## Analysis of

## Three-Dimensional Shapes

 (IN2238, TU München, Summer 2015)
## Euclidean Embeddings

 (27.04.2015)Dr. Emanuele Rodolà<br>rodola@in.tum.de<br>Room 02.09.058, Informatik IX

## Wrap-up

In the previous lectures, we have approached the problem of shape similarity...


## Wrap-up

In the previous lectures, we have approached the problem of shape similarity... and shape matching


Can we find a correspondence?


## Wrap-up

We modeled our shapes as metric spaces, that is, a set of points plus a metric (distance) function defined over it.

$$
\left(X, d_{X}\right)
$$

$$
d_{X}\left(x_{1}, x_{2}\right)
$$

## Wrap-up

We decided that the Gromov-Hausdorff distance captures the notion of shape similarity in the most natural way. Then we turned to the problem of actually computing this distance.

The Gromov-Hausdorff distance between two metric spaces $X$ and $Y$ is defined by

$$
d_{G \mathscr{H}}(X, Y)=\inf _{Z, f, g} d_{\Re}^{Z}(f(X), g(Y))
$$

The infimum is taken over all ambient spaces $Z$ and isometric embeddings (distance preserving) $f: X \rightarrow Z, g: Y \rightarrow Z$
$d_{G \mathscr{H}}$ is a metric on the space of isometry classes of compact metric spaces.

## Wrap-up

Passing to correspondences, we wrote:

$$
\begin{gathered}
d_{G \mathscr{H}}(X, Y)=\inf _{Z, f, g} d_{\mathscr{H}}^{Z}(f(X), g(Y)) \\
d_{G \mathscr{H}}(X, Y)=\frac{1}{2} \inf _{R} \operatorname{dis} R \\
\left.d_{G \mathscr{H}}(\mathbf{X}, \mathbf{Y})=\frac{1}{2} \inf _{R \subset \mathbf{X} \times \mathbf{Y}_{(x, y),\left(x^{\prime}, y^{\prime}\right) \in R}} \sup _{X}\left(x, x^{\prime}\right)-d_{Y}\left(y, y^{\prime}\right) \right\rvert\,
\end{gathered}
$$

## Wrap-up

Since the original problem seems difficult to solve, we had a look at a few possible relaxations. A "relaxation" is an approximation of a difficult problem by another similar problem that is easier to solve.

Hopefully, the solution to the relaxed problem will provide some information about the original solution.


## Wrap-up

First relaxation: replace the max with a sum.

$$
\begin{aligned}
& \frac{1}{2} \min _{R} \max _{i, j \ell, m} C_{(i \ell)(j m)} R_{i j} R_{\ell m} \\
& \frac{1}{2} \min _{R} \sum_{i, j} \sum_{\ell, m} C_{(i \ell)(j m)} R_{i j} R_{\ell m}
\end{aligned}
$$

## Wrap-up

Simplify using the dreadful matrix notation.

$$
\frac{1}{2} \min _{R} \sum_{i, j} \sum_{\ell, m} C_{(i \ell)(j m)} R_{i j} R_{\ell m}
$$



$$
\begin{aligned}
\min _{R \in\left\{0,11^{1 \times n}\right.} & \operatorname{vec}\{R\}^{\mathrm{T}} C \operatorname{vec}\{R\} \\
\text { s.t. } & R \mathbf{1}=\mathbf{1}, R^{\mathrm{T}} \mathbf{1}=\mathbf{1}
\end{aligned}
$$

## Wrap-up

Second relaxation: replace binary solutions with continuous solutions.

$$
R \in\{0,1\}^{n \times n} \rightarrow R \in[0,1]^{n \times n}
$$



## Wrap-up

Second relaxation: replace binary solutions with continuous solutions.

$$
R \in\{0,1\}^{n \times n} \rightarrow R \in[0,1]^{n \times n}
$$



## Wrap-up

Other relaxations: replace the mapping constraints...

$$
\begin{aligned}
& R \mathbf{1}=\mathbf{1}, R^{\mathrm{T}} \mathbf{1}=\mathbf{1} \\
& \|R\|^{2}=1 \\
& \mathbf{1}^{\mathrm{T}} R \mathbf{1}=1
\end{aligned}
$$

...and / or replace the cost function

$$
\begin{aligned}
& C_{(i \ell)(j m)}=\left|d_{\mathbf{x}}\left(x_{i}, x_{j}\right)-d_{\mathbf{Y}}\left(y_{\ell}, y_{m}\right)\right|^{p} \\
& C_{(i \ell)(j m)}=e^{-\beta\left|d_{\mathbf{x}}\left(x_{i}, x_{j}\right)-d_{\mathbf{Y}}\left(y_{\ell}, y_{m}\right)\right|^{2}}
\end{aligned}
$$

## Hausdorff revisited

Recall that the Gromov-Hausdorff distance is defined in terms of the Hausdorff distance:

$$
\begin{aligned}
& d_{\mathscr{H} \mathscr{H}}(X, Y)=\inf _{Z, f, g} d_{\mathcal{H}}^{Z}(f(X), g(Y)) \\
& \text { where } f: X \rightarrow Z, g: Y \rightarrow Z \text { are isometric embeddings }
\end{aligned}
$$

## Hausdorff revisited

Recall that the Gromov-Hausdorff distance is defined in terms of the Hausdorff distance:

The Hausdorff distance between two compact subsets $X, Y \subset\left(Z, d_{Z}\right)$ is defined by

$$
d_{\mathscr{H}}^{Z}(X, Y)=\max \left\{\sup _{x \in X} \operatorname{dist}_{\mathrm{Z}}(x, Y), \sup _{y \in Y} \operatorname{dis}_{\mathrm{Z}}(y, X)\right\}
$$



## Hausdorff revisited

The Hausdorff distance between two compact subsets $X, Y \subset\left(Z, d_{Z}\right)$ is defined by

$$
d_{\mathscr{H}}^{Z}(X, Y)=\max \left\{\sup _{x \in X} \operatorname{dist}_{\mathrm{Z}}(x, Y), \sup _{y \in Y} \operatorname{dist}_{\mathrm{Z}}(y, X)\right\}
$$



## Hausdorff revisited

The Hausdorff distance between two compact subsets $X, Y \subset\left(Z, d_{Z}\right)$ is defined by

$$
d_{\mathscr{H}}^{Z}(X, Y)=\max \left\{\sup _{x \in X} \operatorname{dist}_{\mathrm{Z}}(x, Y), \sup _{y \in Y} \operatorname{dist}_{\mathrm{Z}}(y, X)\right\}
$$



## Hausdorff revisited

The Hausdorff distance between two compact subsets $X, Y \subset\left(Z, d_{Z}\right)$ is defined by

$$
d_{\mathscr{H}}^{Z}(X, Y)=\max \left\{\sup _{x \in X} \operatorname{dist}_{\mathrm{Z}}(x, Y), \sup _{y \in Y} \operatorname{dist}_{\mathrm{Z}}(y, X)\right\}
$$



## Euclidean embeddings

Let's have a look at the Gromov-Hausdorff distance again:

$$
\begin{aligned}
& d_{G \mathscr{H}}(X, Y)=\inf _{Z, f, g} d_{\mathscr{H}}^{Z}(f(X), g(Y)) \\
& \text { where } f: X \rightarrow Z, g: Y \rightarrow Z \text { are isometric embeddings }
\end{aligned}
$$

From the previous example we have seen that optimizing for rigid transformations is much simpler and many effective algorithms exist (ICP).

Then can't we just map each shape to $\mathbf{R}^{3}$ and then solve the resulting rigid problem there?

## Euclidean embeddings

$$
d_{G \mathscr{H}}(X, Y)=\inf _{\mathbf{R}^{3}, f, g} d_{\mathscr{H}}^{\mathbf{R}^{3}}(f(X), g(Y))
$$

where $f: X \rightarrow \mathbf{R}^{3}, g: Y \rightarrow \mathbf{R}^{3}$ are isometric embeddings
In other words, we are looking for something like:


$$
\left(x, d_{X}\right) \Rightarrow(f(X),\|\cdot\|)
$$

$$
(g(Y),\| \|) \curvearrowright\left(Y, d_{Y}\right)
$$

## Euclidean embeddings



$$
\left(X, d_{X}\right) \quad \square \quad(f(X),\|\cdot\|)
$$

Thus, we would like to find a map $f:\left(X, d_{X}\right) \rightarrow\left(\mathbf{R}^{m},\|\cdot\|\right)$ such that

$$
d_{X}\left(x, x^{\prime}\right)=\left\|f(x)-f\left(x^{\prime}\right)\right\|_{2}
$$

for all $x, x^{\prime} \in X$

## Euclidean embeddings



$$
\left(X, d_{X}\right) \quad \Rightarrow \quad(f(X),\|\cdot\|)
$$

The image $f(X)$ is also called the canonical form of $X$.
It defines an equivalence class of shapes up to an isometry in $\mathbf{R}^{m}$ (these correspond to rotations, translations, reflections).

In other words, we are reducing intrinsic isometries into extrinsic isometries.

## Euclidean embeddings

Note that:

- We are assuming $m$ to be arbitrary (i.e. not necessarily $m=3$ ). This allows us to keep the approach general, and to speak about dimensionality reduction.
- Topological noise can significantly alter distances.



## A cartographer's problem

- We still don't know to what extent our shapes $X$ are «isometrically embeddable» into $\mathbf{R}^{m}$ !


$$
d_{S} \stackrel{?}{=} d_{\mathbf{R}^{2}}
$$

Impossible to do without introducing distortion.

## The smallest non-trivial example



$$
\left.D_{S}=\left(\begin{array}{llll}
x_{1} & x_{2} & x_{3} & x_{4} \\
0 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 \\
1 & 1 & 0 & 2 \\
1 & 1 & 2 & 0
\end{array}\right)\right)^{x_{1}} x_{2}
$$

Assume $\left(D_{S}\right)_{i j}=d_{\mathbf{R}^{m}}\left(z_{i}, z_{j}\right)$ and consider the triangle $z_{3}, z_{1}, z_{4}$


Now consider the triangle $z_{3}, z_{2}, z_{4}$


Then $z_{1}=z_{2}$, which contradicts $\left(D_{S}\right)_{12}=d_{\mathbf{R}^{m}}\left(z_{1}, z_{2}\right)=1$

This metric space cannot be embedded into a Euclidean space of any finite dimension!

## Minimum-distortion embedding

Still, we could try to look for an approximate embedding, such that the distortion of $d_{X}$ is mimimal according to some criterion.

One such criterion is the usual metric distortion induced by the mapping $f$ :

$$
\operatorname{dis} f=\sup _{x_{i}, x_{j} \in X}\left|d_{X}\left(x_{i}, x_{j}\right)-d_{\mathbf{R}^{m}}\left(f\left(x_{i}\right), f\left(x_{j}\right)\right)\right|
$$

A minimum-distortion embedding would then be the $f$ minimizing the above.

## Minimum-distortion embedding

$$
\operatorname{dis} f=\sup _{x_{i}, x_{j} \in X}\left|d_{X}\left(x_{i}, x_{j}\right)-d_{\mathbf{R}^{m}}\left(f\left(x_{i}\right), f\left(x_{j}\right)\right)\right|
$$

We can define alternative measures of distortion as well, for instance:

$$
\sigma_{p}(f)=\sum_{i>j}\left|d_{X}\left(x_{i}, x_{j}\right)-d_{\mathbf{R}^{m}}\left(f\left(x_{i}\right), f\left(x_{j}\right)\right)\right|^{p}
$$

Keep in mind the the resulting canonical form $f(X)$ will only be an approximation. The embedding introduces a distortion, which in turn influences the accuracy of our similarity calculations.

## Minimum-distortion embedding

We will consider the quadratic stress $\sigma_{2}(f)$. Then we would like to compute:

$$
f=\underset{f: X \rightarrow \mathbf{R}^{m}}{\arg \min } \sum_{i>j}\left|d_{X}\left(x_{i}, x_{j}\right)-d_{\mathbf{R}^{m}}\left(f\left(x_{i}\right), f\left(x_{j}\right)\right)\right|^{2}
$$

Let us consider a sampling $\left\{x_{1}, \ldots, x_{N}\right\}$ of $N$ points over $X$, and denote their images as $z_{i}=f\left(x_{i}\right)$. Arranging the $z_{i}$ into a $N \times m$ matrix $Z=\left(z_{i}^{j}\right)$, we can rewrite the distortion criterion as

$$
\sigma_{2}\left(Z, D_{X}\right)=\sum_{i>j}\left|d_{X}\left(x_{i}, x_{j}\right)-d_{i j}(Z)\right|^{2}
$$

where $d_{i j}(Z)=\left\|z_{i}-z_{j}\right\|_{2}$
Differently from the matching problem, now $Z$ is the unknown!

## Minimum-distortion embedding

$$
Z^{*}=\underset{Z \in \mathbf{R}^{N \times m}}{\arg \min } \sigma_{2}(Z)
$$

Note that there is no unique solution, in fact applying any Euclidean isometry to $Z^{*}$ will not change the value of $\sigma_{2}$.
Problems of this sort started appearing in psychology in the 1950's, and are usually referred to as multidimensional scaling (MDS) problems.

```
PSYCHOMETRIKA-vol. 27, No. 2
    JuNE, 1962
```

THE ANALYSIS OF PROXIMITIES: MULTIDIMENSIONAL SCALING WITH AN UNKNOWN DISTANCE FUNCTION. I.

Roger N. Shepard

## Multidimensional scaling

Empirical procedures of several diverse kinds have this in common: they start with a fixed set of entities and determine, for every pair of these, a number reflecting how closely the two entities are related psychologically. The nature of the psychological relation depends upon the nature of the entities. If the entities are all stimuli or all responses, we are inclined to think of the relation as one of similarity. A somewhat more objective (though less intuitive) characterization of such a relation, perhaps, is that of substitutability. The statement that stimulus $A$ is more similar to $B$ than to $C$, for example, could be interpreted to say that the psychological (or behavioral) consequences are greater when $C$, rather than $B$, is substituted for $A$. From this standpoint a natural procedure for determining similarities of stimuli or responses is by recording substitution errors during identification learning [ $2,7,12,14,17,18]$. In addition, though, disjunctive reaction time and sorting time have also been proposed as measures of psychological similarity [20]. Finally, of course, individuals have sometimes been instructed simply to rate each pair of stimuli, directly, on a scale of apparent similarity [ 1,6 ]. The notion of similarity is not necessarily restricted to stimuli or responses (in the narrow sense of these words), however. Serviceable measures of similarity may also be found for concepts, attitudes, personality structures, or even social institutions. political svstems. and the like.

## Applications: visualizing the space of shapes



Extrinsic metric (ICP)


Intrinsic metric (Gromov-Hausdorff)

## Applications: intrinsic alignment of shape spaces



## Applications: intrinsic alignment of shape spaces



## Canonical shape analysis

When we consider distance functions defined over our shapes (as opposed to distance between generic entities, see previous examples), then we talk about canonical shape analysis.

This is still an active area of research, with new methods tackling sensitivity to topological noise and sampling, efficiency, and distortion to name some relevant aspects.


## Applications: shape matching



Shape 1


Shape 2


ICP result

Compute canonical forms and then rigidly align them to obtain point-topoint matches.

## Applications: shape retrieval



Compute distance between shapes as the maximum Euclidean distance between corresponding points after rigid alignment.

## Applications: symmetry detection



Computing rigid bilateral symmetries is close to be a solved problem, and it involves solving over planes in $\mathbf{R}^{3}$.

## Minimum-distortion embedding

We are going to solve:

$$
\begin{aligned}
& Z^{*}=\underset{Z \in \mathbf{R}^{\wedge \times m}}{\arg \min } \sigma_{2}(Z) \\
& \sigma_{2}\left(Z, D_{X}\right)=\sum_{i>j}\left|d_{X}\left(x_{i}, x_{j}\right)-d_{i j}(Z)\right|^{2}
\end{aligned}
$$

We will first rewrite the problem above in a more friendly format, and then minimize via gradient descent.

## Quadratic stress

$$
\sigma_{2}\left(Z, D_{X}\right)=\sum_{i>j}\left|d_{X}\left(x_{i}, x_{j}\right)-d_{i j}(Z)\right|^{2}
$$

For any given configuration $Z$, the stress measures how well that configuration matches the data. We look for the configuration of minimum stress.

Let's rewrite the stress function differently:

$$
\begin{aligned}
\sigma_{2}\left(Z, D_{X}\right) & =\sum_{i>j}\left|d_{X}\left(x_{i}, x_{j}\right)-d_{i j}(Z)\right|^{2} \\
& =\sum_{i>j}^{\sum_{i j}} \underbrace{d_{i j}^{2}(Z)}_{\text {Term 1 }}-\underbrace{2 d_{i j}(Z) d_{X}\left(x_{i}, x_{j}\right)}_{\text {Term 2 }}+d_{X}^{2}\left(x_{i}, x_{j}\right)
\end{aligned}
$$

## Quadratic stress (Term 1)

$$
\begin{aligned}
\sum_{i>j} & d_{i j}^{2}(Z)= \\
& =\sum_{i>j}\left\|z_{i}-z_{j}\right\|^{2}=\sum_{i>j} \sum_{k=1}^{m}\left(z_{i}^{k}-z_{j}^{k}\right)^{2} \\
& =\sum_{i>j} \sum_{k=1}^{m}\left(z_{i}^{k}\right)^{2}-2 z_{i}^{k} z_{j}^{k}+\left(z_{j}^{k}\right)^{2} \\
& =\sum_{i>j}\left\langle z_{j}, z_{j}\right\rangle+\left\langle z_{i}, z_{i}\right\rangle-2\left\langle z_{i}, z_{j}\right\rangle=\sum_{i>i}\left\langle z_{j}, z_{j}\right\rangle+\left\langle z_{i}, z_{i}\right\rangle-2 \sum_{i>j}\left\langle z_{i}, z_{j}\right\rangle \\
& =(N-1) \sum_{i=1}^{N}\left\langle z_{i}, z_{i}\right\rangle-\left(\sum_{i, j}\left\langle z_{i}, z_{j}\right\rangle-\sum_{i=1}^{N}\left\langle z_{i}, z_{i}\right\rangle\right) \\
& =N \sum_{i=1}^{N}\left\langle z_{i}, z_{i}\right\rangle-\sum_{i, j}\left\langle z_{i}, z_{j}\right\rangle
\end{aligned}
$$

## Quadratic stress (Term 1)

$$
\begin{aligned}
\sum_{i>j} d_{i j}^{2}(Z) & =N \sum_{i=1}^{N}\left\langle z_{i}, z_{i}\right\rangle-\sum_{i, j}\left\langle z_{i}, z_{j}\right\rangle & & \\
& =N \operatorname{tr}\left(Z Z^{\mathrm{T}}\right)-\operatorname{tr}\left(1_{N \times N} Z Z^{\mathrm{T}}\right) & & 1_{N \times N} \text { is a matrix of ones } \\
& =\operatorname{tr}\left(V Z Z^{\mathrm{T}}\right) & & V_{i j}=\left\{\begin{array}{cc}
-1 & i \neq j \\
N-1 & i=j
\end{array}\right. \\
& =\operatorname{tr}\left(Z^{\mathrm{T}} V Z\right) & &
\end{aligned}
$$

The last step can be done because
$\operatorname{tr}(A B)=\sum_{i=1}^{m}(A B)_{i i}=\sum_{i=1}^{m} \sum_{j=1}^{n} A_{i j} B_{j i}=\sum_{j=1}^{n} \sum_{i=1}^{m} B_{j i} A_{i j}=\sum_{j=1}^{n}(B A)_{j j}=\operatorname{tr}(B A)$

## Quadratic stress (Term 2)

$$
\begin{aligned}
\sigma_{2}\left(Z, D_{X}\right)= & \underbrace{\sum_{i j} d_{i j}^{2}(Z)}_{i>j}-2 d_{i j}(Z) d_{X}\left(x_{i}, x_{j}\right)
\end{aligned}+d_{X}^{2}\left(x_{i}, x_{j}\right)
$$

$$
\begin{aligned}
\sum_{i>j} d_{i j}(Z) d_{X}\left(x_{i}, x_{j}\right) & =\sum_{i>j} d_{X}\left(x_{i}, x_{j}\right) d_{i j}^{-1}(Z) d_{i j}^{2}(Z) \\
& =\sum_{i>j} d_{X}\left(x_{i}, x_{j}\right) d_{i j}^{-1}(Z) \sum_{k=1}^{m}\left(z_{i}^{k}-z_{j}^{k}\right)^{2} \\
& =\sum_{i>j} d_{X}\left(x_{i}, x_{j}\right) d_{i j}^{-1}(Z)\left(\left\langle z_{i}, z_{i}\right\rangle+\left\langle z_{j}, z_{j}\right\rangle-2\left\langle z_{i}, z_{j}\right\rangle\right)
\end{aligned}
$$

## Quadratic stress (Term 2)

$$
\begin{aligned}
\sum_{i>j} d_{i j}(Z) d_{X}\left(x_{i}, x_{j}\right) & =\sum_{i>j} \underbrace{d_{X}\left(x_{i}, x_{j}\right) d_{i j}^{-1}(Z)}_{a_{i j}=a_{j i}\left(a_{i j}=0 \text { for } i=j\right)}\left(\left\langle z_{i}, z_{i}\right\rangle+\left\langle z_{j}, z_{j}\right\rangle-2\left\langle z_{i}, z_{j}\right\rangle\right) \\
& =\sum_{i>j} a_{i j}\left(\left\langle z_{i}, z_{i}\right\rangle+\left\langle z_{j}, z_{j}\right\rangle-2\left\langle z_{i}, z_{j}\right\rangle\right) \\
& =\sum_{i>j} a_{i j}\left(\left\langle z_{i}, z_{i}\right\rangle-\left\langle z_{i}, z_{j}\right\rangle\right)+\sum_{i>j} a_{j i}\left(\left\langle z_{j}, z_{j}\right\rangle-\left\langle z_{j}, z_{i}\right\rangle\right) \\
& =\sum_{i, j} a_{i j}\left(\left\langle z_{i}, z_{i}\right\rangle-\left\langle z_{i}, z_{j}\right\rangle\right)
\end{aligned}
$$

## Quadratic stress (Term 2)

$$
\begin{aligned}
& \begin{aligned}
& \sum_{i>j} d_{i j}(Z) d_{X}\left(x_{i}, x_{j}\right)= \sum_{i, j} a_{i j}\left(\left\langle z_{i}, z_{i}\right\rangle-\left\langle z_{i}, z_{j}\right\rangle\right) \\
&=\operatorname{tr}\left(B Z Z^{\mathrm{T}}\right)=\operatorname{tr}\left(Z^{\mathrm{T}} B Z\right)
\end{aligned} \\
& \text { where } B_{i j}=\left\{\begin{array}{cl}
-a_{i j} & i \neq j \\
-\sum_{k \neq i} B_{i k} & i=j
\end{array}\right.
\end{aligned}
$$

Check:

$$
\begin{aligned}
\left(B Z Z^{\mathrm{T}}\right)_{i i} & =\left(-\sum_{k \neq i} B_{i k}\right)\left\langle z_{i}, z_{i}\right\rangle+\sum_{j \neq i}-a_{i j}\left\langle z_{j}, z_{i}\right\rangle=\left(-\sum_{k \neq i}-a_{i k}\right)\left\langle z_{i}, z_{i}\right\rangle+\sum_{j \neq i}-a_{i j}\left\langle z_{j}, z_{i}\right\rangle \\
& =\sum_{k \neq i} a_{i k}\left\langle z_{i}, z_{i}\right\rangle-\sum_{j \neq i} a_{i j}\left\langle z_{j}, z_{i}\right\rangle=\sum_{j \neq i} a_{i j}\left\langle\left\langle z_{i}, z_{i}\right\rangle-\left\langle z_{j}, z_{i}\right\rangle\right)
\end{aligned}
$$

## Quadratic stress (Term 2)

$$
\begin{aligned}
& \qquad \sum_{i>j} d_{i j}(Z) d_{X}\left(x_{i}, x_{j}\right)=\operatorname{tr}\left(Z^{\mathrm{T}} B Z\right) \\
& \text { where } B_{i j}(Z)=\left\{\begin{array}{cc}
-d_{X}\left(x_{i}, x_{j}\right) d_{i j}^{-1}(Z) & i \neq j, d_{i j}(Z) \neq 0 \\
0 & i \neq j, d_{i j}(Z)=0 \\
-\sum_{k \neq i} B_{i k} & i=j
\end{array}\right.
\end{aligned}
$$

We make explicit the dependence of $B$ on $Z$ by writing $B(Z)$.

## Least-squares MDS

$$
\begin{aligned}
\sigma_{2}(Z) & =\sum_{i>j}\left|d_{X}\left(x_{i}, x_{j}\right)-d_{i j}(Z)\right|^{2} \\
& =\operatorname{tr}\left(Z^{\mathrm{T}} V Z\right)-2 \operatorname{tr}\left(Z^{\mathrm{T}} B(Z) Z\right)+\sum_{i>j} d_{X}^{2}\left(x_{i}, x_{j}\right)
\end{aligned}
$$

Our task is to solve the unconstrained non-convex problem:

$$
\min _{Z \in \mathbf{R}^{N \times m}} \sigma_{2}(Z)
$$

We will use gradient descent.

## Gradient descent

$$
\min f(\mathbf{x})
$$



Allows to find a local minimum of $f$.

Choose starting point $\mathbf{x}^{(0)}$
Iterate: $\mathbf{x}^{(t+1)}=\mathbf{x}^{(t)}-\alpha \nabla f\left(\mathbf{x}^{(t)}\right)$
The recursive equation produces a non-increasing sequence

$$
f\left(\mathbf{x}^{(0)}\right) \geq f\left(\mathbf{x}^{(1)}\right) \geq f\left(\mathbf{x}^{(2)}\right) \cdots
$$

## Gradient of the quadratic stress

$$
\min _{Z \in \mathbf{R}^{N \times m}} \sigma_{2}(Z)
$$

$$
\begin{aligned}
\nabla \sigma_{2}(Z) & =\nabla\left(\operatorname{tr}\left(Z^{\mathrm{T}} V Z\right)-2 \operatorname{tr}\left(Z^{\mathrm{T}} B(Z) Z\right)+\sum_{i>j} d_{X}^{2}\left(x_{i}, x_{j}\right)\right) \\
& =2 V Z-2 B(Z) Z
\end{aligned}
$$

Exercise: Derive the expression given for the gradient $\nabla \sigma_{2}(Z)$

## Gradient descent

$$
\min _{Z \in \mathbf{R}^{N \times m}} \sigma_{2}(Z)
$$

Start with a random configuration of points $Z^{(0)}$

Apply the recursive equations:

$$
Z^{(t+1)}=Z^{(t)}-\alpha \nabla \sigma_{2}\left(Z^{(t)}\right)=Z^{(t)}-2 \alpha\left(V Z^{(t)}-B\left(Z^{(t)}\right) Z^{(t)}\right)
$$

Terminate when $\left|\sigma_{2}\left(Z^{(t+1)}\right)-\sigma_{2}\left(Z^{(t)}\right)\right|<\varepsilon$

## Multidimensional scaling

## Demo Time!



## Suggested reading

- Numerical geometry of non-rigid shapes. Bronstein, Bronstein, Kimmel. Chapters 7.1, 7.2, 7.3, 7.9
- Coulomb shapes: using electrostatic forces for deformation-invariant shape representation. Boscaini et al. 3DOR 2014.
- Cross-collection map inference by intrinsic alignment of shape spaces. Shapira and Ben-Chen. CGF 33(5), 2014.

