

Intro

Votation

B Matri:

ACE

Low Rank Approximation H-score

The Geometry of ACE

Applications
ACE v.s. NN

An Introduction to Alternating Conditional Expectation (ACE) Algorithm

Xiangxiang Xu

Department of Electronic Engineering Tsinghua University

July 15, 2018





- Intro
- Votation
- B Matri
- ACE

Low Rank Approximation H-score

The Geometry of ACE

- 1 Introduction: HGR Correlation
 - 2 Notations
- B Matrix
- 4 The ACE Algorithm
- Low Rank Approximation of BH-score
- 6 The Geometry of ACE
- Applications
 - ACE v.s. NN



Correlation between random variables

Intro

Votatio

B Matri

.

Low Rank

H-score
The Geometry

of ACE

Applications ACE v.s. NN

Pearson correlation coefficient

$$\rho(X,Y) \triangleq \frac{\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}}$$

• $|\rho(X,Y)| \leq 1$

HGR correlation

$$\rho_m(X,Y) \triangleq \max_{f,g} \rho(f(X),g(Y))$$

 $\bullet |\rho(X,Y)| \le \rho_m(X,Y) \le 1$



Intro

Notatio

B Matr

ACE

Low Rank Approximation H-score

The Geometry of ACE

Applications ACE v.s. NN

$\rho_m(X, Y) \triangleq \max_{f, g} \rho(f(X), g(Y)) = \max_{\substack{f : \mathbb{E}[f] = 0, \text{Var}(f) = 1 \\ g : \mathbb{E}[g] = 0, \text{Var}(g) = 1}} \mathbb{E}[f(X)g(Y)]$

Discrete (and finite) alphabet

$$X, Y$$
 take values from $\mathfrak{X} = \{1, \cdots, |\mathfrak{X}|\}$ and $\mathfrak{Y} = \{1, \cdots, |\mathfrak{Y}|\}$, respectively.

$$\begin{split} \mathbb{E}[f(X)g(Y)] &= \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P_{X,Y}(x,y) f(x) g(y) \\ &= \sum_{x=1}^{|\mathcal{X}|} \sum_{y=1}^{|\mathcal{Y}|} f(x) \sqrt{P_X(x)} \cdot \frac{P_{X,Y}(x,y)}{\sqrt{P_X(x)} \sqrt{P_Y(y)}} \cdot g(y) \sqrt{P_Y(y)} \end{split}$$



Intro

Votatio

B Matr

ACE

Low Rank Approximation H-score

The Geometry of ACE

Applications ACE v.s. NN

$$\rho_m(X,Y) \triangleq \max_{f,g} \rho(f(X),g(Y)) = \max_{\substack{f: \mathbb{E}[f]=0, \text{Var}(f)=1\\g: \mathbb{E}[g]=0, \text{Var}(g)=1}} \mathbb{E}[f(X)g(Y)]$$

Discrete (and finite) alphabet

$$X, Y$$
 take values from $\mathfrak{X} = \{1, \cdots, |\mathfrak{X}|\}$ and $\mathfrak{Y} = \{1, \cdots, |\mathfrak{Y}|\}$, respectively.

$$\mathbb{E}[f(X)g(Y)] = \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P_{X,Y}(x,y)f(x)g(y)$$

$$= \sum_{x=1}^{|\mathcal{X}|} \sum_{y=1}^{|\mathcal{Y}|} \underbrace{f(x)\sqrt{P_X(x)}}_{\phi(x)} \cdot \underbrace{\frac{P_{X,Y}(x,y)}{\sqrt{P_X(x)}\sqrt{P_Y(y)}}}_{B(y,x)} \cdot \underbrace{g(y)\sqrt{P_Y(y)}}_{\psi(y)}$$



- Intro
- Notations
- B Matr
- ACE

Low Rank Approximation H-score

The Geometry of ACE

- Introduction: HGR Correlation
- 2 Notations
- B Matrix
- 4 The ACE Algorithm
- Low Rank Approximation of BH-score
- **6** The Geometry of ACE
- Applications
 - ACE v.s. NN



Notations

Notations

- $\phi \leftrightarrow f : \phi(x) \triangleq \sqrt{P_X(x)} f(x)$
- $\psi \leftrightarrow g : \psi(y) \triangleq \sqrt{P_Y(y)}g(y)$

•
$$\boldsymbol{B} \leftrightarrow P_{X,Y}$$
:
 $B(y,x) \triangleq \frac{P_{Y,X}(y,x)}{\sqrt{P_Y(y)}\sqrt{P_X(x)}}$



Intro

Notations

B Matri

ACE

Low Rank

Approximation
H-score

The Geometry of ACE

Applications ACE v.s. NN

Notations

$$\bullet \phi \leftrightarrow f : \phi(x) \triangleq \sqrt{P_X(x)} f(x)$$

•
$$\psi \leftrightarrow g : \psi(y) \triangleq \sqrt{P_Y(y)}g(y)$$

$$\bullet$$
 $\boldsymbol{B} \leftrightarrow P_{X,Y}$:

$$B(y,x) \triangleq \frac{P_{Y,X}(y,x)}{\sqrt{P_Y(y)}\sqrt{P_X(x)}}$$

$$\mathbb{E}[f] = \langle \boldsymbol{\phi}, \boldsymbol{v}_0 \rangle, \ v_0(x) = \sqrt{P_X(x)},$$

$$\mathbb{E}[g] = \langle \boldsymbol{\psi}, \boldsymbol{u}_0 \rangle, \ u_0(y) = \sqrt{P_Y(y)},$$

$$\mathbb{E}[f^2] = \|\boldsymbol{\phi}\|^2, \ \mathbb{E}[g^2] = \|\boldsymbol{\psi}\|^2,$$

$$\mathbb{E}[f(X)g(Y)] = \boldsymbol{\psi}^{\mathrm{T}} \boldsymbol{B} \boldsymbol{\phi}.$$

$$\boldsymbol{B}\boldsymbol{\phi} \leftrightarrow \mathbb{E}[f(X)|Y], \boldsymbol{B}^{\mathrm{T}}\boldsymbol{\psi} \leftrightarrow \mathbb{E}[g(Y)|X]$$



Intro

Notations

B Matri

ACE

Low Rank Approximation H-score

The Geometry of ACE

Applications ACE v.s. NN

Notations

•
$$\phi \leftrightarrow f : \phi(x) \triangleq \sqrt{P_X(x)} f(x)$$

•
$$\psi \leftrightarrow g : \psi(y) \triangleq \sqrt{P_Y(y)}g(y)$$

$$\bullet$$
 $\boldsymbol{B} \leftrightarrow P_{X,Y}$:

$$B(y,x) \triangleq \frac{P_{Y,X}(y,x)}{\sqrt{P_Y(y)}\sqrt{P_X(x)}}$$

$$\mathbb{E}[f] = \langle \boldsymbol{\phi}, \boldsymbol{v}_0 \rangle, \ v_0(x) = \sqrt{P_X(x)},$$

$$\mathbb{E}[g] = \langle \boldsymbol{\psi}, \boldsymbol{u}_0 \rangle, \ u_0(y) = \sqrt{P_Y(y)},$$

$$\mathbb{E}[f^2] = \|\boldsymbol{\phi}\|^2, \ \mathbb{E}[g^2] = \|\boldsymbol{\psi}\|^2,$$

$$\mathbb{E}[f(X)g(Y)] = \boldsymbol{\psi}^{\mathrm{T}} \boldsymbol{B} \boldsymbol{\phi}.$$

$$\boldsymbol{B} \boldsymbol{\phi} \leftrightarrow \mathbb{E}[f(X)|Y], \boldsymbol{B}^{\mathrm{T}} \boldsymbol{\psi} \leftrightarrow \mathbb{E}[g(Y)|X]$$

$$\rho_m(X,Y) = \max_{\substack{f : \mathbb{E}[f] = 0, \text{Var}(f) = 1 \\ g : \mathbb{E}[g] = 0, \text{Var}(g) = 1}} \mathbb{E}[f(X)g(Y)] = \max_{\substack{\phi : \langle \phi, v_0 \rangle = 0, \|\phi\| = 1 \\ \psi : \langle \psi, u_0 \rangle = 0, \|\psi\| = 1}} \psi^{\mathrm{T}} \boldsymbol{B} \phi$$



Intro

Notation

B Matrix

ACE

Low Rank Approximation H-score

The Geometry of ACE

- Introduction: HGR Correlation
- 2 Notations
- B Matrix
- The ACE Algorithm
- 5 Low Rank Approximation of B
 H-score
- 6 The Geometry of ACE
- Applications
 - ACE v.s. NN



Matrix $oldsymbol{B}$

Intro

Notation

B Matrix

ACE.

Low Rank Approximation H-score

The Geometry of ACE

Applications ACE v.s. NN

Definition

$$\boldsymbol{B} \leftrightarrow P_{X,Y}: \quad B(y,x) \triangleq \frac{P_{Y,X}(y,x)}{\sqrt{P_Y(y)}\sqrt{P_X(x)}}$$

Properties

- $\bullet \ m{u}_0 = m{B} m{v}_0, m{v}_0 = m{B}^{\mathrm{T}} m{u}_0.$
- $\|\boldsymbol{B}\|_2 = 1$.
- The SVD of \boldsymbol{B} : $\boldsymbol{B} = \sum_{i=1}^{K} \sigma_{i} \boldsymbol{u}_{i} \boldsymbol{v}_{i}^{\mathrm{T}}, \, \sigma_{0} = 1 \geq \sigma_{1} \geq \cdots \geq \sigma_{K}.$



$\mathsf{Matrix}\ \boldsymbol{B}$

Intro

Notation

B Matrix

ACE

Low Rank Approximation H-score

The Geometry of ACE

Applications ACE v.s. NN

Definition

$$\boldsymbol{B} \leftrightarrow P_{X,Y}: \quad B(y,x) \triangleq \frac{P_{Y,X}(y,x)}{\sqrt{P_Y(y)}\sqrt{P_X(x)}}$$

Properties

• The SVD of
$$\boldsymbol{B}$$
: $\boldsymbol{B} = \sum_{i=0}^K \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\mathrm{T}}, \, \sigma_0 = 1 \geq \sigma_1 \geq \cdots \geq \sigma_K.$

$$\bullet \ \sigma_1 = \max_{\substack{\boldsymbol{\phi}: \langle \boldsymbol{\phi}, \boldsymbol{v}_0 \rangle = 0, \|\boldsymbol{\phi}\| = 1 \\ \boldsymbol{\psi}: \langle \boldsymbol{\psi}, \boldsymbol{u}_0 \rangle = 0, \|\boldsymbol{\psi}\| = 1 }} \boldsymbol{\psi}^{\mathrm{T}} \boldsymbol{B} \boldsymbol{\phi}, \ \text{with} \ \boldsymbol{\phi}^* = \boldsymbol{v}_1, \boldsymbol{\psi}^* = \boldsymbol{u}_1.$$



Matrix $ilde{m{B}}$

Intro

Votatior

B Matrix

ACE

Low Rank Approximation H-score

The Geometry of ACE

Applications ACE v.s. NN

Definition

$$ilde{m{B}} riangleq m{B} - m{u}_0 m{v}_0^{\mathrm{T}}$$

$$\max_{\substack{\boldsymbol{\phi}:\,\langle\boldsymbol{\phi},\boldsymbol{v}_0\rangle=0,\|\boldsymbol{\phi}\|=1\\\boldsymbol{\psi}:\,\langle\boldsymbol{\psi},\boldsymbol{u}_0\rangle=0,\|\boldsymbol{\psi}\|=1}}\boldsymbol{\psi}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{\phi}=\max_{\substack{\boldsymbol{\phi},\boldsymbol{\psi}:\,\|\boldsymbol{\phi}\|=\|\boldsymbol{\psi}\|=1\\\boldsymbol{\psi}:\,\langle\boldsymbol{\psi},\boldsymbol{u}_0\rangle=0,\|\boldsymbol{\psi}\|=1}}\boldsymbol{\psi}^{\mathrm{T}}\tilde{\boldsymbol{B}}\boldsymbol{\phi}$$



- Intro
- Notation
- B Matri

ACE

- Low Rank Approximation H-score
- The Geometry of ACE
- Applications ACE v.s. NN

- Introduction: HGR Correlation
- 2 Notations
- B Matrix
- 4 The ACE Algorithm
- 5 Low Rank Approximation of B
 - H-score
- 6 The Geometry of ACE
 - Applications
 - ACE v.s. NN



Power method and ACE(Alternating Conditional Expectation)

Intro

lotation

B Matri

ACE

Low Rank Approximation H-score

The Geometry of ACE

Applications ACE v.s. NN

Power method

```
egin{aligned} \phi \leftarrow \phi_0 \; ; & /st \; 	ext{Initialization} \; st / \ \phi \leftarrow \phi - \langle \phi, v_0 
angle v_0 \; ; & /st \; \phi \perp v_0 \; st / \ 	ext{repeat} \ & \psi \leftarrow B \phi; \ & \phi \leftarrow B^{	ext{T}} \psi; \ & \phi \leftarrow \phi / \| \phi \|; \end{aligned}
```

until $\psi^{\mathrm{T}} B \phi$ stops to increase:



Power method and ACE(Alternating Conditional Expectation)

Intro

lotation

B Matr

ACE

Low Rank Approximation H-score

The Geometry of ACE

Applications
ACE v.s. NN

Power method

$$egin{aligned} \phi \leftarrow \phi_0 \; ; & /st \; ext{Initialization} \; st/ \ \phi \leftarrow \phi - \langle \phi, v_0
angle v_0 \; ; & /st \; \phi \perp v_0 \; st/ \ ext{repeat} \end{aligned}$$

$$egin{aligned} oldsymbol{\psi} &\leftarrow oldsymbol{B} oldsymbol{\phi}; \ oldsymbol{\phi} &\leftarrow oldsymbol{\phi} / \|oldsymbol{\phi}\| \colon \end{aligned}$$

until $\psi^{\mathrm{T}} B \phi$ stops to increase;

ACE

$$\begin{array}{ll} g \leftarrow g_0 \; ; & / * \; \text{Initialization} \; */ \\ g(y) \leftarrow g(y) - \mathbb{E}[g(Y)] \; ; \, / * \; \text{Center} \; */ \\ \mathbf{repeat} \\ & \mid \; f(X) \leftarrow \mathbb{E}[g(Y)|X]; \end{array}$$

$$| f(X) \leftarrow \mathbb{E}[g(Y)|X];$$

$$| g(Y) \leftarrow \mathbb{E}[f(X)|Y];$$

$$| g(Y) \leftarrow g(Y)/\sqrt{\mathbb{E}[g^2(Y)]};$$

$$| \mathbf{g}[f(X)g(Y)]| | stops to increase$$

until $\mathbb{E}[f(X)g(Y)]$ stops to increase;



ACE for multiple modes

Intro

Votation

B Matri

ACE

Low Rank Approximatio

The Geometry of ACE

Applications ACE v.s. NN ${\bf Orthogonalization}$



- Intro
- Notation
- B Matrix
- ACE

Low Rank Approximation H-score

The Geometry of ACE

- Introduction: HGR Correlation
- 2 Notations
- 3 B Matrix
- 4 The ACE Algorithm
- **6** Low Rank Approximation of B
 - H-score
- 6 The Geometry of ACE
 - Applications
 - ACE v.s. NN



Eckart-Young-Mirsky theorem

Suppose $A = U\Sigma V^{T}$, then $A_r = U_r\Sigma_rV_r^{T}$ is the optimal solution to the following low rank approximation problem:

$$\label{eq:local_equation} \begin{split} & \underset{\hat{A}}{\text{minimize}} & & \| A - \hat{A} \|_{\text{F}}^2 \\ & \text{subject to} & & \text{rank}(\hat{A}) \leq r. \end{split}$$

Corollary

The solutions of

$$\underset{(\boldsymbol{\phi},\boldsymbol{\psi})}{\text{minimize}} \|\tilde{\boldsymbol{B}} - \boldsymbol{\psi} \boldsymbol{\phi}^{\mathrm{T}}\|_{\mathrm{F}}^{2}$$

are
$$\{(c\mathbf{v}_1, \frac{\sigma_1}{c}\mathbf{u}_1) : c \neq 0\}.$$

Notation

B Matri

Low Rank

Approximation
H-score

of ACE



Corollary

The solutions of

$$\underset{(\boldsymbol{\phi},\boldsymbol{\psi}):\boldsymbol{\phi}\perp\boldsymbol{v}_0,\boldsymbol{\psi}\perp\boldsymbol{u}_0}{\text{minimize}} \|\boldsymbol{B}-\boldsymbol{\psi}\boldsymbol{\phi}^{\mathrm{T}}\|_{\mathrm{F}}^2$$

are $\{(c\mathbf{v}_1, \frac{\sigma_1}{c}\mathbf{u}_1) : c \neq 0\}.$

NI.

IVOCACION

B Mati

ACE

Low Rank Approximation H-score

The Geometry of ACE



Corollary

The solutions of

$$\underset{(\boldsymbol{\phi},\boldsymbol{\psi}):\boldsymbol{\phi}\perp\boldsymbol{v}_0,\boldsymbol{\psi}\perp\boldsymbol{u}_0}{\text{minimize}} \|\boldsymbol{B}-\boldsymbol{\psi}\boldsymbol{\phi}^{\mathrm{T}}\|_{\mathrm{F}}^2$$

are $\{(c\mathbf{v}_1, \frac{\sigma_1}{c}\mathbf{u}_1) : c \neq 0\}.$

are
$$\{(c\mathbf{v}_1, \frac{\sigma_1}{c}\mathbf{u}_1) : c \neq 0\}.$$

$$rac{1}{2}\|oldsymbol{B} - oldsymbol{\psi}oldsymbol{\phi}^{ ext{T}}\|_{ ext{F}}^2 = rac{1}{2}\|oldsymbol{B}\|_{ ext{F}}^2 - \left[oldsymbol{\psi}^{ ext{T}}oldsymbol{B}oldsymbol{\phi} - rac{1}{2}\cdot(oldsymbol{\phi}^{ ext{T}}oldsymbol{\phi})\cdot(oldsymbol{\psi}^{ ext{T}}oldsymbol{\psi})
ight]$$

Low Rank Approximation



Low Rank Approximation

SVD & low rank approximation

Corollary

The solutions of

$$\underset{(\boldsymbol{\phi},\boldsymbol{\psi}):\boldsymbol{\phi}\perp\boldsymbol{v}_0,\boldsymbol{\psi}\perp\boldsymbol{u}_0}{\text{minimize}} \|\boldsymbol{B}-\boldsymbol{\psi}\boldsymbol{\phi}^{\mathrm{T}}\|_{\mathrm{F}}^2$$

are $\{(c\mathbf{v}_1, \frac{\sigma_1}{c}\mathbf{u}_1) : c \neq 0\}.$

are
$$\{(c\mathbf{v}_1, \frac{\sigma_1}{c}\mathbf{u}_1) : c \neq 0\}.$$

$$\frac{1}{2}\|\boldsymbol{B} - \boldsymbol{\psi}\boldsymbol{\phi}^{\mathrm{T}}\|_{\mathrm{F}}^{2} = \frac{1}{2}\|\boldsymbol{B}\|_{\mathrm{F}}^{2} - \underbrace{\left[\boldsymbol{\psi}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{\phi} - \frac{1}{2}\cdot(\boldsymbol{\phi}^{\mathrm{T}}\boldsymbol{\phi})\cdot(\boldsymbol{\psi}^{\mathrm{T}}\boldsymbol{\psi})\right]}_{\boldsymbol{H}(\boldsymbol{\phi},\boldsymbol{\psi})}$$



Corollary

The solutions of

$$\min_{(oldsymbol{\phi},oldsymbol{\psi}):oldsymbol{\phi}\perpoldsymbol{v}_0,oldsymbol{\psi}\perpoldsymbol{u}_0} \|oldsymbol{B}-oldsymbol{\psi}oldsymbol{\phi}^{ ext{T}}\|_{ ext{F}}^2$$

are $\{(c\mathbf{v}_1, \frac{\sigma_1}{c}\mathbf{u}_1) : c \neq 0\}.$

$$rac{1}{2}\|oldsymbol{B} - oldsymbol{\psi}oldsymbol{\phi}^{ ext{T}}\|_{ ext{F}}^2 = rac{1}{2}\|oldsymbol{B}\|_{ ext{F}}^2 - \left[oldsymbol{\psi}^{ ext{T}}oldsymbol{B}oldsymbol{\phi} - rac{1}{2}\cdot(oldsymbol{\phi}^{ ext{T}}oldsymbol{\phi})\cdot(oldsymbol{\psi}^{ ext{T}}oldsymbol{\psi})
ight]$$

$$H(oldsymbol{\phi},oldsymbol{\psi})$$

$$H(f(X), g(Y)) \triangleq \mathbb{E}[f(X)g(Y)] - \frac{1}{2} \cdot \text{Var}(f(X)) \cdot \text{Var}(g(Y)).$$

Low Rank Approximation

4 D > 4 A > 4 B > 4 B > B = 900



H-score

Intro

Notation

B Matri

ACE

Low Rank Approximatio H-score

The Geometry of ACE

$$H(f(X), g(Y)) \triangleq \mathbb{E}[f(X)g(Y)] - \frac{1}{2} \cdot \text{Var}(f(X)) \cdot \text{Var}(g(Y)).$$

$$H(f(X)) \triangleq \max_{g} H(f(X), g(Y)) = \frac{1}{2} \cdot \frac{\text{Var}(\mathbb{E}[f(X)|Y]])}{\text{Var}(\mathbb{E}[f(X)])}.$$



Intro

Notation

B Matr

ACE

Low Rank Approximation H-score

The Geometry of ACE

- Introduction: HGR Correlation
- 2 Notations
- 3 B Matrix
- 4 The ACE Algorithm
- Low Rank Approximation of BH-score
- 6 The Geometry of ACE
- Applications
 - ACE v.s. NN



The geometry of ACE

Intro

Notation

B Matri

ACE

Low Rank Approximation H-score

The Geometry of ACE

Applications ACE v.s. NN Principal Angles



Intro

Notation:

B Matri

ACE

Low Rank Approximation H-score

The Geometry of ACE

- Introduction: HGR Correlation
- 2 Notations
- 3 B Matrix
- 4 The ACE Algorithm
- 5 Low Rank Approximation of B
 - H-score
- 6 The Geometry of ACE
- Applications
 - ACE v.s. NN

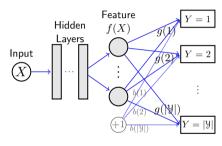


Theoretical results

ACE v.s. NN

H-score v.s. Log-loss

- \bullet Local assumption: X and Y are nearly independent.
- $\ell_{\log} \approx -\frac{1}{2}H(\boldsymbol{f}(X), \boldsymbol{g}(Y)) + C$



$$\boxed{ \tilde{P}_{Y|X}(y|x) \triangleq \frac{e^{f^{\mathrm{T}}(x)g(y) + b(y)}}{\sum_{y' \in \mathbb{Y}} e^{f^{\mathrm{T}}(x)g(y') + b(y')}}} }$$
 Softmax Output



Applications

Intro

lotation

B Matri

ACE

Low Rank Approximation H-score

The Geometry of ACE

Applications ACE v.s. NN

NN for ACE

H-score & ACE-Net:

Warning

Outliers in ACE

Extensions

MACE: Modified(Multiple) ACE Algorithm



References I

Intro

Notatio

B Matr

ACI

Low Rank Approximation H-score

The Geometry of ACE

Applications
ACE v.s. NN

Leo Breiman and Jerome H Friedman.
 Estimating optimal transformations for multiple regression and correlation.
 Journal of the American statistical Association, 80(391):580-598, 1985.

- [2] Hans Gebelein.
 - Das statistische problem der korrelation als variations-und eigenwertproblem und sein zusammenhang mit der ausgleichsrechnung.
 - ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik, 21(6):364–379, 1941.
- [3] Hermann O Hirschfeld.
 - A connection between correlation and contingency.
 - In Mathematical Proceedings of the Cambridge Philosophical Society, volume 31, pages 520–524. Cambridge University Press, 1935.
- [4] Shao-Lun Huang, Anuran Makur, Lizhong Zheng, and Gregory W Wornell. An information-theoretic approach to universal feature selection in high-dimensional inference. In Information Theory (ISIT), 2017 IEEE International Symposium on, pages 1336–1340. IEEE, 2017.



References II

Intro

Notation

B Matr

ACE

Low Rank Approximation H-score

The Geometry of ACE

Applications ACE v.s. NN

- [5] Shao-Lun Huang, Lin Zhang, and Lizhong Zheng. An information-theoretic approach to unsupervised feature selection for high-dimensional data. In *Information Theory Workshop (ITW)*, 2017 IEEE, pages 434–438. IEEE, 2017.
- 6] Anuran Makur, Fabián Kozynski, Shao-Lun Huang, and Lizhong Zheng. An efficient algorithm for information decomposition and extraction. In Communication, Control, and Computing (Allerton), 2015 53rd Annual Allerton Conference on, pages 972–979. IEEE, 2015.
- [7] Alfréd Rényi.

On measures of dependence.

Acta mathematica hungarica, 10(3-4):441-451, 1959.



Intro

Notation

B Matri:

ACE

Low Rank Approximation

The Geometry of ACE

Applications
ACE v.s. NN

Thank you!