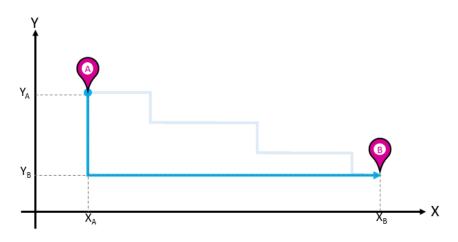
Functions that Preserve Manhattan Distance

Gary Miller, Shyam Narayanan, Mark Sellke

CMU MIT Stanford



Manhattan / Taxicab Distance



Taxicab Distance =
$$|X_A - X_B| + |Y_A - Y_B|$$

A Motivating Puzzle

• Given any n points in any dimension for any n, compute the $\binom{n}{2}$ Manhattan distances between each pair of points.

A Motivating Puzzle

- Given any n points in any dimension for any n, compute the $\binom{n}{2}$ Manhattan distances between each pair of points.
- Raise each distance to the 2/3 power.

A Motivating Puzzle

- Given any n points in any dimension for any n, compute the $\binom{n}{2}$ Manhattan distances between each pair of points.
- Raise each distance to the 2/3 power.
- The result is always a Manhattan distance!

Why?



• Given 3 points on a line at at 0, 1 and 2, compute Manhattan distances between each pair of points.

- Given 3 points on a line at at 0, 1 and 2, compute Manhattan distances between each pair of points.
- Raise each distance to the 2/3 power.

- Given 3 points on a line at at 0, 1 and 2, compute Manhattan distances between each pair of points.
- Raise each distance to the 2/3 power.
- Find 3 points whose Manhattan distances match these powered distances!

Pairwise distances: 1, 1, 2.

Pairwise distances: 1, 1, 2.

Powered distances: 1, 1, 2^{2/3}.

- Pairwise distances: 1, 1, 2.
- Powered distances: 1, 1, 2^{2/3}.

$$a = \left(\frac{1}{2} + \frac{2^{2/3}}{4}, \frac{1}{2} - \frac{2^{2/3}}{4}\right)$$

$$b = (0, 0)$$

$$c = \left(\frac{1}{2} - \frac{2^{2/3}}{4}, \frac{1}{2} + \frac{2^{2/3}}{4}\right).$$

Definition: The function $f(x) = x^{2/3}$ **sends** Manhattan distances to Manhattan distances.

The Core Mystery of our Talk

• Why does $x^{2/3}$ send Manhattan distances to Manhattan distances?

 Our work: A self-contained, 'simpler' proof that Bernstein functions send Manhattan distances to Manhattan distances.

- Our work: A self-contained, 'simpler' proof that Bernstein functions send Manhattan distances to Manhattan distances.
 - This was known before, with a more intricate proof [Schoenberg '37 + Assouad '80]

- Our work: A self-contained, 'simpler' proof that Bernstein functions send Manhattan distances to Manhattan distances.
 - This was known before, with a more intricate proof [Schoenberg '37 + Assouad '80]
- Our Work: Only Benstein Manhattan distances to Manhattan distances.

- Our work: A self-contained, 'simpler' proof that Bernstein functions send Manhattan distances to Manhattan distances.
 - This was known before, with a more intricate proof [Schoenberg '37 + Assouad '80]
- Our Work: Only Benstein Manhattan distances to Manhattan distances.
 - This is new.

- Our work: A self-contained, 'simpler' proof that Bernstein functions send Manhattan distances to Manhattan distances.
 - ► This was known before, with a more intricate proof [Schoenberg '37 + Assouad '80]
- Our Work: Only Benstein Manhattan distances to Manhattan distances.
 - This is new.
- Our talk focuses on the first point.

• Why does $x^{2/3}$ send Manhattan distances to Manhattan distances? A full proof.

- Why does $x^{2/3}$ send Manhattan distances to Manhattan distances? A full proof.
- A deeper understanding of background tools:

- Why does $x^{2/3}$ send Manhattan distances to Manhattan distances? A full proof.
- A deeper understanding of background tools:
 - Manhattan distances

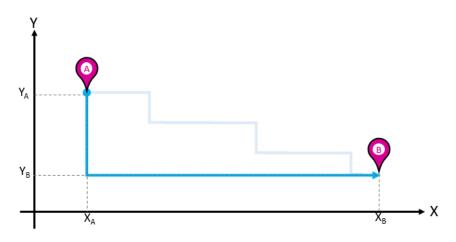
- Why does $x^{2/3}$ send Manhattan distances to Manhattan distances? A full proof.
- A deeper understanding of background tools:
 - Manhattan distances
 - When to determine whether a distance embeds into (squared) and distance.

- Why does $x^{2/3}$ send Manhattan distances to Manhattan distances? A full proof.
- A deeper understanding of background tools:
 - Manhattan distances
 - When to determine whether a distance embeds into (squared) and interpretation (squared).
 Euclidean distance.
 - **3 Bonus**: A hidden use of Representation Theory of the group \mathbb{Z}_2^n .

Background Roadmap

- Manhattan distances
- ② How to determine whether a distance embeds into (squared) Euclidean distance.
- Bonus: A hidden use of group Representation Theory.

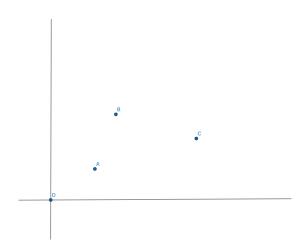
Manhattan / Taxicab Distance

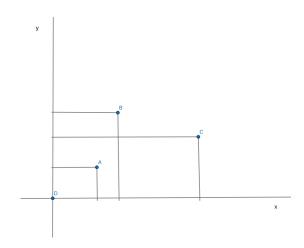


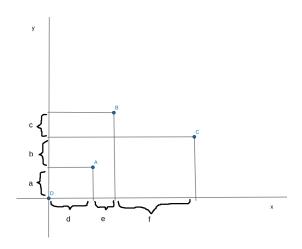
Taxicab Distance =
$$|X_A - X_B| + |Y_A - Y_B|$$

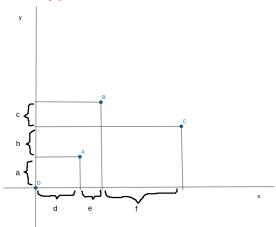
Puzzles on Manhattan distance

 Show that any three point Manhattan metric is equivalent to a Manhattan metric on three corners of some high dimensional hyperbox.

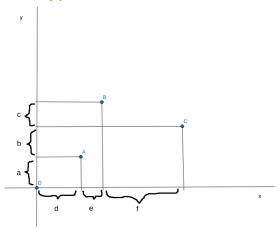




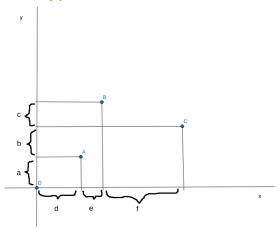




• Let A' be the point (a, 0, 0, d, 0, 0).



- Let A' be the point (a, 0, 0, d, 0, 0).
- Let *B'* be the point (*a*, *b*, *c*, *d*, *e*, 0).



- Let A' be the point (a, 0, 0, d, 0, 0).
- Let *B'* be the point (*a*, *b*, *c*, *d*, *e*, 0).
- Let C' be the point (a, b, 0, d, e, f).

Background Roadmap

- Manhattan distances
- We have to determine whether a distance embeds into (squared) and distance.
- 3 Bonus: A hidden use of group Representation Theory.

What is a criterion for when points embed into **Squared** Euclidean Distance?

• A distance d on a set $\{x_1, \dots x_n\}$ is a **squared** Euclidean distance if and only if the matrix D with:

$$D_{ij}=d(x_i,x_j)$$

What is a criterion for when points embed into **Squared** Euclidean Distance?

• A distance d on a set $\{x_1, \dots x_n\}$ is a **squared** Euclidean distance if and only if the matrix D with:

$$D_{ij}=d(x_i,x_j)$$

• A distance d on a set $\{x_1, \dots x_n\}$ is a **squared** Euclidean distance if and only if the matrix D with:

$$D_{ij} = d(x_i, x_j)$$

has the property that $x^T Dx \le 0$ for all $x \perp 1$ [Schoenberg '35].

• A distance d on a set $\{x_1, \dots x_n\}$ is a **squared** Euclidean distance if and only if the matrix D with:

$$D_{ij}=d(x_i,x_j)$$

has the property that $x^T Dx \le 0$ for all $x \perp 1$ [Schoenberg '35].

A matrix with this property is called a negative type matrix.
 (Classical)

• A distance d on a set $\{x_1, \dots x_n\}$ is a **squared** Euclidean distance if and only if the matrix D with:

$$D_{ij} = d(x_i, x_j)$$

has the property that $x^T Dx \le 0$ for all $x \perp 1$ [Schoenberg '35].

- A matrix with this property is called a negative type matrix.
 (Classical)
- Here, we assume d(x, x) = 0 and d(x, y) = d(y, x). We do not assume metric property.

• A distance d on a set $\{x_1, \dots x_n\}$ is a **squared** Euclidean distance if and only if the matrix D with:

$$D_{ij} = d(x_i, x_j)$$

has the property that $x^T Dx \le 0$ for all $x \perp 1$ [Schoenberg '35].

- A matrix with this property is called a negative type matrix.
 (Classical)
- Here, we assume d(x, x) = 0 and d(x, y) = d(y, x). We do not assume metric property.
- But how can we test for this???

A Sufficient Criterion for Negative Type Matrices

Let D be a symmetric matrix.

A Sufficient Criterion for Negative Type Matrices

- Let D be a symmetric matrix.
- Suppose *D* has eigenvector $\overrightarrow{1}$, and all other eigenvectors have negative eigenvalue.

A Sufficient Criterion for Negative Type Matrices

- Let D be a symmetric matrix.
- Suppose D has eigenvector $\overrightarrow{1}$, and all other eigenvectors have negative eigenvalue.
- Show that *D* is negative type: $x^T Dx \le 0$ for all $x \perp 1$.

Example

• Let $D = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. This has eigenvector $\overrightarrow{1}$, and all other eigenvectors (1, -1) have negative eigenvalue.

Example

- Let $D = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. This has eigenvector $\overrightarrow{1}$, and all other eigenvectors (1, -1) have negative eigenvalue.
- Therefore, it of negative type, and the associated distance is a squared Euclidean distance.

Exercise

• Show that
$$D = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$
 is a negative type matrix!

A Sufficient Criterion for Showing a Distance is Squared Euclidean

Let d be a distance on set $x_1, \ldots x_n$, such that the matrix D with $D_{ii} = d(x_i, x_i)$ satisfies:

- D has $\overrightarrow{1}$ as an eigenvector.
- All other eigenvalues of D are negative.

Then D_{ij} embeds into squared Euclidean distance!

$x^{2/3}$ sends Manhattan distances to Manhattan distances

 Without loss of generality, let's assume the Manhattan distances are the full set of corners on some hyperbox.

$x^{2/3}$ sends Manhattan distances to Manhattan distances

- Without loss of generality, let's assume the Manhattan distances are the full set of corners on some hyperbox.
- We will show that $x^{2/3}$ sends these distances to squared Euclidean distances on a hyperbox.

$x^{2/3}$ sends Manhattan distances to Manhattan distances

- Without loss of generality, let's assume the Manhattan distances are the full set of corners on some hyperbox.
- We will show that $x^{2/3}$ sends these distances to squared Euclidean distances on a hyperbox.
- Squared Euclidean distances on a hyperbox are Manhattan distances!

$x^{2/3}$ sends Manhattan distances to squared Euclidean distances

• We show $x^{2/3}$ sends Manhattan distances on a box to squared Euclidean distance.

$x^{2/3}$ sends Manhattan distances to squared Euclidean distances

- We show $x^{2/3}$ sends Manhattan distances on a box to squared Euclidean distance.
- We do 'proof by example', using a 2 dimensional box.

• Consider a rectangle with side lengths a, b. Let $f(x) = x^{2/3}$.

- Consider a rectangle with side lengths a, b. Let $f(x) = x^{2/3}$.
- Let $x_1 = (0, 0)$, $x_2 = (a, 0)$, $x_3 = (b, 0)$, $x_4 = (a, b)$.

- Consider a rectangle with side lengths a, b. Let $f(x) = x^{2/3}$.
- Let $x_1 = (0, 0), x_2 = (a, 0), x_3 = (b, 0), x_4 = (a, b).$
- The matrix D with $D_{ii} = d(x_i, x_i)$ is:

$$\begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

- Consider a rectangle with side lengths a, b. Let $f(x) = x^{2/3}$.
- Let $x_1 = (0, 0), x_2 = (a, 0), x_3 = (b, 0), x_4 = (a, b).$
- The matrix D with $D_{ii} = d(x_i, x_i)$ is:

$$\begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

• Claim: $\overrightarrow{1}$ is an eigenvector of D. (Why?)

- Consider a rectangle with side lengths a, b. Let $f(x) = x^{2/3}$.
- Let $x_1 = (0,0)$, $x_2 = (a,0)$, $x_3 = (b,0)$, $x_4 = (a,b)$.
- The matrix D with $D_{ij} = d(x_i, x_j)$ is:

$$\begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

- Claim: $\overrightarrow{1}$ is an eigenvector of D. (Why?)
- Claim: All other eigenvalues of *D* are negative or 0. (Why?)

- Consider a rectangle with side lengths a, b. Let $f(x) = x^{2/3}$.
- Let $x_1 = (0, 0), x_2 = (a, 0), x_3 = (b, 0), x_4 = (a, b).$
- The matrix D with $D_{ii} = d(x_i, x_i)$ is:

$$\begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

- Claim: $\overrightarrow{1}$ is an eigenvector of *D*. (Why?)
- Claim: All other eigenvalues of *D* are negative or 0. (Why?)
- Therefore: our Manhattan distances raised to the 2/3 power are squared Euclidean distances.

$$D = \begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

$$f(x) = x^{2/3}$$

• Claim: $\overrightarrow{1}$ is an eigenvector of D with eigenvalue 0 + f(a) + f(b) + f(a + b)

$$D = \begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

$$f(x) = x^{2/3}$$

- Claim: $\overrightarrow{1}$ is an eigenvector of D with eigenvalue 0 + f(a) + f(b) + f(a+b)
- Claim: The other eigenvectors of D are the columns of the 2 by 2 Hadamard matrix:

$$D = \begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

$$f(x) = x^{2/3}$$

- Claim: $\overrightarrow{1}$ is an eigenvector of D with eigenvalue 0 + f(a) + f(b) + f(a + b)
- Claim: The other eigenvectors of D are the columns of the 2 by 2 Hadamard matrix:

Why? You can verify this by calculation.

$$D = \begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

$$f(x) = x^{2/3}$$

- Claim: $\overrightarrow{1}$ is an eigenvector of D with eigenvalue 0 + f(a) + f(b) + f(a + b)
- Claim: The other eigenvectors of D are the columns of the 2 by 2 Hadamard matrix:

- Why? You can verify this by calculation.
- This is the secret use of representation theory: boxes are symmetric under axial reflection, and for any permutation matrix

$$D = \begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

$$f(x) = x^{2/3}$$

- Claim: $\overrightarrow{1}$ is an eigenvector of D with eigenvalue 0 + f(a) + f(b) + f(a + b)
- Claim: The other eigenvectors of D are the columns of the 2 by 2 Hadamard matrix:

- Why? You can verify this by calculation.
- This is the secret use of representation theory: boxes are symmetric under axial reflection, and for any permutation matrix

$$D = \begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

$$f(x) = x^{2/3}$$

• We are knee deep in showing *D* is of negative type. $x^T Dx \le 0$ for all $x \perp 1$.

$$D = \begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

$$f(x) = x^{2/3}$$

- We are knee deep in showing *D* is of negative type. $x^T Dx \le 0$ for all $x \perp 1$.
- We are showing D has $\overrightarrow{1}$ as an eigenvector, and all other eigenvectors are columns of the 2 by 2 Hadamard matrix

$$D = \begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

$$f(x) = x^{2/3}$$

- We are knee deep in showing *D* is of negative type. $x^T Dx \le 0$ for all $x \perp 1$.
- We are showing D has $\overrightarrow{1}$ as an eigenvector, and all other eigenvectors are columns of the 2 by 2 Hadamard matrix

 The associated eigenvalues, we claim, are negative. We will show this by direct computation.

$$D = \begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

satisfies
$$D \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$
.

$$D = \begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

• The eigenvalue λ corresponding to eigenvector v = (1, -1, 1, -1)

satisfies
$$D \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$
.

• This can be computed by evaluating $D_1 v$ where D_1 is the first row of D.

$$D = \begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

satisfies
$$D \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$
.

- This can be computed by evaluating $D_1 v$ where D_1 is the first row of D.
- This is -f(a) + f(b) f(a+b).

$$D = \begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

satisfies
$$D \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$
.

- This can be computed by evaluating $D_1 v$ where D_1 is the first row of D.
- This is -f(a) + f(b) f(a+b).
- $\bullet = -\int_0^a f'(x) dx \int_b^{a+b} f'(x) dx$

$$D = \begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

satisfies
$$D \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$
.

- This can be computed by evaluating $D_1 v$ where D_1 is the first row of D.
- This is -f(a) + f(b) f(a+b).
- $\bullet = -\int_0^a f'(x) dx \int_b^{a+b} f'(x) dx$
- Why is this eigenvalue negative, when $f = x^{2/3}$?

$$D = \begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

• The eigenvalue of D corresponding to $v = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$ equals

$$D_1 v = f(a) + f(b) - f(a+b)$$

$$D = \begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

• The eigenvalue of D corresponding to $v = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$ equals

$$D_1 v = f(a) + f(b) - f(a+b)$$

 $\bullet = \int_0^a \int_0^b f''(x+y) dx dy$

$$D = \begin{pmatrix} 0 & f(a) & f(b) & f(a+b) \\ f(a) & 0 & f(a+b) & f(b) \\ f(b) & f(a+b) & 0 & f(a) \\ f(a+b) & f(b) & f(a) & 0 \end{pmatrix}$$

• The eigenvalue of D corresponding to $v = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$ equals

$$D_1 v = f(a) + f(b) - f(a+b)$$

- $\bullet = \int_0^a \int_0^b f''(x+y) dx dy$
- Why is this eigenvalue negative?

• We showed that our matrix D has eigenvector $\overrightarrow{1}$ and all other eigenvalues negative. This means it is negative type.

- We showed that our matrix D has eigenvector and all other eigenvalues negative. This means it is negative type.
- D being negative type is equivalent to the distance being a squared Euclidean distance. (Presented earlier without proof).

- We showed that our matrix D has eigenvector $\overrightarrow{1}$ and all other eigenvalues negative. This means it is negative type.
- D being negative type is equivalent to the distance being a squared Euclidean distance. (Presented earlier without proof).
- Thus, $x^{2/3}$ sends Manhattan distance to squared Euclidean distance.

- We showed that our matrix D has eigenvector $\overrightarrow{1}$ and all other eigenvalues negative. This means it is negative type.
- D being negative type is equivalent to the distance being a squared Euclidean distance. (Presented earlier without proof).
- Thus, $x^{2/3}$ sends Manhattan distance to squared Euclidean distance.
- What property of $x^{2/3}$ did we use? What is the general property of functions f?

$x^{2/3}$ sends Manhattan-on-a-box to Manhattan

• We showed a weaker result than the one we wanted: $x^{2/3}$ sends Manhattan-on-a-box to squared Euclidean.

$x^{2/3}$ sends Manhattan-on-a-box to Manhattan

- We showed a weaker result than the one we wanted: $x^{2/3}$ sends Manhattan-on-a-box to squared Euclidean.
- But I promised you that $x^{2/3}$ would send Manhattan-on-a-box to squared-Euclidean-on-a-box.

$x^{2/3}$ sends Manhattan-on-a-box to Manhattan

- We showed a weaker result than the one we wanted: $x^{2/3}$ sends Manhattan-on-a-box to squared Euclidean.
- But I promised you that $x^{2/3}$ would send Manhattan-on-a-box to squared-Euclidean-on-a-box.
- We show this by backing out the actual embedding.

• Let's say you have your squared Euclidean distance in a matrix D.

- Let's say you have your squared Euclidean distance in a matrix D.
- You can find the matrix of dot products by computing: $M = -\Pi D\Pi/2$ where Π is the projection matrix off the all ones vector (Well known, exercise for the reader).

- Let's say you have your squared Euclidean distance in a matrix D.
- You can find the matrix of dot products by computing: $M = -\Pi D\Pi/2$ where Π is the projection matrix off the all ones vector (Well known, exercise for the reader).
- Any matrix P with $PP^T = M$, has row vectors which realize the squared Euclidean distance D. (Well known, exercise for the reader).

Let J be the all ones matrix.

- Let J be the all ones matrix.
- Let P be the matrix whose i^{th} row is p_i .

- Let J be the all ones matrix.
- Let P be the matrix whose i^{th} row is p_i .
- Let p_i be points such that $D_{ij} = \|p_i p_j\|_2^2 = |p_i|_2^2 + |p_j|_2^2 2\langle p_i, p_j \rangle$.

- Let J be the all ones matrix.
- Let P be the matrix whose i^{th} row is p_i .
- Let p_i be points such that $D_{ij} = \|p_i p_j\|_2^2 = |p_i|_2^2 + |p_j|_2^2 2\langle p_i, p_j \rangle$.

- Let J be the all ones matrix.
- Let P be the matrix whose i^{th} row is p_i .
- Let p_i be points such that $D_{ij} = \|p_i p_j\|_2^2 = |p_i|_2^2 + |p_j|_2^2 2\langle p_i, p_j \rangle$.

$$D = \begin{pmatrix} |p_1|^2 & 0 & 0 & \dots \\ 0 & |p_2|^2 & 0 & \dots \\ \dots & & & & \\ 0 & 0 & \dots & |p_n|^2 \end{pmatrix} J + J \begin{pmatrix} |p_1|^2 & 0 & 0 & \dots \\ 0 & |p_2|^2 & 0 & \dots \\ \dots & & & & \\ 0 & 0 & \dots & |p_n|^2 \end{pmatrix}$$

$$-2PP^{7}$$

- Let J be the all ones matrix.
- Let P be the matrix whose i^{th} row is p_i .
- Let p_i be points such that $D_{ij} = \|p_i p_j\|_2^2 = |p_i|_2^2 + |p_j|_2^2 2\langle p_i, p_j \rangle$.

$$D = \begin{pmatrix} |p_1|^2 & 0 & 0 & \dots \\ 0 & |p_2|^2 & 0 & \dots \\ \dots & & & & \\ 0 & 0 & \dots & |p_n|^2 \end{pmatrix} J + J \begin{pmatrix} |p_1|^2 & 0 & 0 & \dots \\ 0 & |p_2|^2 & 0 & \dots \\ \dots & & & & \\ 0 & 0 & \dots & |p_n|^2 \end{pmatrix}$$

$$-2PP^{T}$$

• Recall that we defined $M = -\Pi D\Pi/2$ where Π projects off the all ones vector.

- Let J be the all ones matrix.
- Let P be the matrix whose i^{th} row is p_i .
- Let p_i be points such that $D_{ij} = \|p_i p_j\|_2^2 = |p_i|_2^2 + |p_j|_2^2 2\langle p_i, p_j \rangle$.

$$D = \begin{pmatrix} |p_1|^2 & 0 & 0 & \dots \\ 0 & |p_2|^2 & 0 & \dots \\ \dots & & & & \\ 0 & 0 & \dots & |p_n|^2 \end{pmatrix} J + J \begin{pmatrix} |p_1|^2 & 0 & 0 & \dots \\ 0 & |p_2|^2 & 0 & \dots \\ \dots & & & & \\ 0 & 0 & \dots & |p_n|^2 \end{pmatrix}$$

$$-2PP^{T}$$

- Recall that we defined $M = -\Pi D\Pi/2$ where Π projects off the all ones vector.
- Then $M = \Pi P P^T \Pi$

- Let J be the all ones matrix.
- Let P be the matrix whose i^{th} row is p_i .
- Let p_i be points such that $D_{ij} = \|p_i p_j\|_2^2 = |p_i|_2^2 + |p_j|_2^2 2\langle p_i, p_j \rangle$.

$$D = \begin{pmatrix} |p_1|^2 & 0 & 0 & \dots \\ 0 & |p_2|^2 & 0 & \dots \\ \dots & & & & \\ 0 & 0 & \dots & |p_n|^2 \end{pmatrix} J + J \begin{pmatrix} |p_1|^2 & 0 & 0 & \dots \\ 0 & |p_2|^2 & 0 & \dots \\ \dots & & & & \\ 0 & 0 & \dots & |p_n|^2 \end{pmatrix}$$

$$-2PP^{T}$$

- Recall that we defined $M = -\Pi D\Pi/2$ where Π projects off the all ones vector.
- Then $M = \Pi P P^T \Pi$

- Let J be the all ones matrix.
- Let P be the matrix whose i^{th} row is p_i .
- Let p_i be points such that $D_{ij} = \|p_i p_j\|_2^2 = |p_i|_2^2 + |p_j|_2^2 2\langle p_i, p_j \rangle$.

$$D = \begin{pmatrix} |p_1|^2 & 0 & 0 & \dots \\ 0 & |p_2|^2 & 0 & \dots \\ \dots & & & & \\ 0 & 0 & \dots & |p_n|^2 \end{pmatrix} J + J \begin{pmatrix} |p_1|^2 & 0 & 0 & \dots \\ 0 & |p_2|^2 & 0 & \dots \\ \dots & & & & \\ 0 & 0 & \dots & |p_n|^2 \end{pmatrix}$$

$$-2PP^{T}$$

- Recall that we defined $M = -\Pi D\Pi/2$ where Π projects off the all ones vector.
- Then $M = \Pi P P^T \Pi = P P^T$

- Let's say you have your squared Euclidean distance in a matrix D.
- You can find the matrix of dot products by computing: $M = -\Pi D\Pi/2$ where Π is the projection matrix off the all ones vector (Well known, exercise for the reader).
- Any matrix P with $PP^T = M$, has row vectors which realize the squared Euclidean distance D. (Well known, exercise for the reader).

$$h_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, h_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, h_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, h_4 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}.$$

• Let's try this on our matrix D on a 2 dimensional box, after applying $x^{2/3}$ to the distances. The eigenvectors of D are

$$h_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, h_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, h_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, h_4 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}.$$

• We know $D = \lambda_1 h_1 h_1^T + \lambda_2 h_2 h_2^T + \lambda_3 h_3 h_3^T + \lambda_4 h_4 h_4^T$.

$$h_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, h_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, h_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, h_4 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}.$$

- We know $D = \lambda_1 h_1 h_1^T + \lambda_2 h_2 h_2^T + \lambda_3 h_3 h_3^T + \lambda_4 h_4 h_4^T$.
- Here, λ_i is the eigenvalue corresponding to eigenvector h_i .

$$h_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, h_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, h_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, h_4 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}.$$

- We know $D = \lambda_1 h_1 h_1^T + \lambda_2 h_2 h_2^T + \lambda_3 h_3 h_3^T + \lambda_4 h_4 h_4^T$.
- Here, λ_i is the eigenvalue corresponding to eigenvector h_i .
- Then M, the matrix of dot products, equals $-\Pi D\Pi/2$ where Π is the projection matrix off $\overrightarrow{1}$.

$$h_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, h_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, h_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, h_4 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}.$$

- We know $D = \lambda_1 h_1 h_1^T + \lambda_2 h_2 h_2^T + \lambda_3 h_3 h_3^T + \lambda_4 h_4 h_4^T$.
- Here, λ_i is the eigenvalue corresponding to eigenvector h_i .
- Then M, the matrix of dot products, equals $-\Pi D\Pi/2$ where Π is the projection matrix off $\overrightarrow{1}$.
- $M = \frac{1}{2}(-\lambda_2 h_2 h_2^T \lambda_3 h_3 h_3^T \lambda_4 h_4 h_4^T)$ (Why?)

• We know matrix D of squared Euclidean distances satisfies $D = \lambda_1 h_1 h_1^T + \lambda_2 h_2 h_2^T + \lambda_3 h_3 h_3^T + \lambda_4 h_4 h_4^T$.

• Let
$$h_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$
, $h_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$, $h_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$, $h_4 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$.

 $M = \frac{1}{2}(-\lambda_2 h_2 h_2^T - \lambda_3 h_3 h_3^T - \lambda_4 h_4 h_4^T)$

• We know matrix D of squared Euclidean distances satisfies $D = \lambda_1 h_1 h_1^T + \lambda_2 h_2 h_2^T + \lambda_3 h_3 h_3^T + \lambda_4 h_4 h_4^T$.

• Let
$$h_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$
, $h_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$, $h_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$, $h_4 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$.

- $M = \frac{1}{2}(-\lambda_2 h_2 h_2^T \lambda_3 h_3 h_3^T \lambda_4 h_4 h_4^T)$
- Claim: $M = PP^T$, where

• We know matrix D of squared Euclidean distances satisfies $D = \lambda_1 h_1 h_1^T + \lambda_2 h_2 h_2^T + \lambda_3 h_3 h_3^T + \lambda_4 h_4 h_4^T$.

• Let
$$h_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$
, $h_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$, $h_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$, $h_4 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$.

- $M = \frac{1}{2}(-\lambda_2 h_2 h_2^T \lambda_3 h_3 h_3^T \lambda_4 h_4 h_4^T)$
- Claim: $M = PP^T$, where

$$P = \begin{pmatrix} \sqrt{-\lambda_2/2} & \sqrt{-\lambda_3/2} & \sqrt{-\lambda_4/2} \\ -\sqrt{-\lambda_2/2} & \sqrt{-\lambda_3/2} & -\sqrt{-\lambda_4/2} \\ \sqrt{-\lambda_2/2} & -\sqrt{-\lambda_3/2} & -\sqrt{-\lambda_4/2} \\ -\sqrt{-\lambda_2/2} & -\sqrt{-\lambda_3/2} & \sqrt{-\lambda_4/2} \end{pmatrix}$$

• We know matrix D of squared Euclidean distances satisfies $D = \lambda_1 h_1 h_1^T + \lambda_2 h_2 h_2^T + \lambda_3 h_3 h_3^T + \lambda_4 h_4 h_4^T$.

• Let
$$h_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$
, $h_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$, $h_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$, $h_4 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$.

- $M = \frac{1}{2} (-\lambda_2 h_2 h_2^T \lambda_3 h_3 h_3^T \lambda_4 h_4 h_4^T)$
- Claim: $M = PP^T$, where

•
$$P = \begin{pmatrix} \sqrt{-\lambda_2/2} & \sqrt{-\lambda_3/2} & \sqrt{-\lambda_4/2} \\ -\sqrt{-\lambda_2/2} & \sqrt{-\lambda_3/2} & -\sqrt{-\lambda_4/2} \\ \sqrt{-\lambda_2/2} & -\sqrt{-\lambda_3/2} & -\sqrt{-\lambda_4/2} \\ -\sqrt{-\lambda_2/2} & -\sqrt{-\lambda_3/2} & \sqrt{-\lambda_4/2} \end{pmatrix}$$

• The rows of *P* are on the corners of a box!

• We know matrix D of squared Euclidean distances satisfies $D = \lambda_1 h_1 h_1^T + \lambda_2 h_2 h_2^T + \lambda_3 h_3 h_3^T + \lambda_4 h_4 h_4^T$.

• Let
$$h_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$
, $h_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$, $h_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$, $h_4 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$.

- $M = \frac{1}{2} (-\lambda_2 h_2 h_2^T \lambda_3 h_3 h_3^T \lambda_4 h_4 h_4^T)$
- Claim: $M = PP^T$, where

•
$$P = \begin{pmatrix} \sqrt{-\lambda_2/2} & \sqrt{-\lambda_3/2} & \sqrt{-\lambda_4/2} \\ -\sqrt{-\lambda_2/2} & \sqrt{-\lambda_3/2} & -\sqrt{-\lambda_4/2} \\ \sqrt{-\lambda_2/2} & -\sqrt{-\lambda_3/2} & -\sqrt{-\lambda_4/2} \\ -\sqrt{-\lambda_2/2} & -\sqrt{-\lambda_3/2} & \sqrt{-\lambda_4/2} \end{pmatrix}$$

- The rows of *P* are on the corners of a box!
- The box corners are $(\pm\sqrt{-\lambda_2/2},\pm\sqrt{-\lambda_3/2},\pm\sqrt{-\lambda_4/2})$.

Proof Recap!

- Without loss of generality, we assumed the Manhattan distances are the full set of corners on some hyperbox.
- We showed that $x^{2/3}$ sends these distances to squared Euclidean distances on a hyperbox.
- Squared Euclidean distances on a hyperbox are Manhattan distances!
- Therefore: $x^{2/3}$ sends Manhattan distances to Manhattan distances!

Bernstein Functions

Theorem: A function sends Manhattan distances to Manhattan distances if and only if it is **Bernstein.**

Bernstein Functions

Theorem: A function sends Manhattan distances to Manhattan distances if and only if it is **Bernstein**.

A **Bernstein function** is a (positive valued) function with positive derivative, negative second derivative, positive third derivative, negative fourth derivative... on $x \ge 0$.

Examples of Bernstein Functions:

• $f(x) = x^s$ for any $0 \le s \le 1$.

- $f(x) = x^s$ for any $0 \le s \le 1$.
- $f(x) = 1 e^{-tx}$ for any t > 0.

- $f(x) = x^s$ for any $0 \le s \le 1$.
- $f(x) = 1 e^{-tx}$ for any t > 0.
- $f(x) = 1 \frac{1}{1+x}$.

- $f(x) = x^s$ for any $0 \le s \le 1$.
- $f(x) = 1 e^{-tx}$ for any t > 0.
- $f(x) = 1 \frac{1}{1+x}$.
- $f(x) = \log(1 + x)$.

- $f(x) = x^s$ for any $0 \le s \le 1$.
- $f(x) = 1 e^{-tx}$ for any t > 0.
- $f(x) = 1 \frac{1}{1+x}$.
- $f(x) = \log(1 + x)$.
- Any positive linear combination of the above.