates

- Smoothness in gluing: the changes of coordinates
  $\varphi_j \circ \varphi_i^{-1} : \varphi_i(V_i \cap V_j) \to \varphi_j(V_i \cap V_j)$
  are smooth.
- Set
  $\varphi_j(P) = (y^1(P), \ldots, y^n(P))$
  then
  $\varphi_j \circ \varphi_i^{-1}(x_1, \ldots, x_n) = (y_1, \ldots, y_n)$
  and the $n \times n$ Jacobian matrices
  $\left( \frac{\partial y^k}{\partial x^h} \right)_{k,h}$
  are invertible.
- Maps $\varphi_i^{-1} : W_i \to V_i$ are local parametrization of $M$. 
**Differentiable Manifold**

A **differentiable manifold** $M$ of dim $n$ is a smooth set $M = \bigcup_i V_i$, homeomorphisms $\varphi_i : V_i \to W_i \subset \mathbb{R}^n$, such that for any $P \in V_i \cap V_j$, the changes of coordinates $\varphi_j \circ \varphi_i^{-1} : \varphi_i(V_i \cap V_j) \to \varphi_j(V_i \cap V_j)$ are smooth.

- Charts or local coordinates $\varphi_i(P) = (x_1(P), \ldots, x_n(P))$
- Maps $\varphi_i^{-1} : W_i \to V_i$ are local parametrization of $M$. 

Smoothness in gluing: the changes of coordinates $\varphi_j \circ \varphi_i^{-1} : \varphi_i(V_i \cap V_j) \to \varphi_j(V_i \cap V_j)$ are smooth.

- Set $\varphi_j(P) = (y^1(P), \ldots, y^n(P))$, then $\varphi_j \circ \varphi_i^{-1}(x_1, \ldots, x_n) = (y_1, \ldots, y_n)$ and the $n \times n$ Jacobian matrices $\left( \frac{\partial y^k}{\partial x^h} \right)_{k,h}$ are invertible.
Differentiable Manifold

- **Definitions**

**Differentiable Manifold** $M$ of dim $n$:

- smoothly glued open pieces of Euclidean space $\mathbb{R}^n$ via $M = \bigcup_i V_i$, homeomorphisms $\varphi_i : V_i \to W_i \subset \mathbb{R}^n$,

- $\varphi_i(P) = (x_1(P), \ldots, x_n(P))$

Charts or local coordinates

- Smoothness in gluing: the changes of coordinates

$$\varphi_j \circ \varphi_i^{-1} : \varphi_i(V_i \cap V_j) \to \varphi_j(V_i \cap V_j)$$

are smooth.

- Set

$$\varphi_j(P) = (y^1(P), \ldots, y^n(P)),$$

then

$$\varphi_j \circ \varphi_i^{-1}(x_1, \ldots, x_n) = (y_1, \ldots, y_n)$$

and the $n \times n$ Jacobian matrices

$$\left( \frac{\partial y^k}{\partial x^h} \right)_{k,h}$$

are invertible.

- Maps $\varphi_i^{-1} : W_i \to V_i$ are local parametrization of $M$. 

François Lauze (University of Copenhagen)

Differential Geometry
Differentiable Manifold

- **Differential manifold** $M$ of dim $n$:
  - smoothly glued open pieces of Euclidean space $\mathbb{R}^n$ via $M = \bigcup_i V_i$, homeomorphisms $\varphi_i : V_i \rightarrow W_i \subset \mathbb{R}^n$,
  - $\varphi_i(P) = (x_1(P), \ldots, x_n(P))$

Charts or local coordinates

- Smoothness in gluing: the changes of coordinates
  $$\varphi_j \circ \varphi_i^{-1} : \varphi_i(V_i \cap V_j) \rightarrow \varphi_j(V_i \cap V_j)$$

are smooth.

- Set
  $$\varphi_j(P) = (y^1(P), \ldots, y^n(P)),$$

then
  $$\varphi_j \circ \varphi_i^{-1}(x_1, \ldots, x_n) = (y_1, \ldots, y_n)$$

and the $n \times n$ Jacobian matrices $(\frac{\partial y^k}{\partial x^h})_{k,h}$ are invertible.

- Maps $\varphi_i^{-1} : W_i \rightarrow V_i$ are local parametrization of $M$. 
Differentiable Manifold

- **Differential manifold** $M$ of dim $n$:
  - smoothly glued open pieces of Euclidean space $\mathbb{R}^n$ via $M = \bigcup_i V_i$, homeomorphisms $\varphi_i : V_i \to W_i \subset \mathbb{R}^n$,
  - $\varphi_i(P) = (x_1(P), \ldots, x_n(P))$
- Charts or local coordinates
- **Smoothness in gluing**: the changes of coordinates
  $$\varphi_j \circ \varphi_i^{-1} : \varphi_i(V_i \cap V_j) \to \varphi_j(V_i \cap V_j)$$
  are smooth.
- Set
  $$\varphi_j(P) = (y^1(P), \ldots, y^n(P)),$$
  then
  $$\varphi_j \circ \varphi_i^{-1}(x_1, \ldots, x_n) = (y_1, \ldots, y_n)$$
  and the $n \times n$ Jacobian matrices
  $$\left( \frac{\partial y^k}{\partial x^h} \right)_{k,h}$$
  are invertible.
- Maps $\varphi_i^{-1} : W_i \to V_i$ are local parametrization of $M$. 

François Lauze (University of Copenhagen)
Differentiable maps

- $f : M \rightarrow N$ is **differentiable** if its expression in any coordinates for $M$ and $N$ is.
- $\varphi$ local coordinates at $P \in M$, $\psi$ local coordinates at $f(P) \in N$

\[ \varphi^{-1} \circ f \circ \psi \text{ differentiable}. \]
Differentiable maps

- \( f : M \rightarrow N \) is **differentiable** if its expression in any coordinates for \( M \) and \( N \) is.
- \( \varphi \) local coordinates at \( P \in M \), \( \psi \) local coordinates at \( f(P) \in N \)

\[
\varphi^{-1} \circ f \circ \psi \text{ differentiable.}
\]
Differentiable maps

- $f : M \to N$ is **differentiable** if its expression in any coordinates for $M$ and $N$ is.
- $\varphi$ local coordinates at $P \in M$, $\psi$ local coordinates at $f(P) \in N$

$$\varphi^{-1} \circ f \circ \psi$$

differentiable.
Differentiable maps

- $f : M \rightarrow N$ is **differentiable** if its expression in any coordinates for $M$ and $N$ is.
- $\varphi$ local coordinates at $P \in M$, $\psi$ local coordinates at $f(P) \in N$

\[ \varphi^{-1} \circ f \circ \psi \text{ differentiable.} \]
First Examples

- The Euclidean space $\mathbb{R}^n$ is a manifold: take $\varphi = \text{Id}$ as global coordinate system!
- The sphere $S^2 = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}$

For instance the projection from North Pole, given, for a point $P = (x, y, z) \neq N$ of the sphere, by

$$\varphi_N(P) = \left( \frac{x}{1 - z}, \frac{y}{1 - z} \right)$$

is a (large) local coordinate system (around the south pole).
First Examples

- The Euclidean space $\mathbb{R}^n$ is a manifold: take $\varphi = \text{Id}$ as global coordinate system!
- The sphere $S^2 = \{(x, y, z), x^2 + y^2 + z^2 = 1\}$

For instance the projection from North Pole, given, for a point $P = (x, y, z) \neq N$ of the sphere, by

$$\varphi_N(P) = \left( \frac{x}{1 - z}, \frac{y}{1 - z} \right)$$

is a (large) local coordinate system (around the south pole).
First Examples

- The Euclidean space \( \mathbb{R}^n \) is a manifold: take \( \varphi = Id \) as global coordinate system!
- The sphere \( S^2 = \{ (x, y, z), x^2 + y^2 + z^2 = 1 \} \)

For instance the projection from North Pole, given, for a point \( P = (x, y, z) \neq N \) of the sphere, by

\[
\varphi_N(P) = \left( \frac{x}{1 - z}, \frac{y}{1 - z} \right)
\]

is a (large) local coordinate system (around the south pole).
First Examples

- The Euclidean space $\mathbb{R}^n$ is a manifold: take $\varphi = \text{Id}$ as global coordinate system!
- The sphere $S^2 = \{(x, y, z), x^2 + y^2 + z^2 = 1\}$

For instance the projection from North Pole, given, for a point $P = (x, y, z) \neq N$ of the sphere, by

$$\varphi_N(P) = \left(\frac{x}{1-z}, \frac{y}{1-z}\right)$$

is a (large) local coordinate system (around the south pole).
Examples

The Moebius strip

\[ u \in [0, 2\pi], \ v \in \left[ \frac{1}{2}, \frac{1}{2} \right] \]

\[
\begin{pmatrix}
\cos(u) \left( 1 + \frac{1}{2} \nu \cos\left(\frac{u}{2}\right) \right) \\
\sin(u) \left( 1 + \frac{1}{2} \nu \cos\left(\frac{u}{2}\right) \right) \\
\frac{1}{2} \nu \sin(u)
\end{pmatrix}
\]

The 2D-torus

\[ (u, v) \in [0, 2\pi]^2, \ R \gg r > 0 \]

\[
\begin{pmatrix}
\cos(u) \left( R + r \cos(v) \right) \\
\sin(u) \left( R + r \cos(v) \right) \\
r \sin(v)
\end{pmatrix}
\]
Examples

The Moebius strip

\[ u \in [0, 2\pi], \ v \in \left[ \frac{1}{2}, \frac{1}{2} \right] \]

\[
\begin{pmatrix}
\cos(u) \left( 1 + \frac{1}{2} v \cos\left( \frac{u}{2} \right) \right) \\
\sin(u) \left( 1 + \frac{1}{2} v \cos\left( \frac{u}{2} \right) \right) \\
\frac{1}{2} v \sin\left( \frac{u}{2} \right)
\end{pmatrix}
\]

The 2D-torus

\[(u, v) \in [0, 2\pi]^2, \ R \gg r > 0\]

\[
\begin{pmatrix}
\cos(u) \left( R + r \cos(v) \right) \\
\sin(u) \left( R + r \cos(v) \right) \\
r \sin(v)
\end{pmatrix}
\]
Examples

The Moebius strip

\[ u \in [0, 2\pi], \quad v \in \left[\frac{1}{2}, \frac{1}{2}\right] \]

\[
\begin{pmatrix}
\cos(u) \left(1 + \frac{1}{2} v \cos\left(\frac{u}{2}\right)\right) \\
\sin(u) \left(1 + \frac{1}{2} v \cos\left(\frac{u}{2}\right)\right) \\
\frac{1}{2} v \sin\left(\frac{u}{2}\right)
\end{pmatrix}
\]

The 2D-torus

\[(u, v) \in [0, 2\pi]^2, \quad R \gg r > 0\]

\[
\begin{pmatrix}
\cos(u) (R + r \cos(v)) \\
\sin(u) (R + r \cos(v)) \\
r \sin(v)
\end{pmatrix}
\]
Outline

1 Motivation
   • Non Linearity
   • Statistics on Non Linear Data

2 Recalls
   • Geometry
   • Topology
   • Calculus on $\mathbb{R}^n$

3 Differentiable Manifolds
   • Definitions
   • Building Manifolds
   • Tangent Space

4 Riemannian Manifolds
   • Metric
   • Gradient Field
   • Length of curves
   • Geodesics
   • Covariant derivatives
Submanifolds of $\mathbb{R}^n$

- Take $f : U \in \mathbb{R}^n \rightarrow \mathbb{R}^q$, $q \leq n$ smooth.
- Set $M = f^{-1}(0)$.
- If for all $x \in M$, $f$ is a submersion at $x$, $M$ is a manifold of dimension $n - q$.
- Example:
  \[
  f(x_1, \ldots, x_n) = 1 - \sum_{i=1}^{n} x_i^2 :
  \]
  $f^{-1}(0)$ is the $(n - 1)$-dimensional unit sphere $S^{n-1}$.
- Many common examples of manifolds in practice are of that type.
Submanifolds of $\mathbb{R}^n$

- Take $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^q$, $q \leq n$ smooth.
- Set $M = f^{-1}(0)$.
- If for all $x \in M$, $f$ is a submersion at $x$, $M$ is a manifold of dimension $n - q$.
- Example:
  $$f(x_1, \ldots, x_n) = 1 - \sum_{i=1}^{n} x_i^2 :$$
- $f^{-1}(0)$ is the $(n - 1)$-dimensional unit sphere $S^{n-1}$.
- Many common examples of manifolds in practice are of that type.
Submanifolds of $\mathbb{R}^n$

- Take $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^q$, $q \leq n$ smooth.
- Set $M = f^{-1}(0)$.
- If for all $x \in M$, $f$ is a submersion at $x$, $M$ is a manifold of dimension $n - q$.
- Example:

  $$f(x_1, \ldots, x_n) = 1 - \sum_{i=1}^{n} x_i^2 :$$

  $f^{-1}(0)$ is the $(n - 1)$-dimensional unit sphere $S^{n-1}$.
- Many common examples of manifolds in practice are of that type.
Submanifolds of $\mathbb{R}^n$

- Take $f : U \subset \mathbb{R}^n \to \mathbb{R}^q$, $q \leq n$ smooth.
- Set $M = f^{-1}(0)$.
- If for all $x \in M$, $f$ is a submersion at $x$, $M$ is a manifold of dimension $n - q$.
- Example:
  \[
  f(x_1, \ldots, x_n) = 1 - \sum_{i=1}^{n} x_i^2 : \\
  f^{-1}(0) \text{ is the (}n - 1\text{)-dimensional unit sphere } S^{n-1}.
  \]
- Many common examples of manifolds in practice are of that type.
Submanifolds of $\mathbb{R}^n$

- Take $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^q$, $q \leq n$ smooth.
- Set $M = f^{-1}(0)$.
- If for all $x \in M$, $f$ is a submersion at $x$, $M$ is a manifold of dimension $n - q$.

Example:

$$f(x_1, \ldots, x_n) = 1 - \sum_{i=1}^{n} x_i^2 :$$

$f^{-1}(0)$ is the $(n - 1)$-dimensional unit sphere $S^{n-1}$.

- Many common examples of manifolds in practice are of that type.
Submanifolds of $\mathbb{R}^n$

- Take $f : U \subset \mathbb{R}^n \to \mathbb{R}^q$, $q \leq n$ smooth.
- Set $M = f^{-1}(0)$.
- If for all $x \in M$, $f$ is a submersion at $x$, $M$ is a manifold of dimension $n - q$.
- Example:
  \[ f(x_1, \ldots, x_n) = 1 - \sum_{i=1}^{n} x_i^2 : \]
  
  $f^{-1}(0)$ is the $(n - 1)$-dimensional unit sphere $S^{n-1}$.
- Many common examples of manifolds in practice are of that type.
Product Manifolds

- $M$ and $N$ manifolds, so is $M \times N$.
- Just consider the products of charts of $M$ and $N$!
- Example: $M = S^1$, $N = \mathbb{R}$: cylinder.
- Example: $M = N = S^1$: the torus!
Product Manifolds

- \( M \) and \( N \) manifolds, so is \( M \times N \).
- Just consider the products of charts of \( M \) and \( N \)!
- Example: \( M = S^1, N = \mathbb{R} \): cylinder.
- Example: \( M = N = S^1 \): the torus!
Product Manifolds

- $M$ and $N$ manifolds, so is $M \times N$.
- Just consider the products of charts of $M$ and $N$!
  - Example: $M = \mathbb{S}^1$, $N = \mathbb{R}$: cylinder.
  - Example: $M = N = \mathbb{S}^1$: the torus!
Product Manifolds

- $M$ and $N$ manifolds, so is $M \times N$.
- Just consider the products of charts of $M$ and $N$!
- Example: $M = S^1$, $N = \mathbb{R}$: cylinder.
- Example: $M = N = S^1$: the torus!
Product Manifolds

- $M$ and $N$ manifolds, so is $M \times N$.
- Just consider the products of charts of $M$ and $N$!
- Example: $M = S^1$, $N = \mathbb{R}$: cylinder.
- Example: $M = N = S^1$: the torus!
Outline

1 Motivation
   - Non Linearity
   - Statistics on Non Linear Data

2 Recalls
   - Geometry
   - Topology
   - Calculus on $\mathbb{R}^n$

3 Differentiable Manifolds
   - Definitions
   - Building Manifolds
     - Tangent Space

4 Riemannian Manifolds
   - Metric
   - Gradient Field
   - Length of curves
   - Geodesics
   - Covariant derivatives
Informally: a tangent vector at $P \in M$: draw a curve $c : (-\epsilon, \epsilon) \rightarrow M$, $c(0) = P$, then $\dot{c}(0)$ is a tangent vector.
Informally: a tangent vector at $P \in M$: draw a curve $c : (-\epsilon, \epsilon) \rightarrow M$, $c(0) = P$, then $\dot{c}(0)$ is a tangent vector.
Informally: a tangent vector at $P \in M$: draw a curve $c : (-\epsilon, \epsilon) \to M$, $c(0) = P$, then $\dot{c}(0)$ is a tangent vector.
A bit more formally

- \( c : c : (-\epsilon, \epsilon) \to M, c(0) = P \). In chart \( \varphi \), the map \( t \mapsto \varphi \circ c(t) \) is a curve in Euclidean space, and so is \( t \mapsto \psi \circ c(t) \).
- Set \( v = \frac{d}{dt} (\varphi \circ c)|_0 \), \( w = \frac{d}{dt} (\psi \circ c)|_0 \) then

\[
w = J_0 (\varphi^{-1} \circ \psi) \; v.
\]

- Use this relation to identify vectors in different coordinate systems!
A bit more formally

\[ c : c : (−\epsilon, \epsilon) \rightarrow M, c(0) = P. \] In chart \( \varphi \), the map \( t \mapsto \varphi \circ c(t) \) is a curve in Euclidean space, and so is \( t \mapsto \psi \circ c(t) \).

- set \( v = \frac{d}{dt} (\varphi \circ c)|_0 \), \( w = \frac{d}{dt} (\psi \circ c)|_0 \) then

\[ w = J_0 (\varphi^{-1} \circ \psi) v. \]

- Use this relation to identify vectors in different coordinate systems!
A bit more formally

\[ c : c : (−\epsilon, \epsilon) \rightarrow M, \ c(0) = P. \text{ In chart } \varphi, \text{ the map } t \mapsto \varphi \circ c(t) \text{ is a curve in Euclidean space, and so is } t \mapsto \psi \circ c(t). \]

\[ \text{set } v = \frac{d}{dt} (\varphi \circ c)|_0, \ w = \frac{d}{dt} (\psi \circ c)|_0 \text{ then} \]

\[ w = J_0 (\varphi^{-1} \circ \psi) \ v. \]

Use this relation to identify vectors in different coordinate systems!
A bit more formally

- \( c : c : (-\epsilon, \epsilon) \to M, \, c(0) = P \). In chart \( \varphi \), the map \( t \mapsto \varphi \circ c(t) \) is a curve in Euclidean space, and so is \( t \mapsto \psi \circ c(t) \).
- Set \( v = \frac{d}{dt} (\varphi \circ c)|_0 \), \( w = \frac{d}{dt} (\psi \circ c)|_0 \) then
  \[
  w = J_0 (\varphi^{-1} \circ \psi) \, v.
  \]
- Use this relation to identify vectors in different coordinate systems!
The set of tangent vectors to the n-dimensional manifold \( M \) at point \( P \) is the tangent space of \( M \) at \( P \) denoted \( T_\mathcal{P}M \).

It is a vector space of dimension \( n \): let \( \theta \) a local parametrization of \( M \), \( \theta(x_1, \ldots, x_n) \in M \) with \( \theta(0) = P \). Define curves

\[
x_i : t \mapsto \theta(0, \ldots, 0, t, 0, 0)
\]

They go through \( P \) when \( t = 0 \) and follow the axes. Their derivative at 0 are denoted \( \partial_{x_i} \). They form a basis of \( T_\mathcal{P}M \).
The set of tangent vectors to the n-dimensional manifold $M$ at point $P$ is the **tangent space of $M$ at $P$** denoted $T_PM$.

It is a vector space of dimension $n$: let $\theta$ a local parametrization of $M$, $\theta(x_1, \ldots, x_n) \in M$ with $\theta(0) = P$. Define curves $x_i : t \mapsto \theta(0, \ldots, 0, t, 0, 0)$.

They go through $P$ when $t = 0$ and follow the axes. Their derivative at 0 are denoted $\partial_{x_i}$. They form a basis of $T_PM$. 
Tangent space

- The set of tangent vectors to the n-dimensional manifold \( M \) at point \( P \) is the tangent space of \( M \) at \( P \) denoted \( T_P M \).

- It is a vector space of dimension \( n \): let \( \theta \) a local parametrization of \( M \), \( \theta(x_1, \ldots, x_n) \in M \) with \( \theta(0) = P \). Define curves
  \[ x_i : t \mapsto \theta(0, \ldots, 0, t, 0, 0) \]

  They go through \( P \) when \( t = 0 \) and follow the axes. Their derivative at 0 are denoted \( \partial x_i \). They form a basis of \( T_P M \).
Tangent space

- The set of tangent vectors to the n-dimensional manifold $M$ at point $P$ is the **tangent space of $M$ at $P$** denoted $T_PM$.
- It is a vector space of dimension $n$: let $\theta$ a local parametrization of $M$, $\theta(x_1, \ldots, x_n) \in M$ with $\theta(0) = P$. Define curves
  \[ x_i : t \mapsto \theta(0, \ldots, 0, t, 0, 0) \]
  They go through $P$ when $t = 0$ and follow the axes. Their derivative at 0 are denoted $\partial_{x_i}$. They form a basis of $T_PM$. 

Tangent space

- The set of tangent vectors to the n-dimensional manifold $M$ at point $P$ is the **tangent space of $M$ at $P$** denoted $T_PM$.
- It is a vector space of dimension $n$: let $\theta$ a local parametrization of $M$, $\theta(x_1, \ldots, x_n) \in M$ with $\theta(0) = P$. Define curves

$$x_i : t \mapsto \theta(0, \ldots, 0, t, 0, 0)$$

- They go through $P$ when $t = 0$ and follow the axes. Their derivative at $0$ are denoted $\partial_{x_i}$. They form a basis of $T_PM$. 

![Diagram of tangent space](image-url)
A vector field is a smooth map that sends $P \in M$ to a vector $v(P) \in T_PM$. 
Differential of a differentiable map

- $f : M \rightarrow N$ differentiable, $P \in M$, $f(P) \in N$
- $d_P f : T_PM \rightarrow T_{f(P)} N$ linear map corresponding to the Jacobian matrix of $f$ in local coordinates.
- When $N = \mathbb{R}$, $d_P f$ is a linear form $T_PM \rightarrow \mathbb{R}$. 
Differential of a differentiable map

- $f : M \to N$ differentiable, $P \in M$, $f(P) \in N$
- $d_P f : T_P M \to T_{f(P)} N$ linear map corresponding to the Jacobian matrix of $f$ in local coordinates.
- When $N = \mathbb{R}$, $d_P f$ is a linear form $T_P M \to \mathbb{R}$. 
Differential of a differentiable map

- $f: M \to N$ differentiable, $P \in M$, $f(P) \in N$
- $d_P f : T_P M \to T_{f(P)} N$ linear map corresponding to the Jacobian matrix of $f$ in local coordinates.
- When $N = \mathbb{R}$, $d_P f$ is a linear form $T_P M \to \mathbb{R}$.
Outline

1 Motivation
   - Non Linearity
   - Statistics on Non Linear Data

2 Recalls
   - Geometry
   - Topology
   - Calculus on $\mathbb{R}^n$

3 Differentiable Manifolds
   - Definitions
   - Building Manifolds
   - Tangent Space

4 Riemannian Manifolds
   - Metric
   - Gradient Field
   - Length of curves
   - Geodesics
   - Covariant derivatives
A Riemannian metric on a $n$–dimensional manifold is a smooth family $g_P$ of inner products on the tangent spaces $T_PM$ of $M$, $u,v \in T_PM \mapsto g_P(u,v) := \langle u,v \rangle_P \in \mathbb{R}$. With it, one can compute length of vectors in tangent spaces, check orthogonality of them...

With a local parametrization $\theta(x) = (x_1, \ldots, x_n) \to M$, it corresponds to a smooth family of positive definite matrices:

$$g_x = \begin{pmatrix}
g_{x11} & \cdots & g_{x1n} \\
\vdots & \ddots & \vdots \\
g_{xn1} & \cdots & g_{xnn}
\end{pmatrix}$$

$$u = \sum_{i=1}^{n} u_i \partial_{x_i}, \quad v = \sum_{i=1}^{n} v_i \partial_{x_i}, \quad \langle u,v \rangle_x = (u_1, \ldots, u_n) g_x(v_1, \ldots, v_n)^t$$
A Riemannian metric on a $n$–dimensional manifold is a smooth family $g_P$ of inner products on the tangent spaces $T_P M$ of $M$, $u, v \in T_P M \mapsto g_P(u, v) := \langle u, v \rangle_P \in \mathbb{R}$. With it, one can compute length of vectors in tangent spaces, check orthogonality of them...

With a local parametrization $\theta(x) = (x_1, \ldots, x_n) \to M$, it corresponds to a smooth family of positive definite matrices:

$$g_x = \begin{pmatrix} g_{x11} & \cdots & g_{x1n} \\ \vdots & \ddots & \vdots \\ g_{xnn} \end{pmatrix}$$

$$u = \sum_{i=1}^n u_i \partial_{x_i}, \quad v = \sum_{i=1}^n v_i \partial_{x_i}, \quad \langle u, v \rangle_x = (u_1, \ldots, u_n) g_x (v_1, \ldots, v_n)^t$$
A Riemannian metric on a $n$–dimensional manifold is a smooth family $g_P$ of inner products on the tangent spaces $T_PM$ of $M$, $u, v \in T_PM \mapsto g_P(u, v) := \langle u, v \rangle_P \in \mathbb{R}$. With it, one can compute length of vectors in tangent spaces, check orthogonality of them...

With a local parametrization $\theta(x) = (x_1, \ldots, x_n) \to M$, it corresponds to a smooth family of positive definite matrices:

$$g_x = \begin{pmatrix} g_{x11} & \cdots & g_{x1n} \\ \vdots & \ddots & \vdots \\ g_{xn1} & \cdots & g_{xnn} \end{pmatrix}$$

$$u = \sum_{i=1}^{n} u_i \partial x_i, \ v = \sum_{i=1}^{n} v_i \partial x_i \quad \langle u, v \rangle_x = (u_1, \ldots, u_n)g_x(v_1, \ldots, v_n)^t$$
A differential manifold with a Riemannian metric is a **Riemannian manifold**.
Outline

1. Motivation
   - Non Linearity
   - Statistics on Non Linear Data

2. Recalls
   - Geometry
   - Topology
   - Calculus on $\mathbb{R}^n$

3. Differentiable Manifolds
   - Definitions
   - Building Manifolds
   - Tangent Space

4. Riemannian Manifolds
   - Metric
   - Gradient Field
   - Length of curves
   - Geodesics
   - Covariant derivatives
Gradient, Gradient vector field.

- $M$ Riemannian, $f : M \to \mathbb{R}$ differentiable. Then
  \[ d_P f(h) = \langle v, h \rangle_P, \]
  for a unique $v$.

- $v := \nabla f_P$ is the gradient of $f$ at $P$.
- $P \mapsto \nabla f_P$ is the gradient vector field of $f$.
- One can thus make gradient descent/ascent... Not possible without Riemannian structure.
Gradient, Gradient vector field.

- $M$ Riemannian, $f : M \rightarrow \mathbb{R}$ differentiable. Then
  \[ d_P f(h) = \langle v, h \rangle_P, \quad \text{for a unique } v. \]

- $v := \nabla f_P$ is the gradient of $f$ at $P$.
- $P \mapsto \nabla f_P$ is the gradient vector field of $f$.
- One can thus make gradient descent/ascent... Not possible without Riemannian structure.
Riemannian Manifolds

Gradient Field

Gradient, Gradient vector field.

- $M$ Riemannian, $f : M \to \mathbb{R}$ differentiable. Then

  $$d_P f(h) = \langle v, h \rangle_P,$$

  for a unique $v$.

- $v := \nabla f_P$ is the gradient of $f$ at $P$.
- $P \mapsto \nabla f_P$ is the gradient vector field of $f$.
- One can thus make gradient descent/ascent... Not possible without Riemannian structure.
Gradient, Gradient vector field.

- $M$ Riemannian, $f : M \to \mathbb{R}$ differentiable. Then
  
  $$d_P f(h) = \langle v, h \rangle_P,$$

  for a unique $v$.

- $v := \nabla f_P$ is the gradient of $f$ at $P$.
- $P \mapsto \nabla f_P$ is the gradient vector field of $f$.

One can thus make gradient descent/ascent... Not possible without Riemannian structure.
Gradient, Gradient vector field.

- $M$ Riemannian, $f : M \to \mathbb{R}$ differentiable. Then 
  \[ d_P f(h) = \langle v, h \rangle_P, \quad \text{for a unique } v. \]

- $v := \nabla f_P$ is the gradient of $f$ at $P$.
- $P \mapsto \nabla f_P$ is the gradient vector field of $f$.
- One can thus make gradient descent/ascent... Not possible without Riemannian structure.
Outline

1 Motivation
   - Non Linearity
   - Statistics on Non Linear Data

2 Recalls
   - Geometry
   - Topology
   - Calculus on $\mathbb{R}^n$

3 Differentiable Manifolds
   - Definitions
   - Building Manifolds
   - Tangent Space

4 Riemannian Manifolds
   - Metric
   - Gradient Field
   - Length of curves
   - Geodesics
   - Covariant derivatives
Curve $c : [a, b] \rightarrow M$, $M$ Riemannian. For all $t$, $c'(t) \in T_{c(t)}M$ is the velocity of $c$ at time $t$.

- It has length $\|\dot{c}(t)\| = \sqrt{\langle c'(t), c'(t) \rangle_{c(t)}}$

- Define the length of $c$ as

$$\ell(c) = \int_{a}^{b} \|\dot{c}(t)\| \, dt$$

as in the Euclidean case, by now with variable inner products.
Curve \( c : [a, b] \to M, M \) Riemannian. For all \( t \), \( c'(t) \in T_{c(t)}M \) is the velocity of \( c \) at time \( t \).

- It has length \( \|\dot{c}(t)\| = \sqrt{\langle c'(t), c'(t) \rangle_{c(t)}} \)
- Define the length of \( c \) as

\[
\ell(c) = \int_{a}^{b} \|\dot{c}(t)\| \, dt
\]

as in the Euclidean case, by now with variable inner products.
Curve $c : [a, b] \rightarrow M$, $M$ Riemannian. For all $t$, $c'(t) \in T_{c(t)}M$ is the velocity of $c$ at time $t$.

It has length $\|\dot{c}(t)\| = \sqrt{\langle c'(t), c'(t) \rangle_{c(t)}}$

Define the length of $c$ as

$$\ell(c) = \int_{a}^{b} \|\dot{c}(t)\| \, dt$$

as in the Euclidean case, by now with variable inner products.
Curve $c : [a, b] \rightarrow M$, $M$ Riemannian. For all $t$, $c'(t) \in T_{c(t)}M$ is the velocity of $c$ at time $t$.

It has length $\|\dot{c}(t)\| = \sqrt{\langle c'(t), c'(t) \rangle_{c(t)}}$

Define the **length of $c$** as

$$\ell(c) = \int_{a}^{b} \|\dot{c}(t)\| \, dt$$

as in the Euclidean case, by now with variable inner products.
Outline

1 Motivation
   • Non Linearity
   • Statistics on Non Linear Data

2 Recalls
   • Geometry
   • Topology
   • Calculus on $\mathbb{R}^n$

3 Differentiable Manifolds
   • Definitions
   • Building Manifolds
   • Tangent Space

4 Riemannian Manifolds
   • Metric
   • Gradient Field
   • Length of curves
   • Geodesics
   • Covariant derivatives
Geodesics

- Restricted definition: Riemannian Geodesics are curves of (locally) minimal length among curves with fixed endpoints say $P$ and $Q$.
- They are also minimizers of the curve energy:

  $$E(c) = \int_{a}^{b} \|\dot{c}(t)\|^2 \, dt$$

- The shortest length of a curve joining $P$ and $Q$ is the geodesic distance $d(P, Q)$. 
Geodesics

- **Restricted definition**: Riemannian Geodesics are curves of (locally) minimal length among curves with fixed endpoints say $P$ and $Q$.
- They are also minimizers of the curve energy:

$$E(c) = \int_{a}^{b} \| \dot{c}(t) \|^2 \, dt$$

- The shortest length of a curve joining $P$ and $Q$ is the geodesic distance $d(P, Q)$. 

**Geodesics**

- Restricted definition: **Riemannian Geodesics** are curves of (locally) minimal length among curves with fixed endpoints say $P$ and $Q$.
- They are also minimizers of the *curve energy*:

$$E(c) = \int_{a}^{b} \| \dot{c}(t) \|^2 \, dt$$

- The shortest length of a curve joining $P$ and $Q$ is the *geodesic distance* $d(P, Q)$. 

Geodesics

- Restricted definition: **Riemannian Geodesics** are curves of (locally) minimal length among curves with fixed endpoints say $P$ and $Q$.
- They are also minimizers of the **curve energy**:

  \[ E(c) = \int_{a}^{b} \| \dot{c}(t) \|^2 \, dt \]

- The shortest length of a curve joining $P$ and $Q$ is the **geodesic distance** $d(P, Q)$. 
How to characterize geodesics?

- In $\mathbb{R}^n$, The calculus of variations for curve energy gives: $\ddot{c} = 0$.
- In a general manifold: problem to define $\ddot{c}$:

$$
\ddot{c}(0) = \lim_{t \to 0} \frac{\dot{c}(t) - \dot{c}(0)}{t}
$$

$\dot{c}(t) \in T_{c(t)}M$ and $\dot{c}(0) \in T_{c(0)}M$: these tangent spaces are distinct!

- Need for a “device” that “connects” tangent spaces of close enough points. Such a device is called an affine connection.
How to characterize geodesics?

- In $\mathbb{R}^n$, the calculus of variations for curve energy gives: $\ddot{c} = 0$.
- In a general manifold: problem to define $\ddot{c}$:

$$\ddot{c}(0) = \lim_{t \to 0} \frac{\dot{c}(t) - \dot{c}(0)}{t}$$

$\dot{c}(t) \in T_{c(t)}M$ and $\dot{c}(0) \in T_{c(0)}M$: these tangent spaces are distinct!

- Need for a “device” that “connects” tangent spaces of close enough points. Such a device is called an affine connection.
How to characterize geodesics?

- In $\mathbb{R}^n$, the calculus of variations for curve energy gives: $\ddot{c} = 0$.
- In a general manifold: problem to define $\ddot{c}$:

$$
\ddot{c}(0) = \lim_{t \to 0} \frac{\dot{c}(t) - \dot{c}(0)}{t}
$$

$\dot{c}(t) \in T_{c(t)}M$ and $\dot{c}(0) \in T_{c(0)}M$: these tangent spaces are distinct!

Need for a “device” that “connects” tangent spaces of close enough points. Such a device is called an affine connection.
How to characterize geodesics?

- In $\mathbb{R}^n$, the calculus of variations for curve energy gives: $\ddot{c} = 0$.
- In a general manifold: problem to define $\ddot{c}$:

$$
\ddot{c}(0) = \lim_{t \to 0} \frac{\dot{c}(t) - \dot{c}(0)}{t}
$$

$\dot{c}(t) \in T_{c(t)}M$ and $\dot{c}(0) \in T_{c(0)}M$: these tangent spaces are distinct!

Need for a “device” that “connects” tangent spaces of close enough points. Such a device is called an affine connection.
How to characterize geodesics?

- In $\mathbb{R}^n$, the calculus of variations for curve energy gives: $\ddot{c} = 0$.
- In a general manifold: problem to define $\ddot{c}$:

$$\ddot{c}(0) = \lim_{t \to 0} \frac{\dot{c}(t) - \dot{c}(0)}{t}$$

$\dot{c}(t) \in T_{c(t)}M$ and $\dot{c}(0) \in T_{c(0)}M$: these tangent spaces are distinct!

- Need for a “device” that “connects” tangent spaces of close enough points. Such a device is called an affine connection.
Outline

1 Motivation
   - Non Linearity
   - Statistics on Non Linear Data

2 Recalls
   - Geometry
   - Topology
   - Calculus on $\mathbb{R}^n$

3 Differentiable Manifolds
   - Definitions
   - Building Manifolds
   - Tangent Space

4 Riemannian Manifolds
   - Metric
   - Gradient Field
   - Length of curves
   - Geodesics
   - Covariant derivatives
Covariant Derivative

- Allows to differentiate a vector field along a curve: given a curve \( \gamma(t) \in M \), \( X \) a vector field,

\[
\frac{D}{dt} X(t) = \dot{X}(t) \in T_{\gamma(t)}
\]

We ask that \( \frac{D}{dt} \) depends only on the value \( \dot{\gamma}(t) \) and not on the behaviour of \( \gamma \) around \( \gamma(t) \). The computation \( \frac{D}{dt} X(t) \) depends on values of \( X \) around \( \gamma(t) \).

- Many choices are possible, but exactly one is compatible with the Riemannian structure in the sense that

\[
\frac{d}{dt} \langle X, Y \rangle = \langle \frac{DX}{dt}, Y \rangle + \langle X, \frac{DY}{dt} \rangle
\]

plus another property. Levi-Civita connexion.
Concrete construction

Assume $M \subset \mathbb{R}^n$. A vector field $X$ on $M$ can be seen as a vector field on $\mathbb{R}^n$ and a curve $\gamma$ on $M$ can be seen as a curve in $\mathbb{R}^n$. Then

1. Compute the usual derivative

\[
\tilde{X}(t) = \frac{d}{dt} X(\gamma(t))
\]

its a vector field on $\mathbb{R}^n$ but not a tangent vector field on $M$ in general.

2. Project $\tilde{X}(t)$ orthogonally on $T_{\gamma(t)} M \subset \mathbb{R}^3$. The result is $DX/dt$!
A curve $\gamma$ is geodesic if its covariant is acceleration 0.

$$\ddot{\gamma}(t) = \frac{D\dot{\gamma}(t)}{dt} = 0!$$

This is in fact a second order ODE: given initial position $\gamma(0)$ and velocity $\dot{\gamma}(0)$ there is a unique solution.
The uniqueness above leads to the following definition: given $P \in M$, $v \in T_PM$, the exponential map $\text{Exp}_P(v)$ is the solution at time 1 of the previous ODE. For small enough $v$: diffeomorphism.

- The curve $t \mapsto \text{Exp}_P(tv)$, $t \in [0, 1]$ is geodesic, its length is $\|v\|$. 

François Lauze (University of Copenhagen)
The inverse map of the exponential map is called the Log map! For $Q \in M$ “not too far from $P$”, $\text{Log}_P(Q)$ is the vector $v$ of $T_P M$ s.t. $\text{Exp}_P(v) = Q$.

The exponential map is relatively easy to compute. The Log map is generally much more complicated, but badly needed in many optimization problems!
• Boothby: Introduction to Differential Manifolds and Riemannian Geometry, Wiley
• do Carmo: Riemannian Geometry, Birkhäuser.
• Hulin-Lafontaine: Riemannian Geometry, Springer
• Small: The statistical theory of shapes, Springer.