

反復重みつけによる ロバストなマッチング相関分析

大阪大学 大学院 基礎工学研究科

奥野彰文 下平英寿

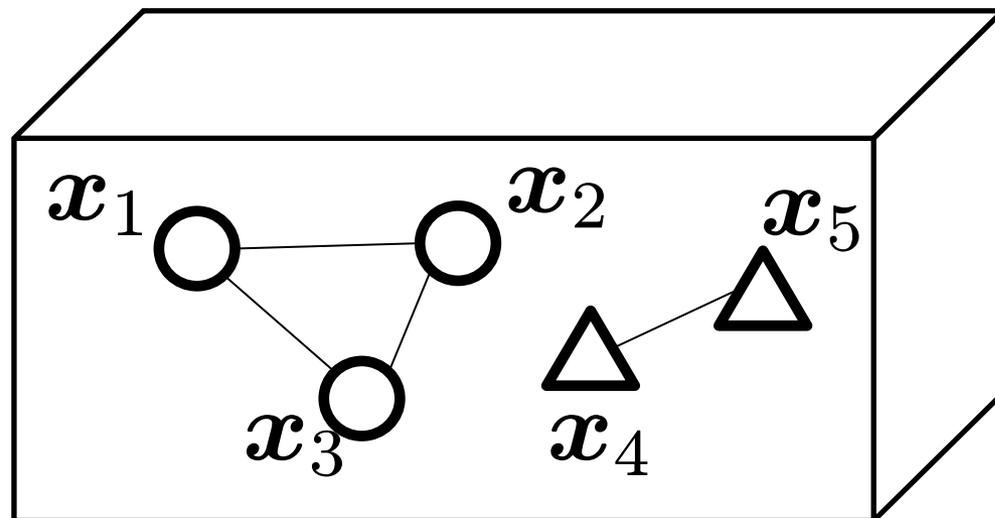
マッチング相関分析

(a.k.a. Spectral Graph Embedding (Yan et al. (2007)[2]))

既知:

データ行列 X

隣接行列 W



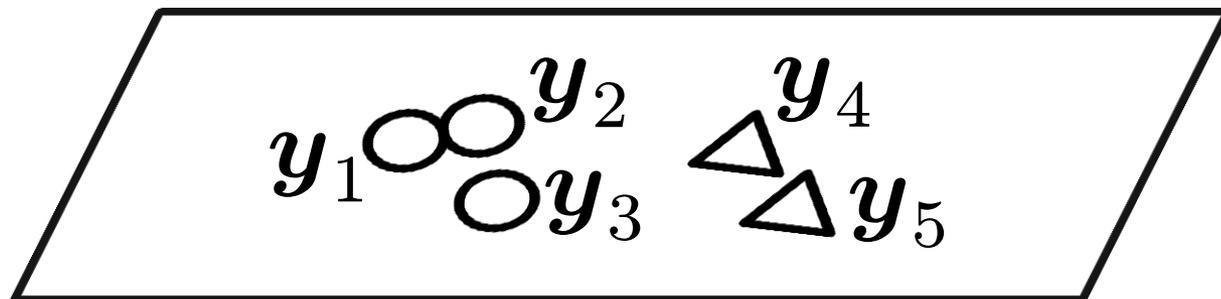
\mathbb{R}^p



$$y_i = A^\top x_i$$

推定:

射影子 A



\mathbb{R}^K

損失関数

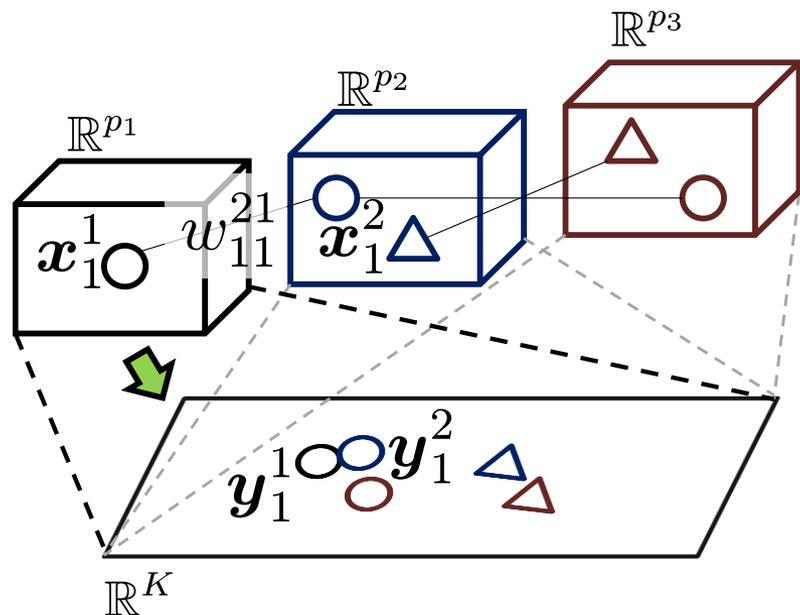
$$\begin{aligned} L_0(\mathbf{A}; \mathbf{X}, \mathbf{W}) &\propto \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_{ij} \|\mathbf{A}^\top \mathbf{x}_i - \mathbf{A}^\top \mathbf{x}_j\|_2^2 \\ &= \text{tr} \mathbf{A}^\top \mathbf{X}^\top (\mathbf{W} - \mathbf{M}) \mathbf{X} \mathbf{A} \end{aligned}$$

制約条件(自明な解を取り除く)

$$\mathbf{A}^\top \mathbf{M} \mathbf{A} = \mathbf{I}_K$$

固有値分解によって解ける

多ドメインマッチング相関分析, Shimodaira(2016)[3]



ドメイン数 $D \in \mathbb{N}$

データベクトル $\mathbf{x}_i^d \in \mathbb{R}^{p_d}$

マッチングウェイト $w_{ij}^{de} \geq 0$

推定

$$\mathbf{A}^d : \mathbf{x}_i^d \mapsto (\mathbf{A}^d)^\top \mathbf{x}_i^d =: \mathbf{y}_i^d$$

\cap

\mathbb{R}^{p_d}

(ドメイン d)

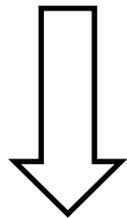
\cap

\mathbb{R}^K

(共通空間)

$$\{\mathbf{X}^d\}, \{\mathbf{W}^{de}\}$$

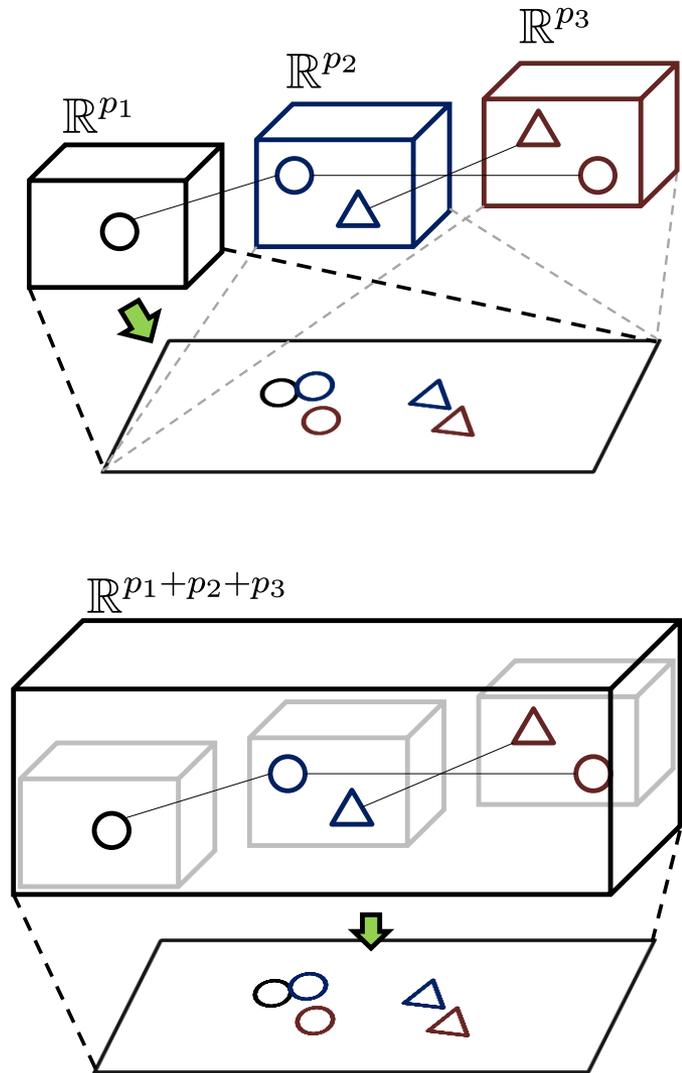
$(d, e \in \{1, \dots, D\})$



Simple Coding

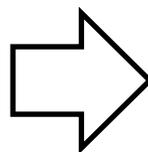
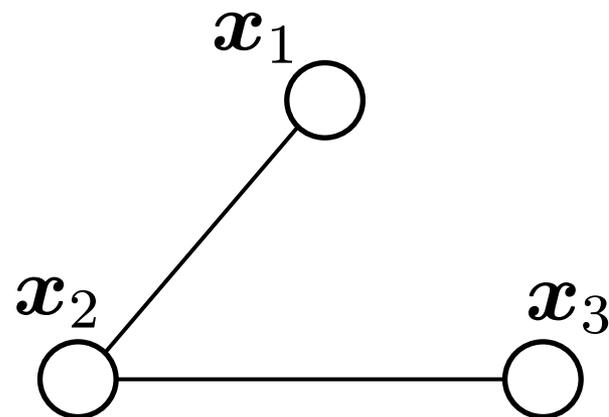
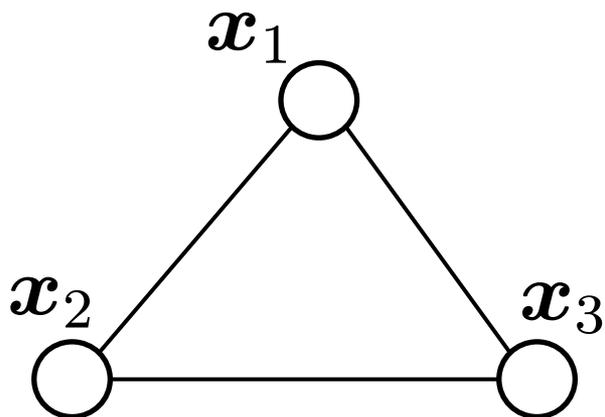
\mathbf{X}, \mathbf{W}

1ドメインの場合のみ
考えれば良い



※同様の定式化がNori et al. (2012)[4]によって独立に提案されている

マッチングウェイトのリサンプリング



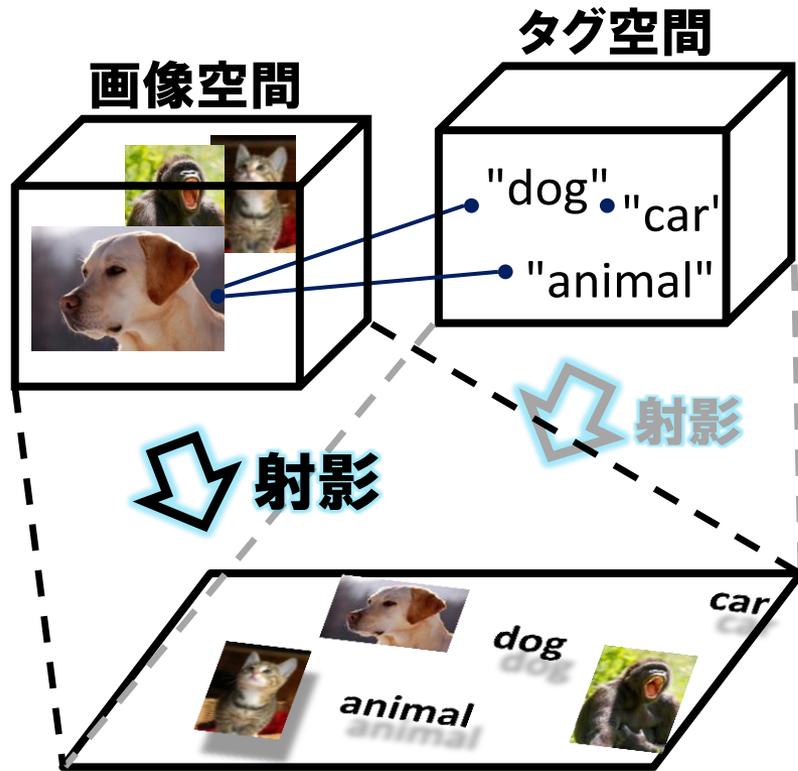
$$\bar{W} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

$$W_0 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

辺のリサンプリングに相当

問題設定

マッチングウェイト

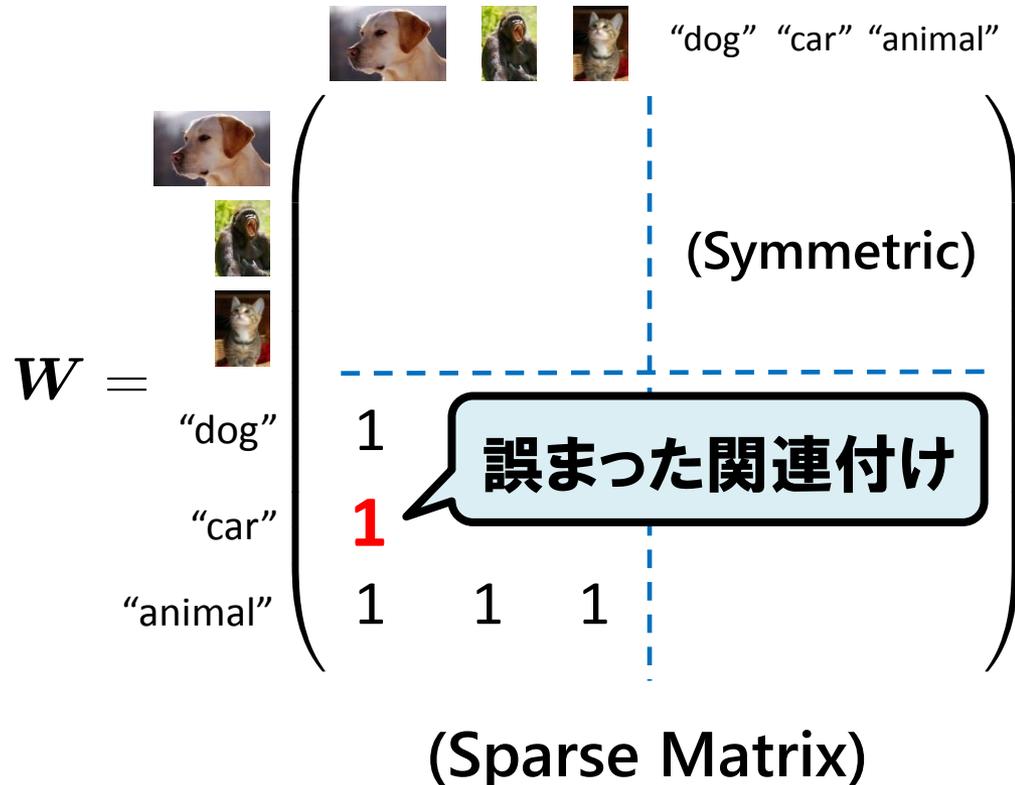
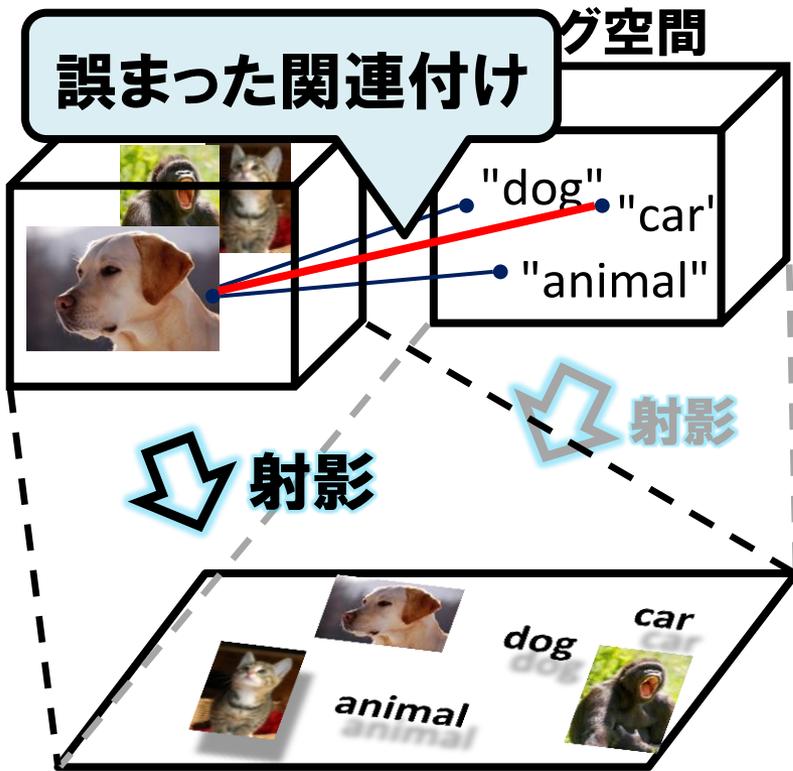


$$W = \begin{matrix} \begin{matrix} \text{dog} & \text{car} & \text{animal} \end{matrix} \\ \begin{matrix} \text{dog} \\ \text{car} \\ \text{animal} \end{matrix} \end{matrix} \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & 1 \end{pmatrix}$$

(Symmetric)

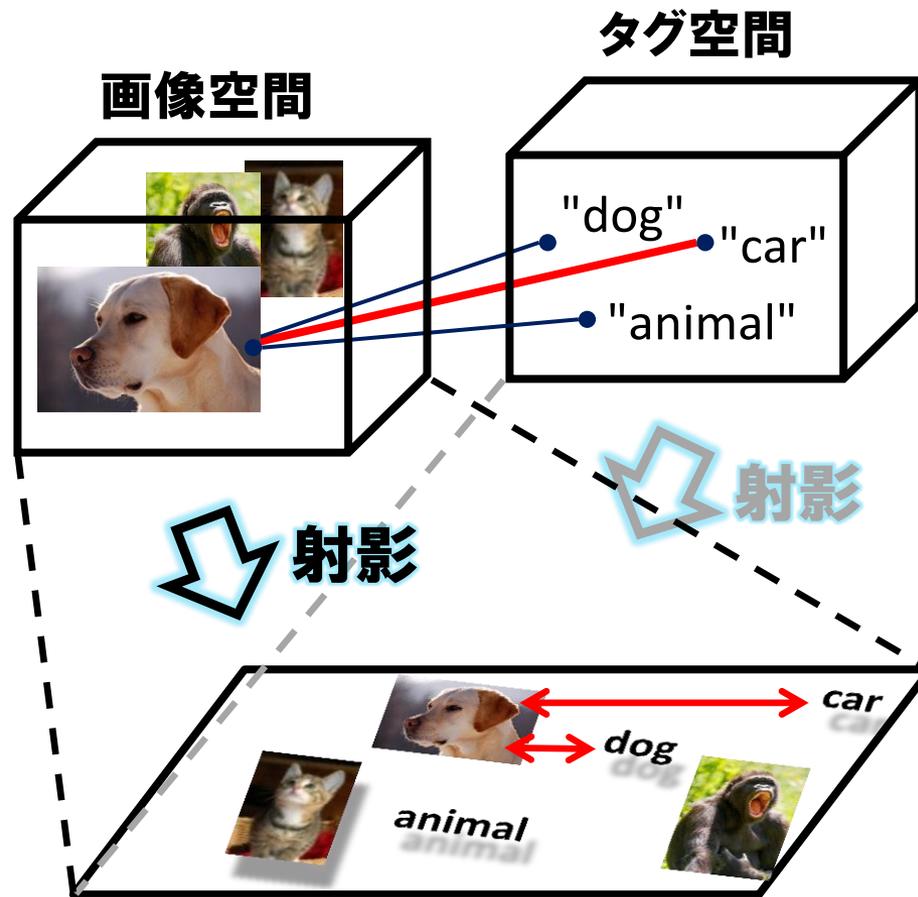
(Sparse Matrix)

マッチングウェイト



誤った関連付けの影響を軽減したい

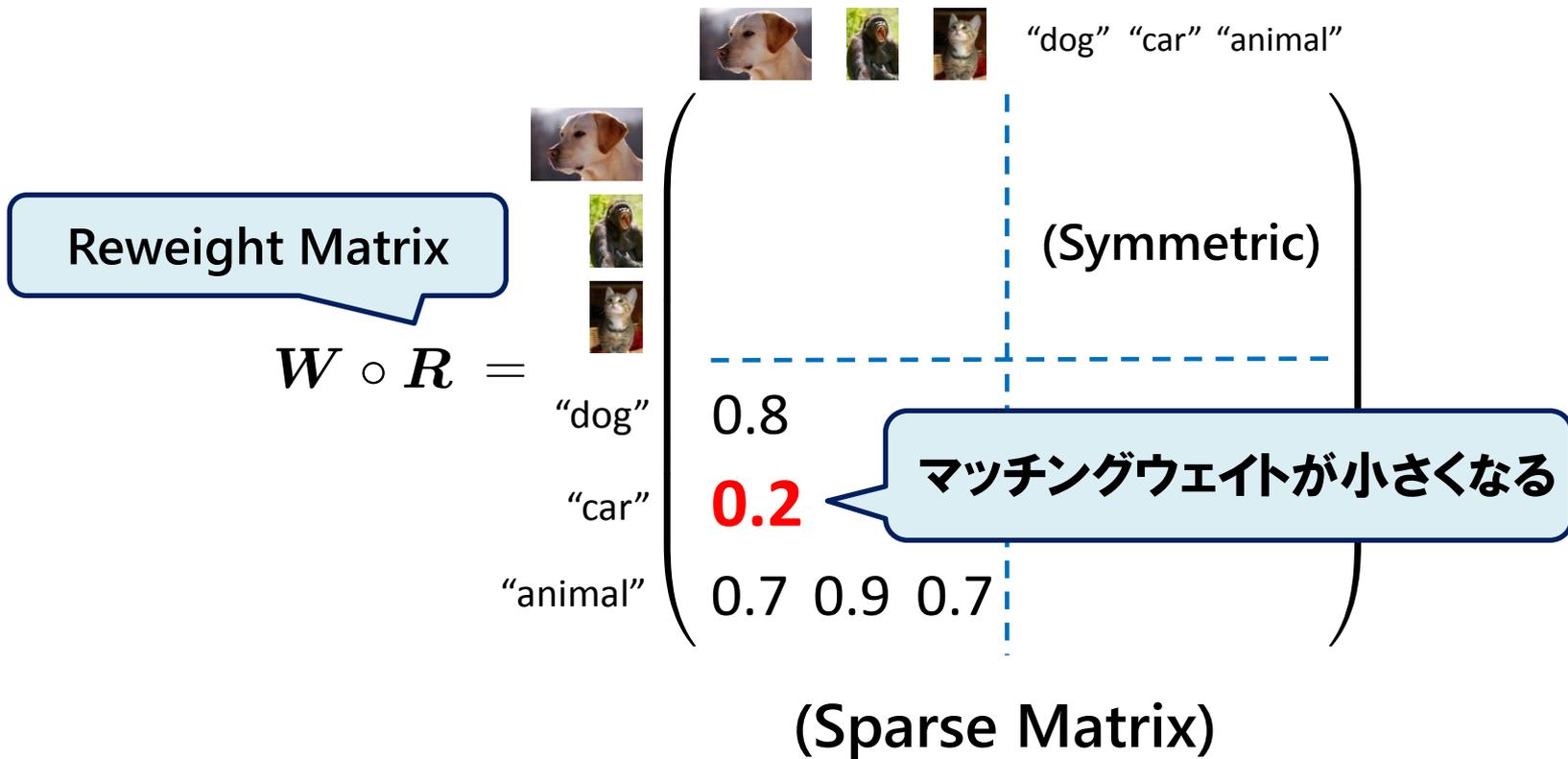
提案手法



マッチングウェイトを重み付け

$$w^{(1)} = w^{(0)} \cdot \exp \left(-\gamma \left\| \mathcal{H} \left(\begin{array}{c} \text{dog image} \end{array} \right) - \mathcal{H} \left(\begin{array}{c} \text{"car"} \end{array} \right) \right\|_2^2 \right)$$

射影の「距離」



マッチングウェイトを重み付け

$$w^{(1)} = w^{(0)} \cdot \exp \left(-\gamma \underbrace{\| \mathcal{H}(\text{dog image}) - \mathcal{H}(\text{"car"}) \|_2^2}_{\text{射影の「距離」}} \right)$$

提案手法: 反復重み付きマッチング相関分析

• 初期値

$$\hat{A}^{(0)} = \arg \min_{A: V[XA]=I} L_0(A; X, W)$$

• 重みの計算

$$R(A) := (r_{ij}(A)), \quad r_{ij}(A) = \exp(-\gamma \|Ax_i - Ax_j\|_2^2)$$

• 射影子の更新

$$\hat{A}^{(t+1)} = \arg \min_{A: V[XA]=I} L_0(A; X, \underline{W \circ R(A^{(t)})})$$

提案手法は以下の損失を単調減少させる

γ -損失関数

$$L_\gamma(\mathbf{A}; \mathbf{X}, \mathbf{W}) = -\frac{1}{\gamma} \log \sum_{i=1}^n \sum_{j=1}^n \tilde{w}_{ij} \exp(-\gamma \|\mathbf{A}^\top \mathbf{x}_i - \mathbf{A}^\top \mathbf{x}_j\|_2^2)$$
$$\left(\tilde{w}_{ij} = w_{ij} / \sum_{i=1}^n \sum_{j=1}^n w_{ij} \right)$$

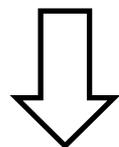
提案手法は γ -損失関数を単調に減少させる
MM-アルゴリズム[7]であることを示した。

$$L_\gamma(\hat{\mathbf{A}}^{(t)}; \mathbf{X}, \mathbf{W}) \geq L_\gamma(\hat{\mathbf{A}}^{(t+1)}; \mathbf{X}, \mathbf{W})$$

$$\hat{\mathbf{A}}_\gamma := \arg \min_{\mathbf{A}: \mathbf{V}[\mathbf{X}\mathbf{A}] = \mathbf{I}} L_\gamma(\mathbf{A}; \mathbf{X}, \mathbf{W})$$

既存手法: γ -相互エントロピー (Robust)[10]

$$d_\gamma(Q, P) = -\frac{1}{\gamma} \log \int p(\mathbf{x})^\gamma dQ(\mathbf{x}) + \frac{1}{1+\gamma} \log \int p(\mathbf{x})^{1+\gamma} d\mathbf{x} \quad (\gamma > 0)$$

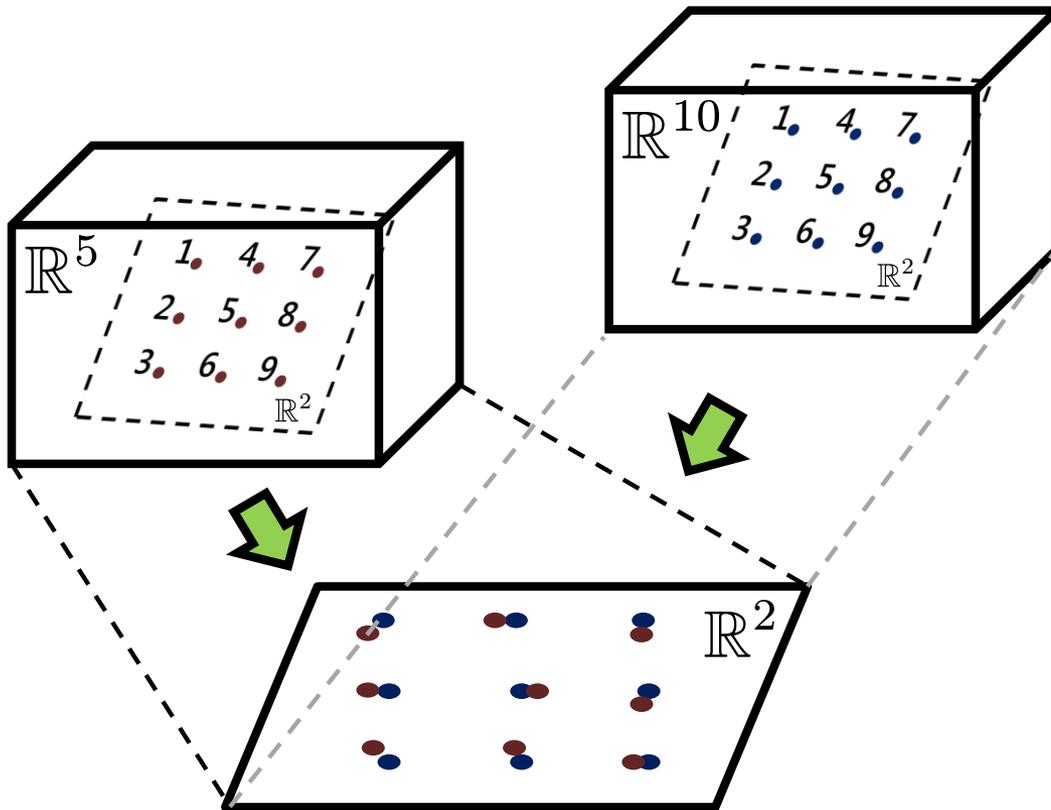


分散既知の正規分布を仮定

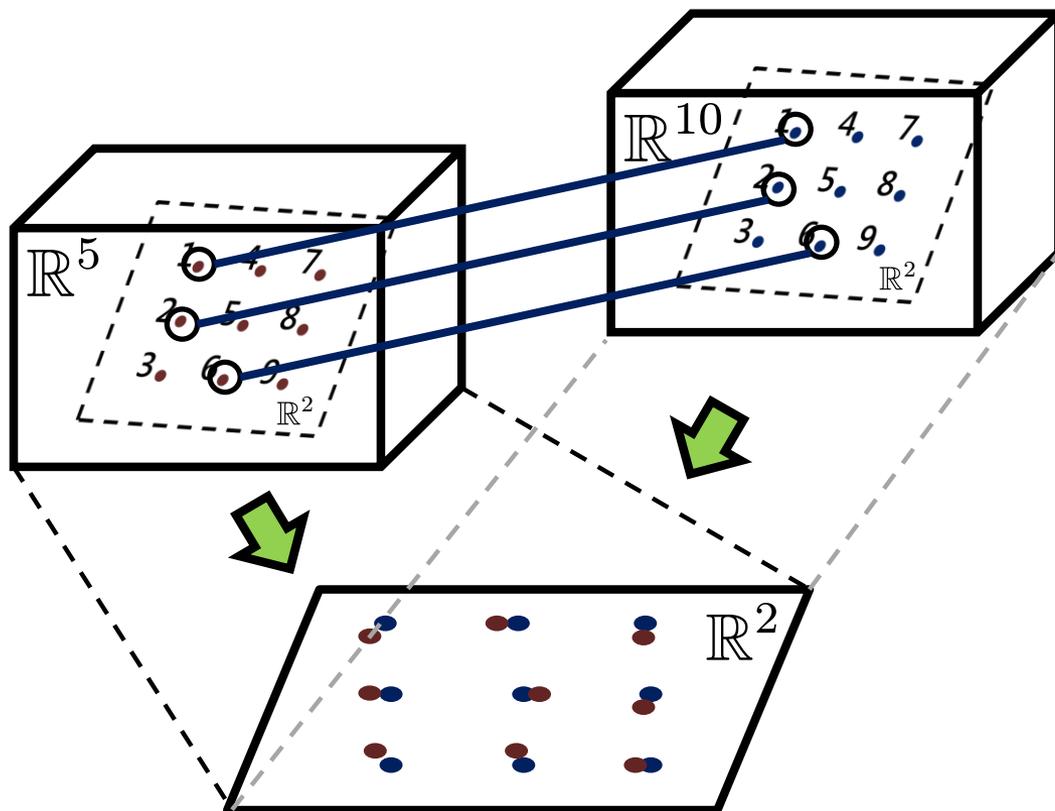
$$d_\gamma(\hat{Q}, P) = -\frac{1}{\gamma} \log \frac{1}{n} \sum_{i=1}^n \exp(-\gamma \|\mathbf{x}_i - \boldsymbol{\mu}\|_2^2) + \text{Const.}$$

数值実験

数値実験の設定



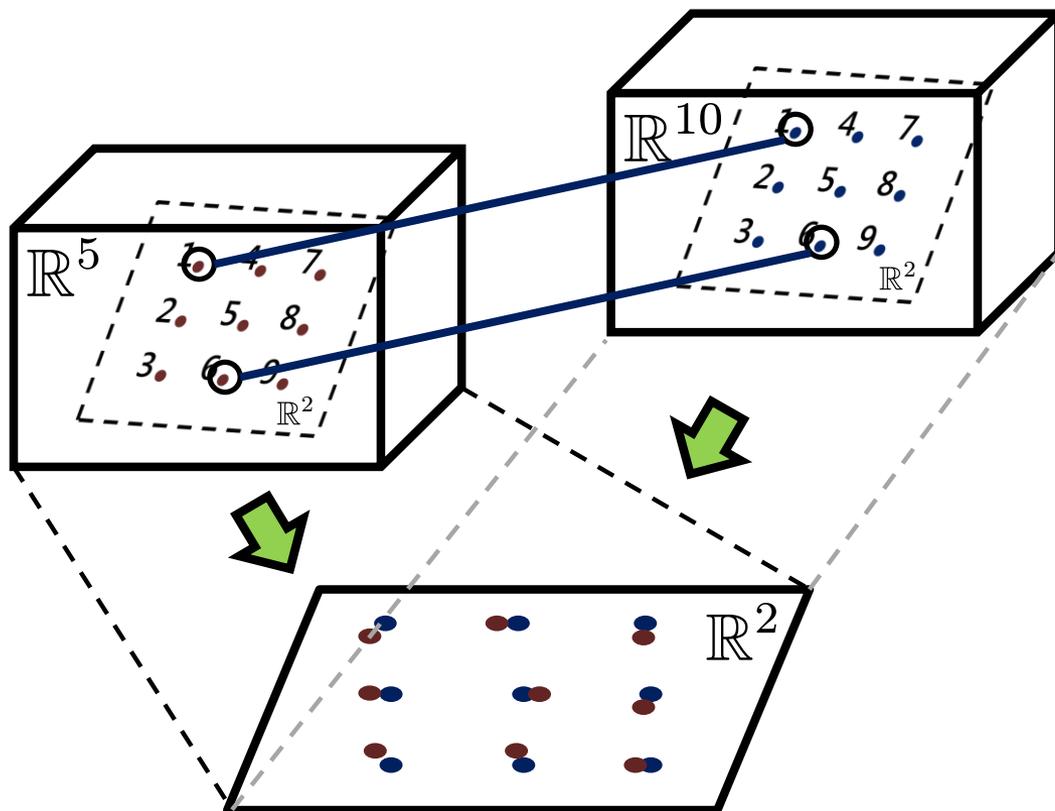
数値実験の設定



STEP1:

\bar{W} を生成

数値実験の設定



STEP1:

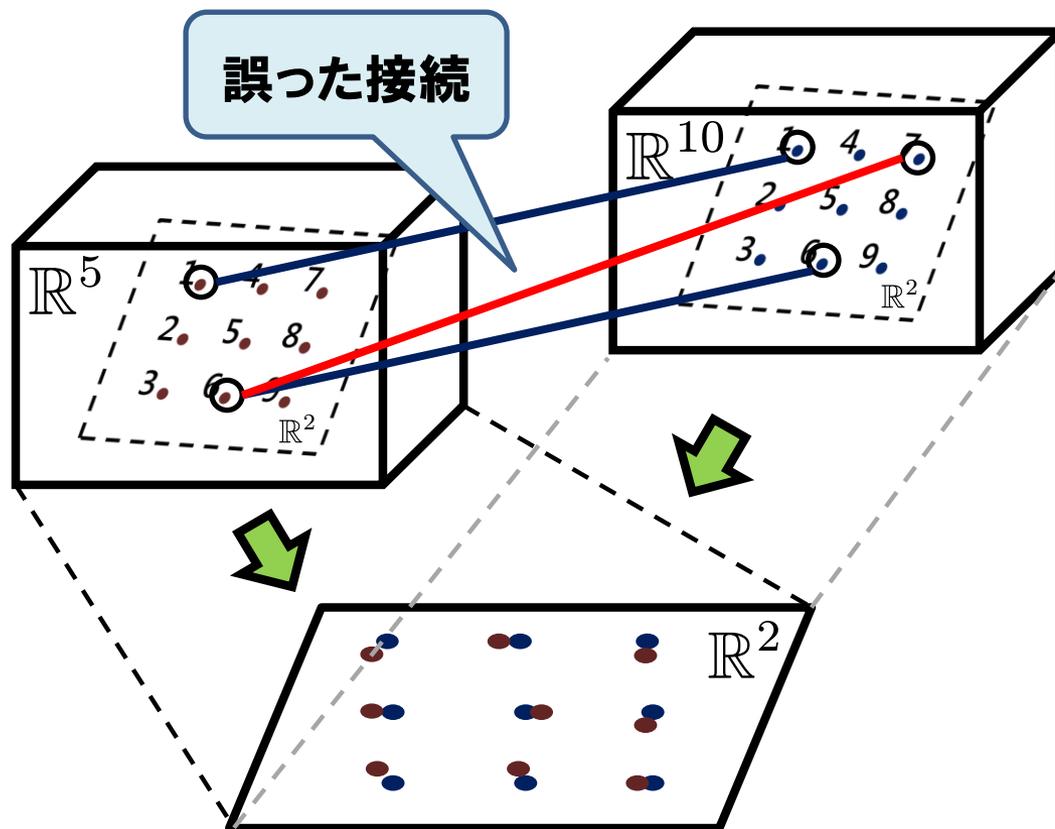
\bar{W} を生成

STEP2:

\bar{W} のリサンプリング

$= W_0$

数値実験の設定



STEP1:

\bar{W} を生成

STEP2:

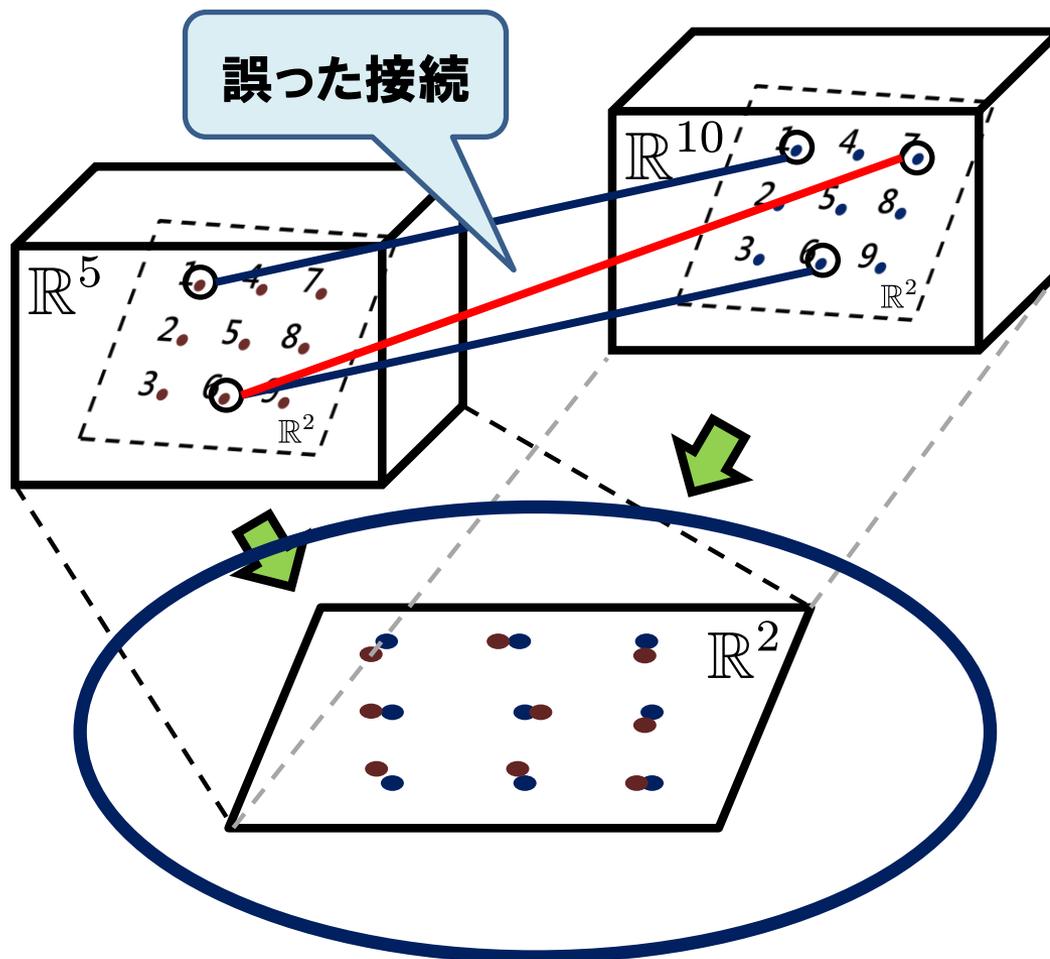
\bar{W} のリサンプリング

$$= W_0$$

STEP3:

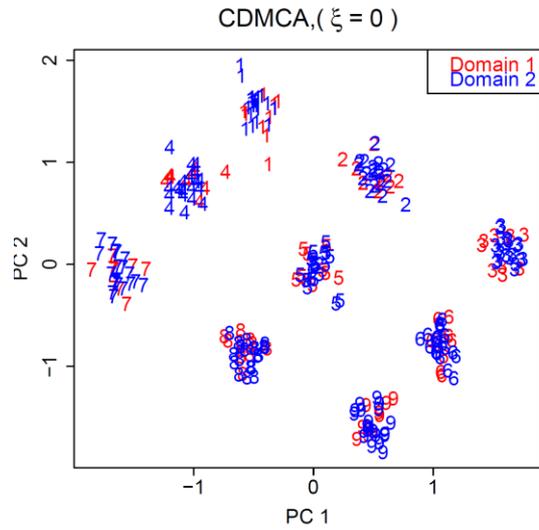
$$W_\xi = \underbrace{W_0}_{\text{正しい}} + \underbrace{C_\xi}_{\text{誤り}}$$

数値実験1

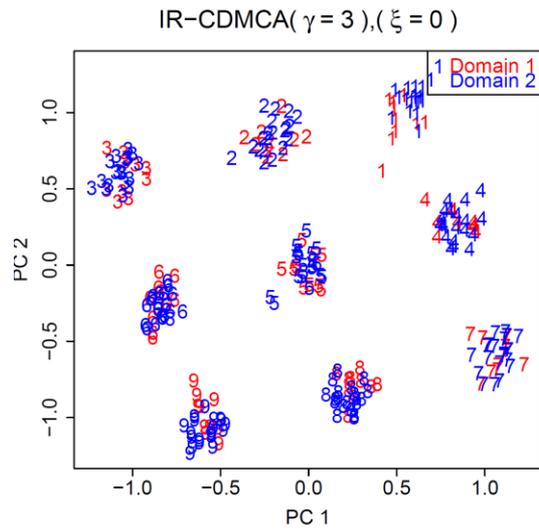


共通空間への射影はどうか

既存手法

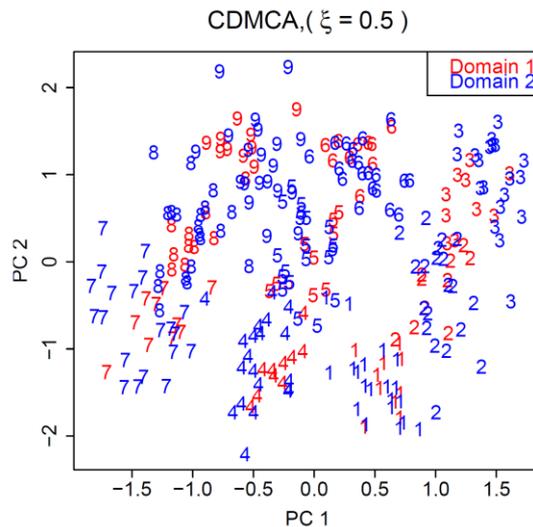
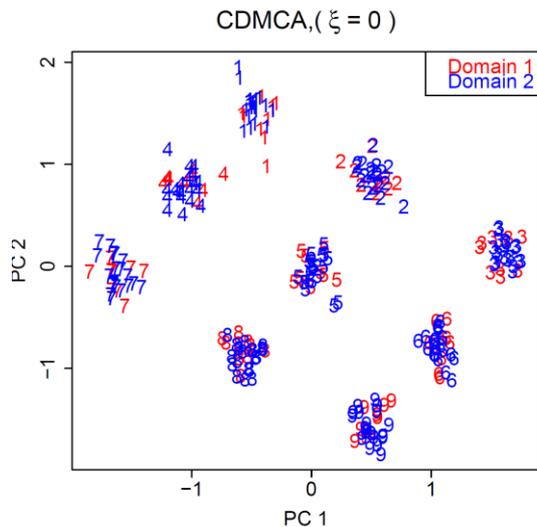


提案手法 ($\gamma = 3$)

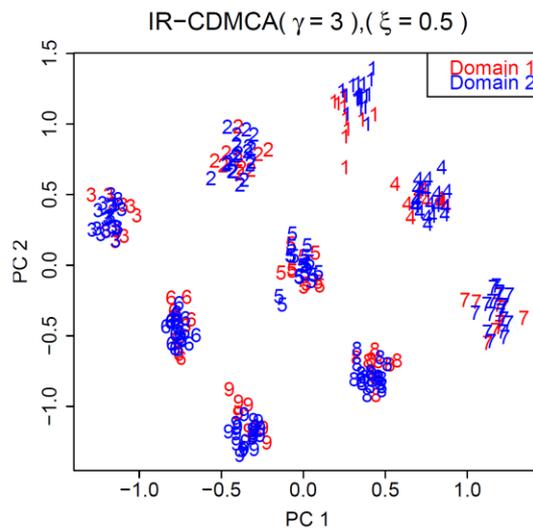
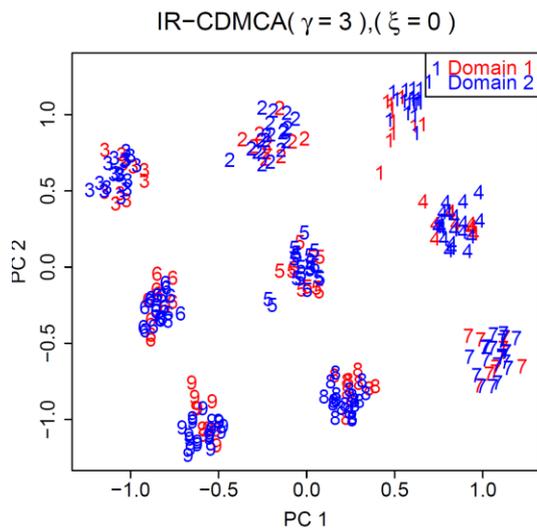


正しい接続のみ

既存手法



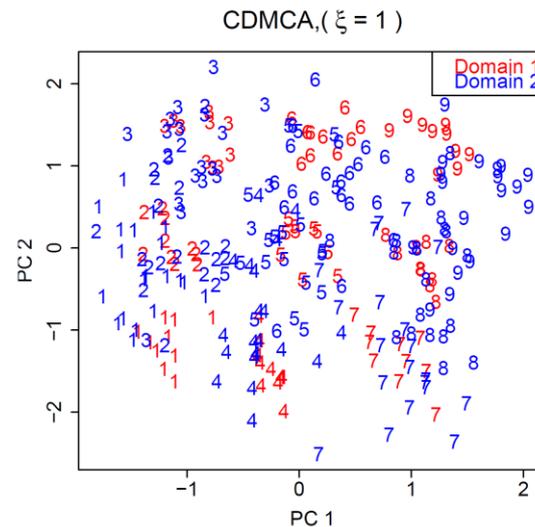
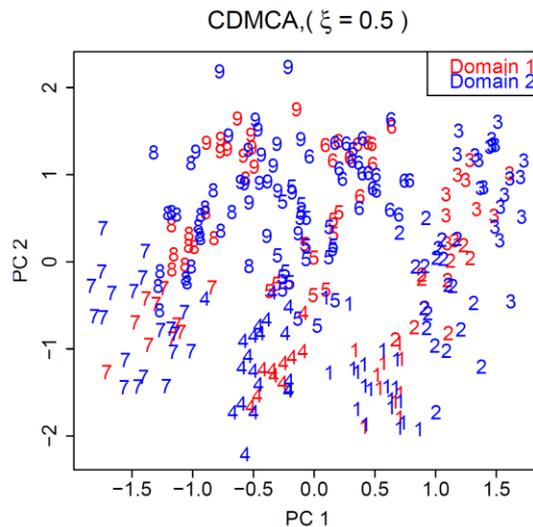
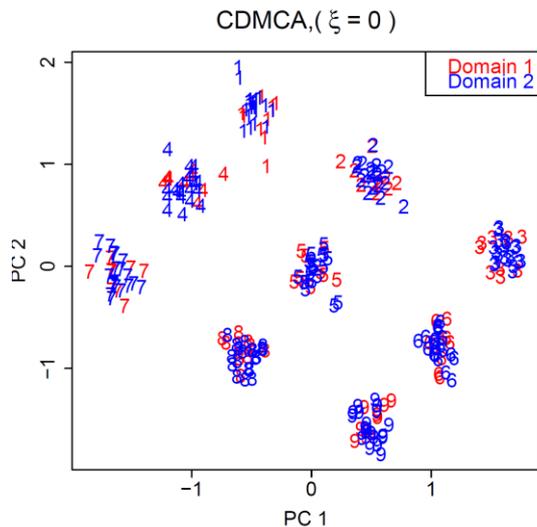
提案手法 ($\gamma = 3$)



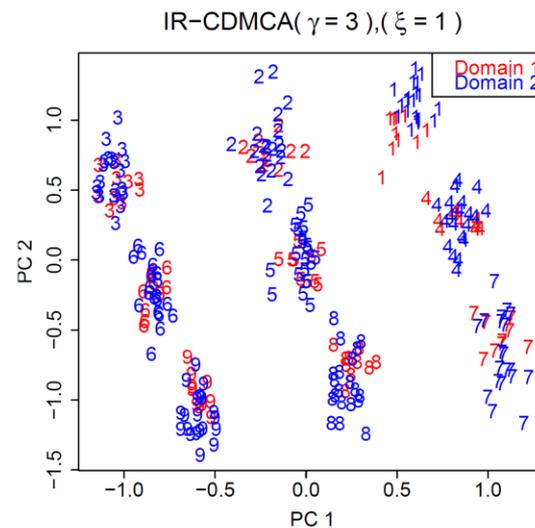
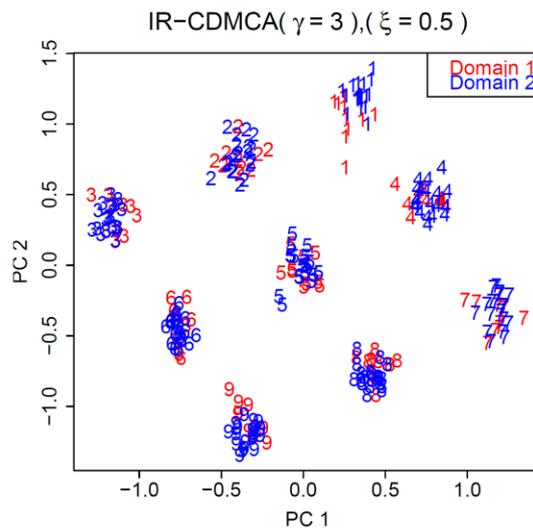
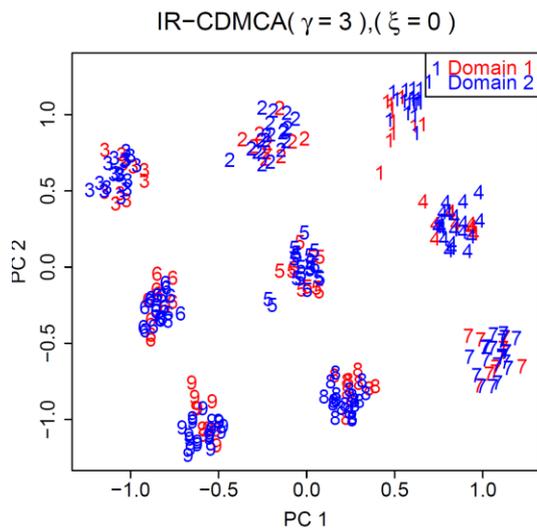
正しい接続のみ

$\xi = 0.5$

既存手法



提案手法 ($\gamma = 3$)

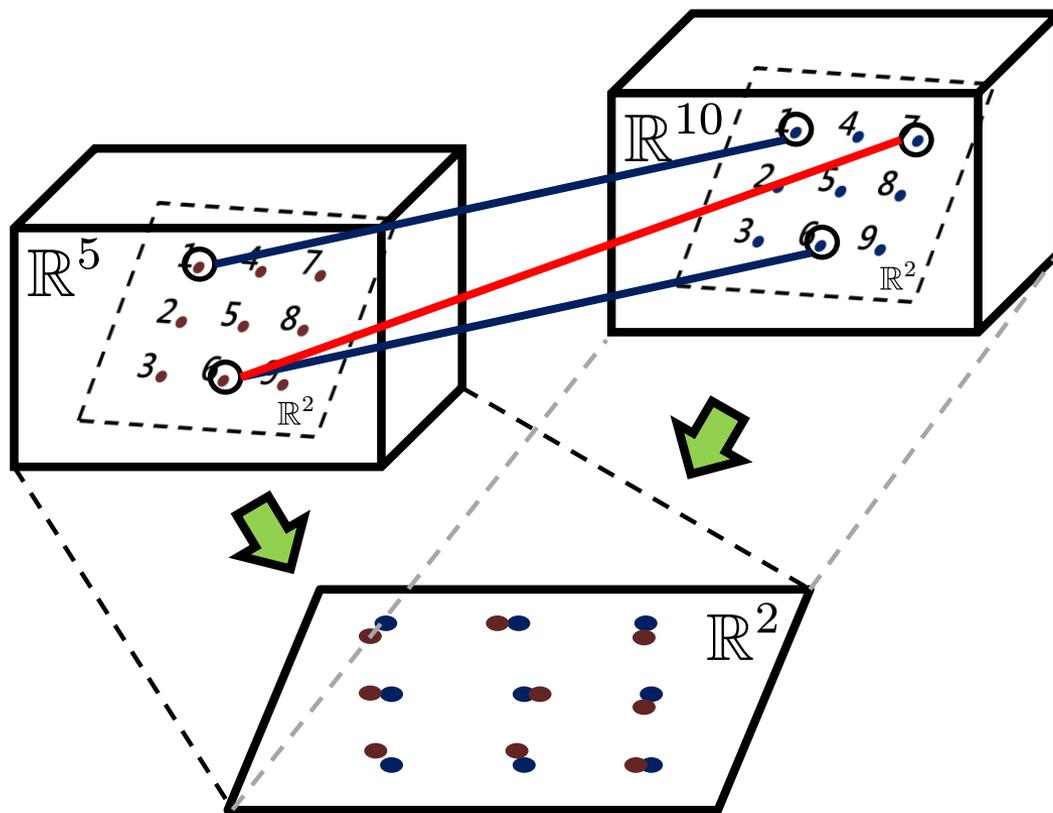


正しい接続のみ

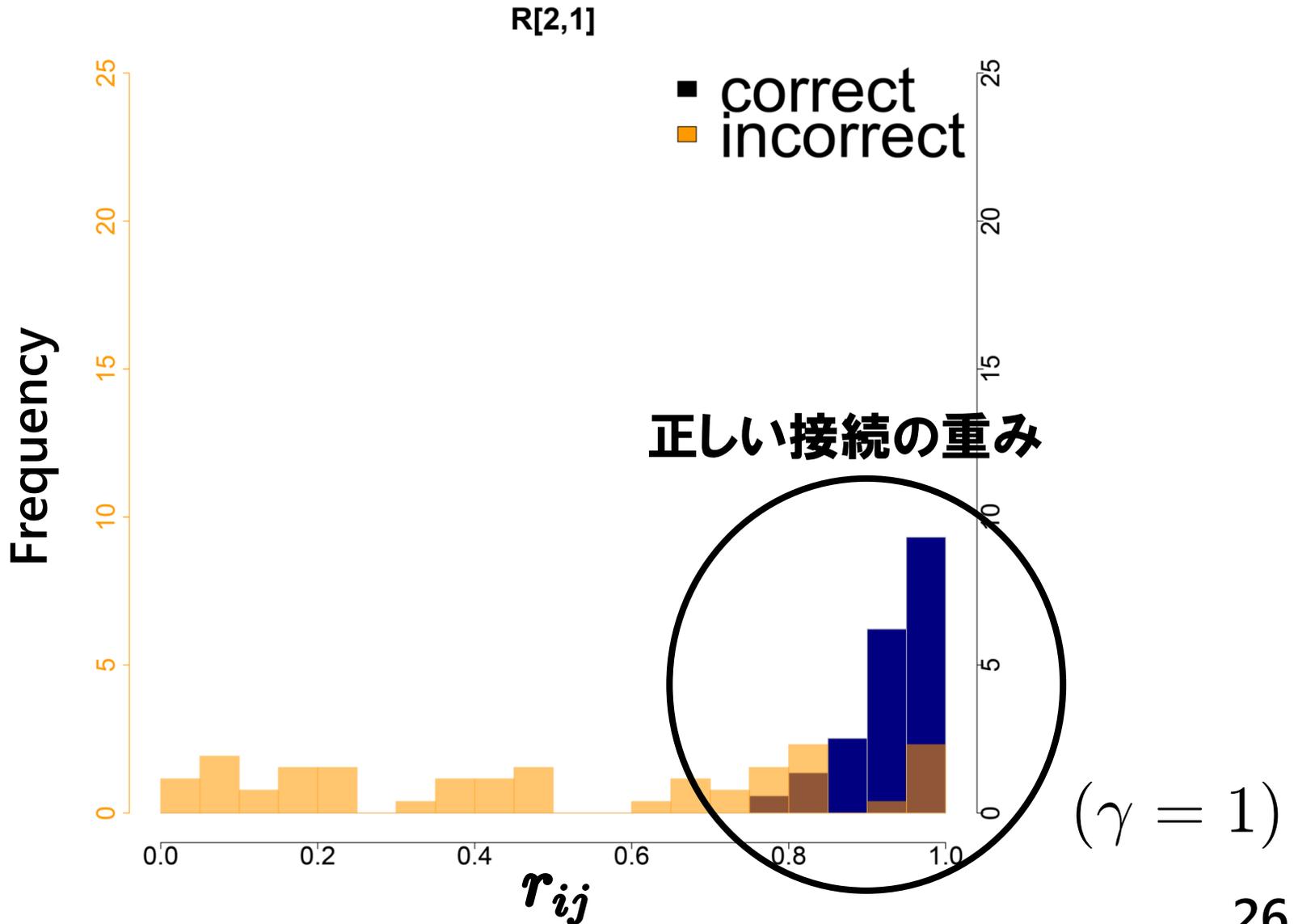
$\xi = 0.5$

$\xi = 1.0$

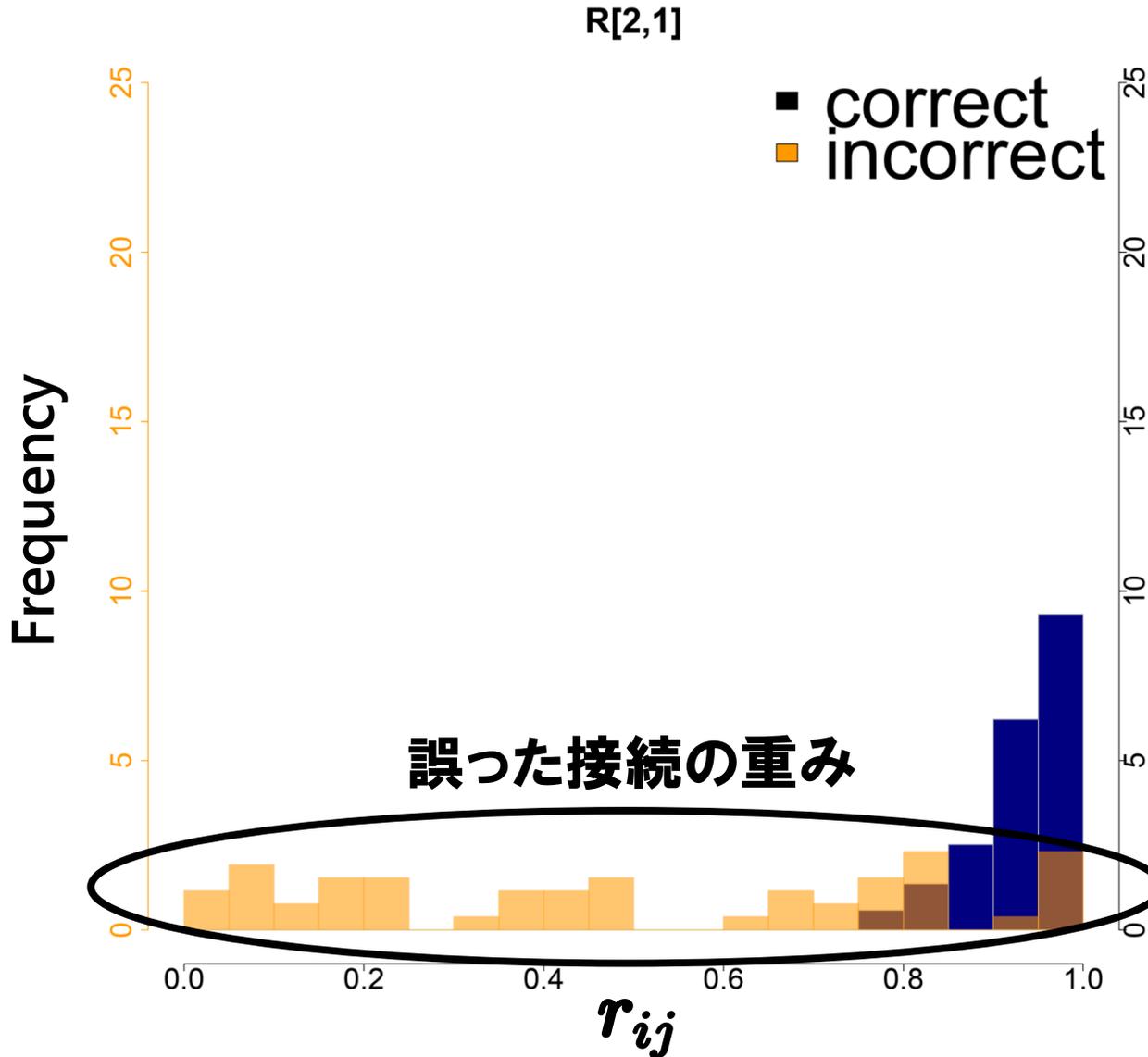
数値実験2



重みのヒストグラム ($\xi = 0.5$)



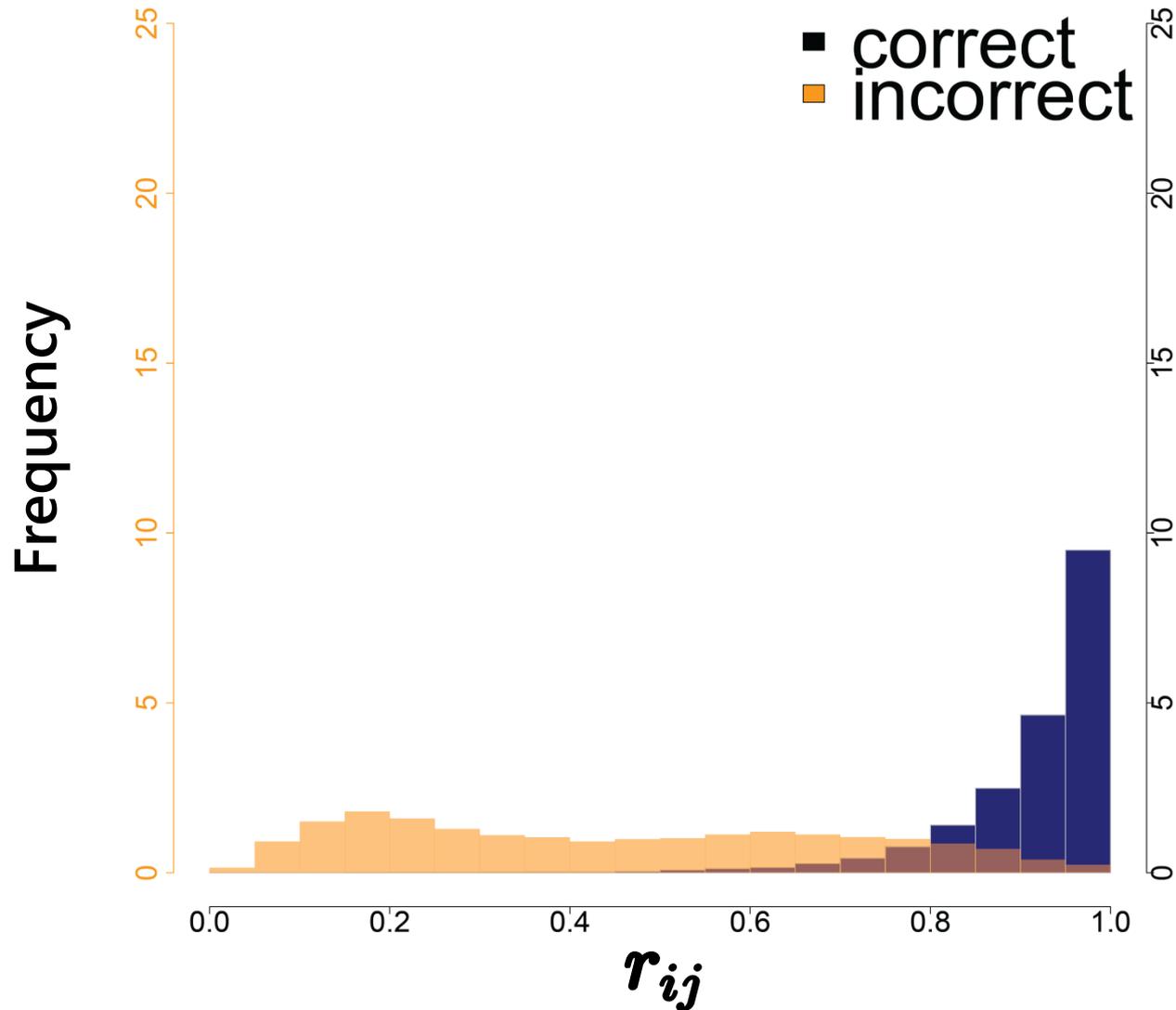
重みのヒストグラム ($\xi = 0.5$)



($\gamma = 1$)

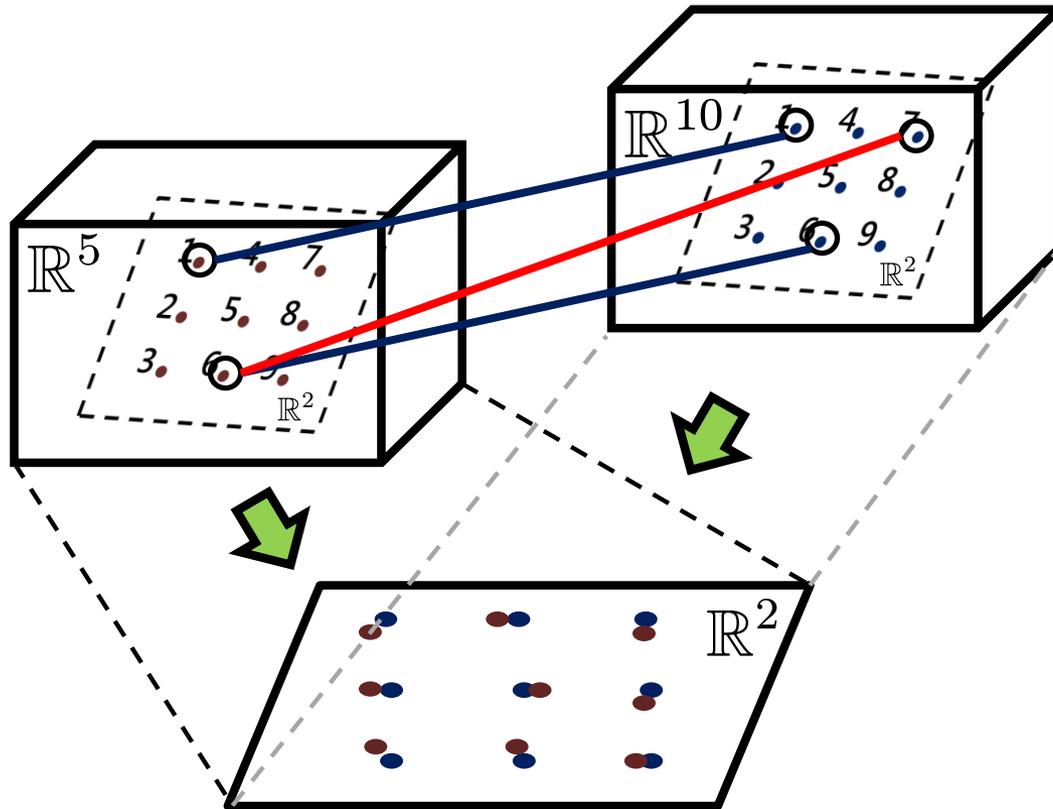
50回の平均 ($\xi = 0.5$)

R[2,1]



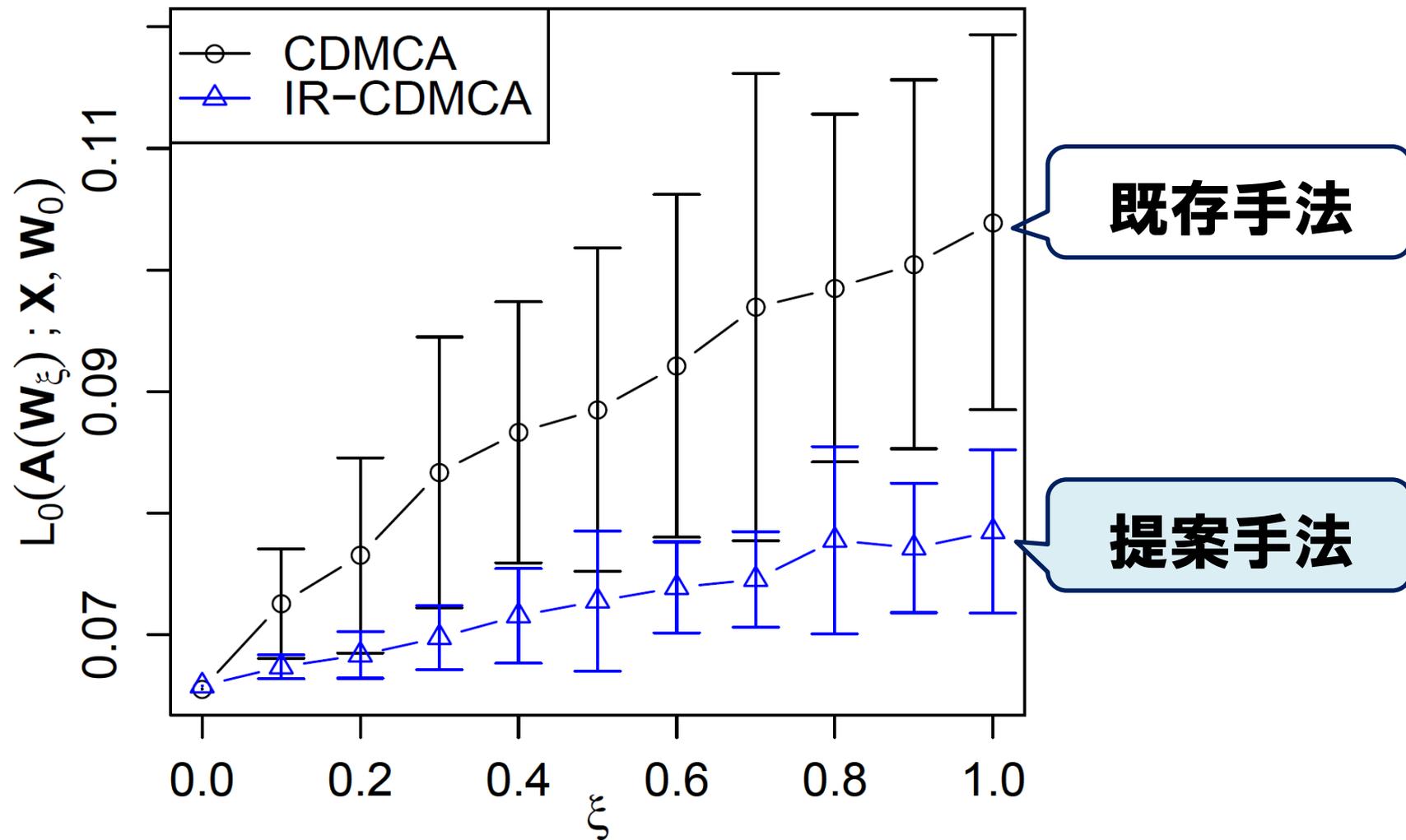
($\gamma = 1$)

数値実験3



誤ったマッチングウェイトで射影を計算

⇒ **正しい**マッチングウェイトとの損失関数を見る



Cross Validationで γ を選ぶ

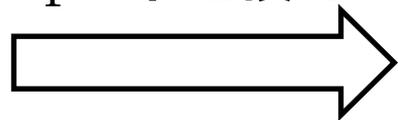
Shimodaira(2016)[3] : Matching Weightをリサンプリング

$$\mathbf{W} = \underbrace{\mathbf{W}^{\text{train}:t}}_{\alpha} + \underbrace{\mathbf{W}^{\text{test}:t}}_{1-\alpha} \quad (t = 1, 2, \dots, T)$$

$$\text{CV} := \frac{1}{T} \sum_{t=1}^T L_0(\hat{\mathbf{A}}_{\gamma}(\mathbf{W}^{\text{train}:t}); \mathbf{X}, \mathbf{W}^{\text{test}:t})$$

$$\text{CV} := \frac{1}{T} \sum_{t=1}^T L_0(\hat{\mathbf{A}}_\gamma(\mathbf{W}^{\text{train}:t}); \mathbf{X}, \mathbf{W}^{\text{test}:t})$$

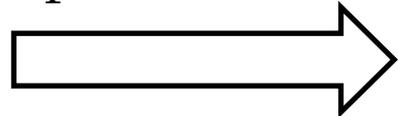
L_Γ に置き換え



$$\frac{1}{T} \sum_{t=1}^T L_\Gamma(\hat{\mathbf{A}}_\gamma(\mathbf{W}^{\text{train}:t}); \mathbf{X}, \mathbf{W}^{\text{test}:t})$$

$$CV := \frac{1}{T} \sum_{t=1}^T L_0(\hat{A}_\gamma(W^{\text{train}:t}); X, W^{\text{test}:t})$$

L_Γ に置き換え



$$\frac{1}{T} \sum_{t=1}^T L_\Gamma(\hat{A}_\gamma(W^{\text{train}:t}); X, W^{\text{test}:t})$$

推定に用いるパラメータの選択

(評価に用いる)パラメータを固定

$$\text{CV} := \frac{1}{T} \sum_{t=1}^T L_{\gamma}(\hat{\mathbf{A}}_{\gamma}(\mathbf{W}^{\text{train}:t}); \mathbf{X}, \mathbf{W}^{\text{test}:t})$$

L_{Γ} に置き換え



$$\frac{1}{T} \sum_{t=1}^T L_{\Gamma}(\hat{\mathbf{A}}_{\gamma}(\mathbf{W}^{\text{train}:t}); \mathbf{X}, \mathbf{W}^{\text{test}:t})$$

$$\text{CV} := \frac{1}{T} \sum_{t=1}^T L_0(\hat{\mathbf{A}}_\gamma(\mathbf{W}^{\text{train}:t}); \mathbf{X}, \mathbf{W}^{\text{test}:t})$$

L_Γ に置き換え


$$\frac{1}{T} \sum_{t=1}^T L_\Gamma(\hat{\mathbf{A}}_\gamma(\mathbf{W}^{\text{train}:t}); \mathbf{X}, \mathbf{W}^{\text{test}:t})$$

Leave One Out Cross Validationをすると...

$$\begin{aligned} L_\gamma(\mathbf{A}; \mathbf{X}, \mathbf{W}^{\text{test}}) &= -\frac{1}{\gamma} \log \exp(-\gamma \|\mathbf{A}^\top \mathbf{x}_i - \mathbf{A}^\top \mathbf{x}_j\|_2^2) \\ &= \|\mathbf{A}^\top \mathbf{x}_i - \mathbf{A}^\top \mathbf{x}_j\|_2^2 \\ &= L_0(\mathbf{A}; \mathbf{X}, \mathbf{W}^{\text{test}}) \end{aligned}$$

$$\frac{1}{T} \sum_{t=1}^T L_{\Gamma}(\hat{\mathbf{A}}_{\gamma}(\mathbf{W}^{\text{train}:t}); \mathbf{X}, \mathbf{W}^{\text{test}:t})$$

$$= -\frac{1}{\Gamma} \log \prod_{t=1}^T \left(\sum_{i=1}^n \sum_{j=1}^n \tilde{w}_{ij}^{\text{test}:t} \exp \left(-\Gamma \|\hat{\mathbf{A}}_{\gamma}(\mathbf{x}_i - \mathbf{x}_j)\|_2^2 \right) \right)^{1/T}$$

相乘平均

CV統計量 (提案)

$$-\frac{1}{\Gamma} \log \frac{1}{T} \sum_{t=1}^T \left(\sum_{i=1}^n \sum_{j=1}^n \tilde{w}_{ij}^{\text{test}:t} \exp \left(-\Gamma \|\hat{\mathbf{A}}_{\gamma}(\mathbf{x}_i - \mathbf{x}_j)\|_2^2 \right) \right)$$

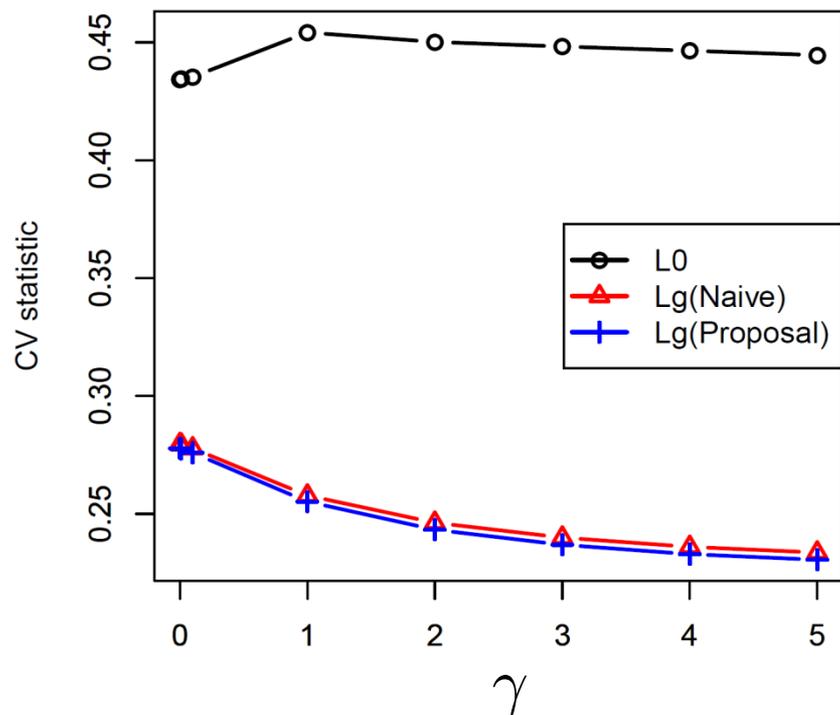
相加平均

数値実験4

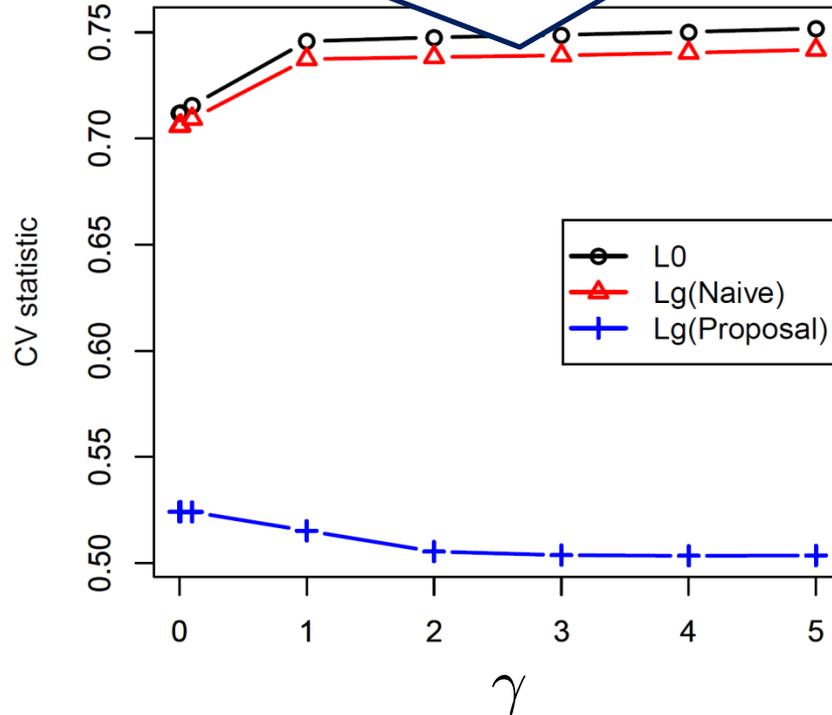
α : トレーニングデータの割合

($\Gamma = 1$)

テストデータの割合が小さい(\equiv Leave One Out)



$\alpha = 0.70$

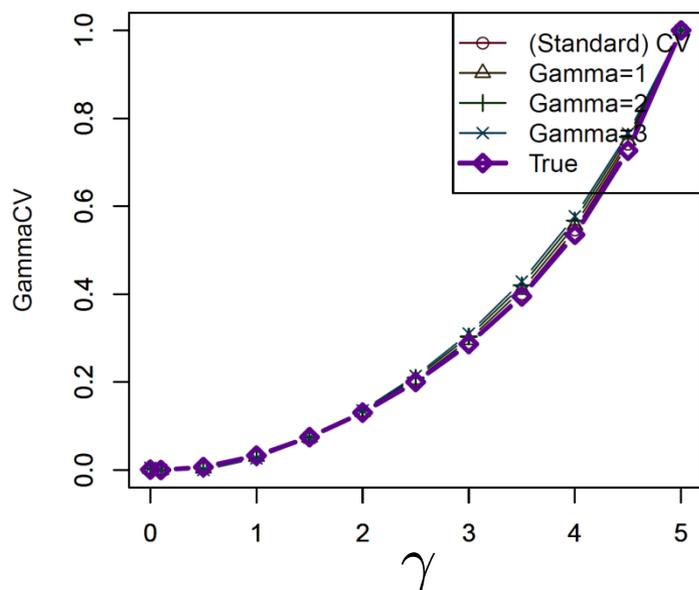


$\alpha = 0.99$

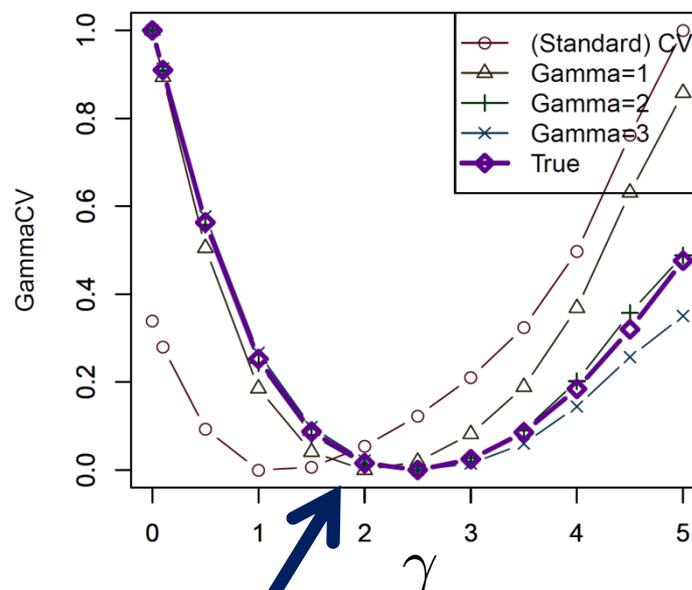
数値実験: CV統計量の計算

[0,1]に値を取るよう, 各CVと $L_0(\hat{Y}(W_\xi); W_0)$ (=True)を規格化してプロット

(誤った接続の割合=0.1)



(誤った接続の割合=1.0)

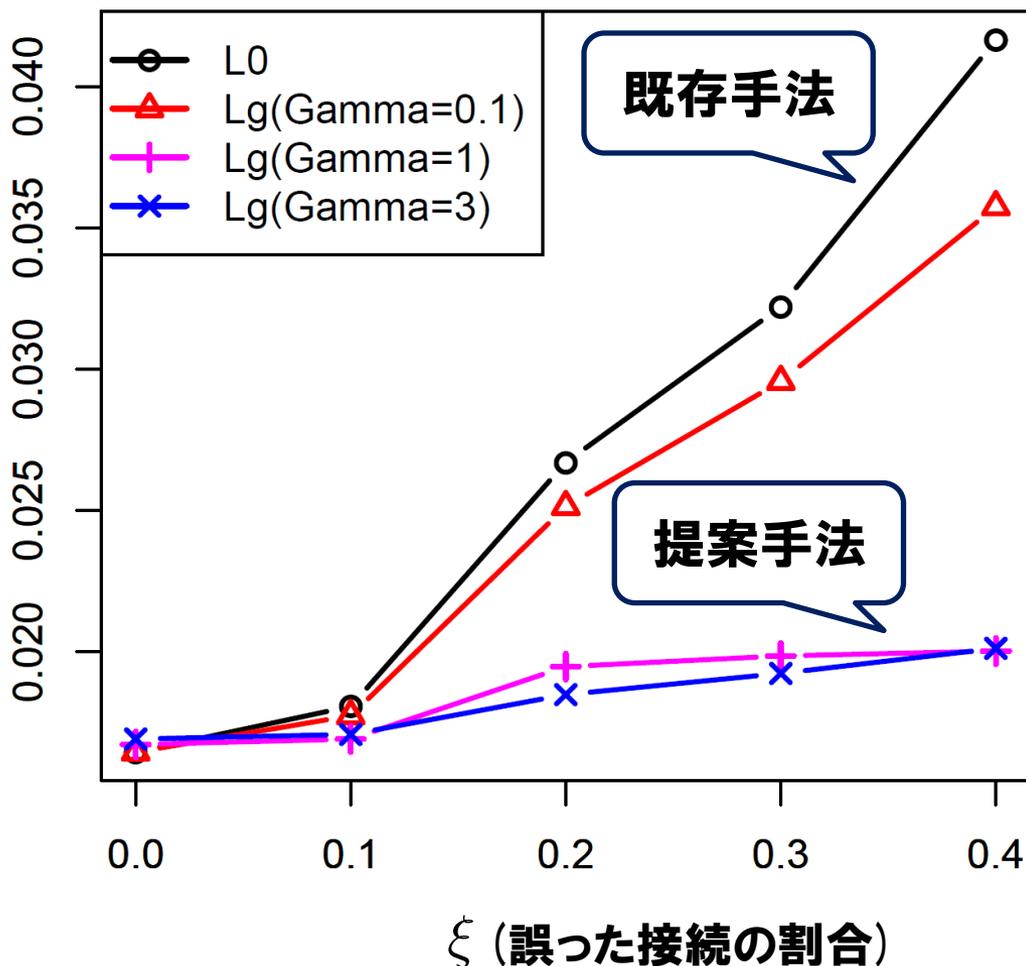


ロバスト版のCVは, 誤った関連付けの影響を軽減できる

数値実験: CV統計量の評価 (50回の平均)

W_ξ を用いて推定+パラメータ選択 \Rightarrow W_0 によって評価 (50回平均)

$$\gamma \in \{0, 0.1, 1, 3\}$$



実データ解析 (PI100データセット)

実データ：PI100 Dataset [8]

MSNショッピングから収集された5000枚の画像



(1カテゴリあたり) **100枚**ずつ × **50カテゴリ**(=タグ)
= **5000枚**の画像 からなる

PI100データセットの構成

(※画像1枚あたり、タグは1つ。)



画像特徴量

HOG特徴量[9]

(Histogram of Oriented Gradient)

タグ特徴量

ランダムベクトル

(各要素を正規分布から生成)

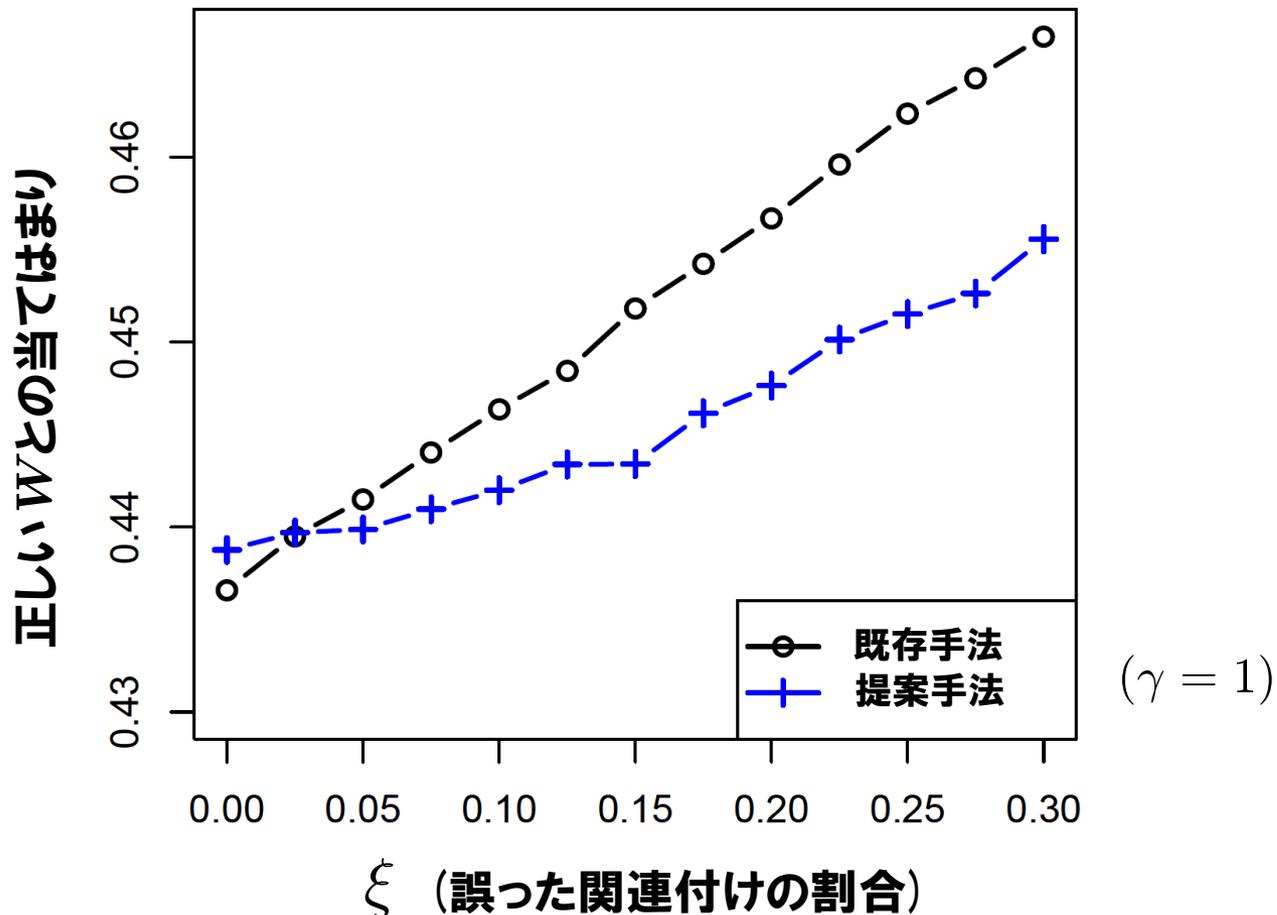
$$p_{\text{img}} = 100, p_{\text{tag}} = 30,$$

$$n_{\text{img}} = 5000, n_{\text{tag}} = 50$$

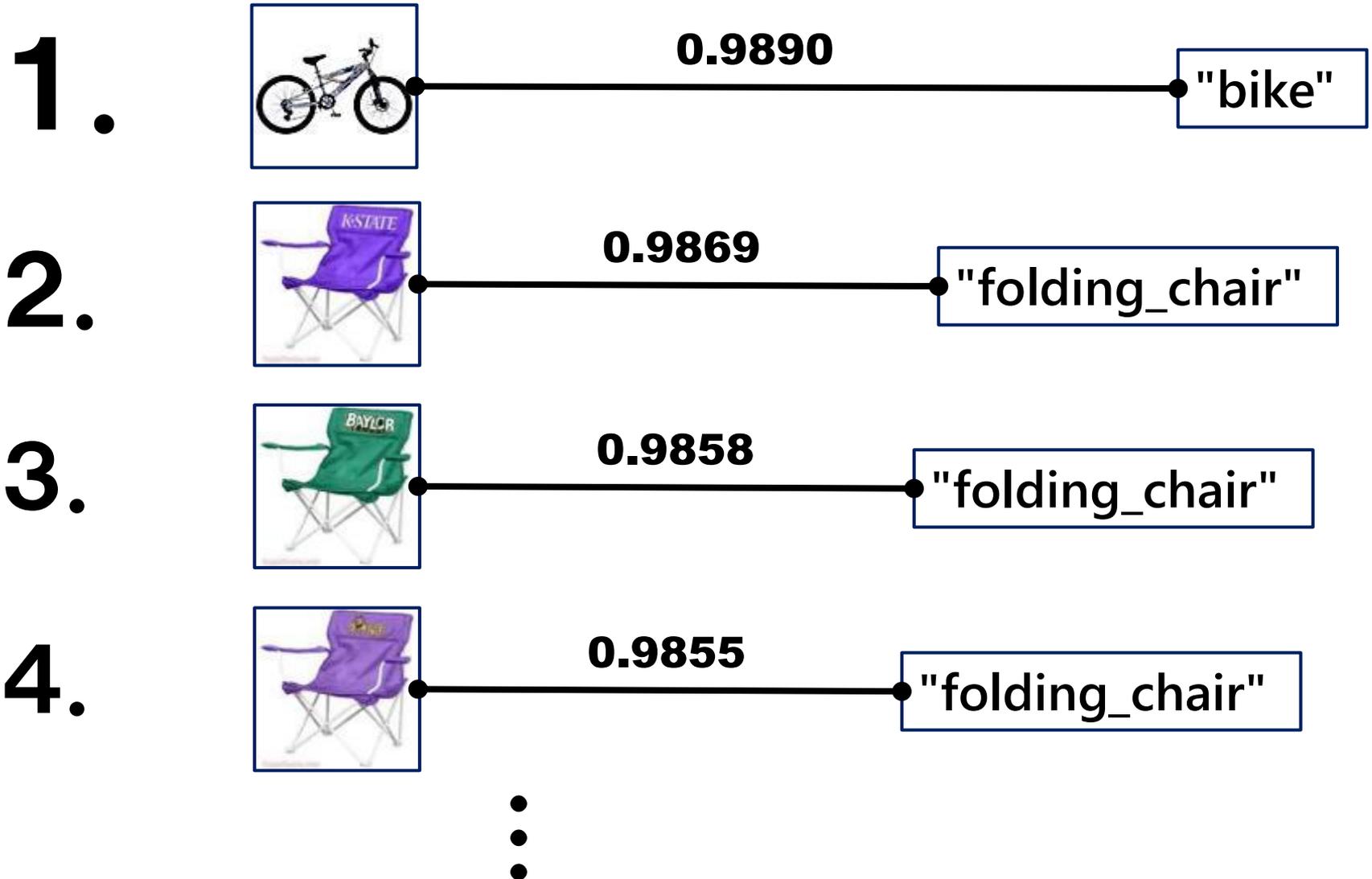
PI100データセット:汚染下での誤差

正しい W との当てはまりの良さ $L_0(\underbrace{\hat{Y}(W_\xi)}_{\text{誤まった } W \text{ で推定}}; \underbrace{W_0}_{\text{正しい } W \text{ で評価}})$ のグラフ

誤まった W で推定 正しい W で評価



(接続の)重みが大きい順にランキング



•
•
•

7497.



0.0296

"goblet"

7498.



0.0267

"glasses"

7499.



0.0157

"computer_keyboard"

7500.



0.0077

"handheld_GPS"

■ まとめ

1. 重み付きマッチング相関分析を提案した。
2. 提案手法はMM-Algorithmであり, 射影の更新が収束することを示した。
3. 提案手法が誤った関連付けの影響を軽減することを数値実験と実データで確認した。
4. ロバストなCVを提案し, パラメータが適切に選択できることを数値実験で確認した。

参考文献

- [1] Belkin, M. & Niyogi, P. Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering. *Adv. Neural Inf. Process. Syst.* **14**, 585-591 (2001).
- [2] Yan, S. *et al.* Graph embedding and extensions: A general framework for dimensionality reduction. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**, 40–51 (2007).
- [3] Shimodaira, H. Cross-validation of matching correlation analysis by resampling matching weights. *Neural Networks.* **75**, 126-140 (2016).
- [4] Nori, N., Bollegala, D. & Kashima, H. Multinomial relation prediction in social data: a dimension reduction approach. Proc. Twenty-Sixth AAAI Conf. Artif. Intell. 115-121 (2012).
- [5] Hotelling, B. Y. H. Relations between two sets of variates. *Biometrika* **28**, 321-377 (1936).

- [6] Kettenring, J. R. Analysis of Several Sets of Variables. *Biometrika*, **58**(3), 433-451 (1971).
- [7] Ortega, J. M., & Werner, C. R. Iterative solution of nonlinear equations in several variables. Academic Press, *New York* (1970); reissued as *Classics Appl. Math.* **30**, SIAM, Philadelphia, (2000).
- [8] Xie, X. *et al.* Mobile search with multimodal queries. *Proceedings of the IEEE*. **96**(4). (2008).
- [9] Navneet, D. & Bill, T. Histogram of Oriented Gradients for Human Detection. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. **1** (2005).
- [10] Fujisawa, H. & Eguchi, S. Robust parameter estimation with a small bias against heavy contamination. *J. Multivar. Anal.* **99**, 2053–2081 (2008).
- [11] Chew, S. H. A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox. *Econometrica* **51**, 1065-1092 (1983).

- [P1] 奥野 彰文, 下平 英寿. 「ロバストな多ドメインマッチング相関分析の提案と画像認識への応用」. ポスター発表. 第18回情報論的学習理論ワークショップ. 筑波. 2015年11月.
- [P2] 奥野 彰文, 福井 一輝, 下平 英寿. 「多ドメインでのマッチング相関分析のロバスト化とその応用」. 口頭発表. 2015年度統計関連学会連合大会. 岡山. 2015年9月.
- [P3] Okuno, A. & Shimodaira, H. "Robust cross-domain matching: Analyzing multi-domain data vectors under mismatched associations". Poster presentation at the Machine Learning Summer School. Kyoto. August, 2015.
- [P4] 福井 一輝, 奥野 彰文, 下平 英寿. 「マッチング相関分析を用いた画像-マルチタグ間の相互検索」. ポスター発表. 第18回画像の認識・理解シンポジウム. 大阪. 2015年7月.
- [P5] 奥野 彰文, 下平 英寿. 「ロバストな情報量規準の提案とその応用」. 2014年度統計関連学会連合大会. 東京. 2014年9月.

參考資料

Simple Coding (Shimodaira, (2016) [3])

D :ドメイン数 K :共通空間の次元 $\sum_{d=1}^D n_d = N, \sum_{d=1}^D p_d = P$
 n_d :ドメインdの次元 p_d :ドメインdのデータ数

$$\mathbf{W}^{de} = (w_{ij}^{de}) \in \mathbb{R}^{n_d \times n_e}$$

$$\mathbf{W} = [\mathbf{W}^{de}] \in \mathbb{R}^{N \times N}$$

$$\mathbf{X}^d = (\mathbf{x}_1^d \cdots \mathbf{x}_{n_d}^d)^\top \in \mathbb{R}^{n_d \times p_d}$$

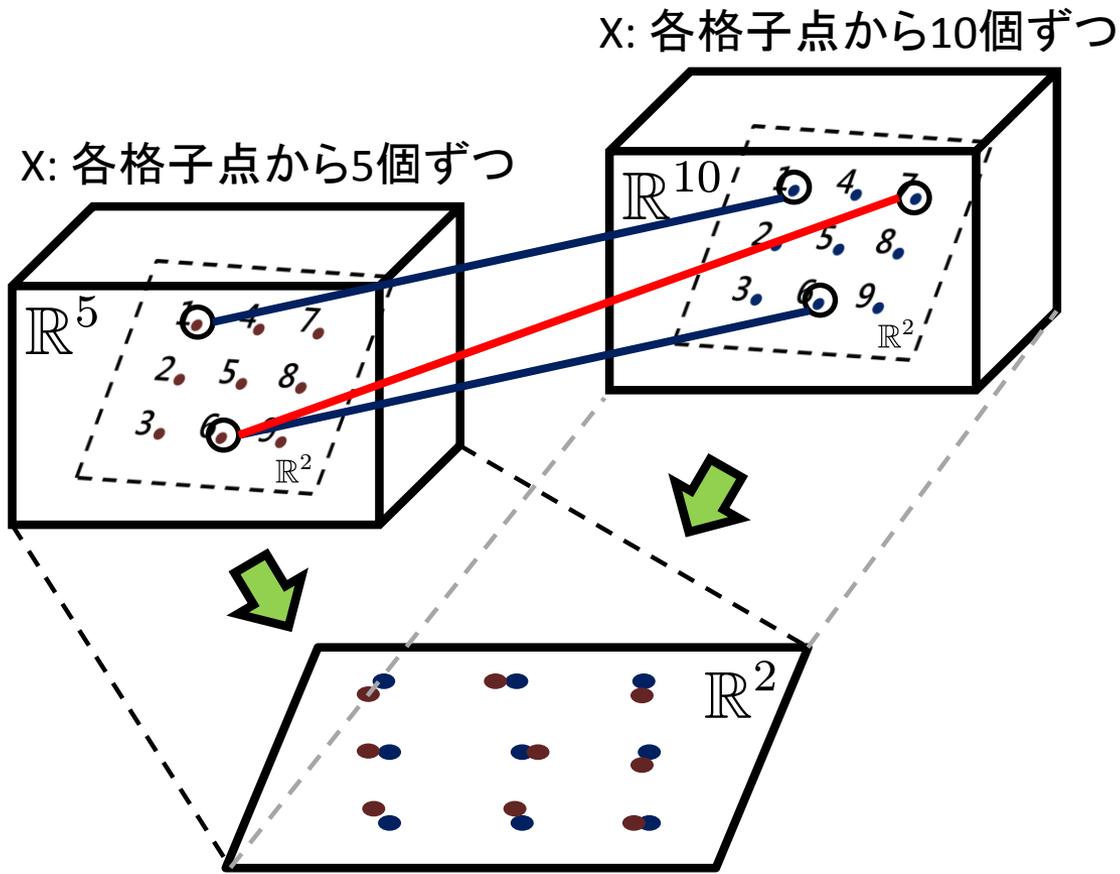
$$\mathbf{X} = \text{Diag}[\mathbf{X}_1, \cdots, \mathbf{X}_D] \in \mathbb{R}^{N \times P}$$

$$\mathbf{A}^d \in \mathbb{R}^{p_d \times K}$$

$$\mathbf{A} = ((\mathbf{A}^1)^\top \cdots (\mathbf{A}^D)^\top)^\top \in \mathbb{R}^{P \times K}$$

データの生成手順

- (1) \mathbb{R}^Q 上の $\underbrace{D \times D \cdots \times D}_{Q \text{ 個}}$ グリッドの点 $\mathbf{x} \in \{1, \dots, \ell\}^Q$ を順に並べ, $(\mathbf{x}_1^0)^\top, \dots, (\mathbf{x}_{\ell^Q}^0)^\top$ とする.
 $i = \ell^Q + 1, \ell^Q + 2, \dots$ についても, $\mathbf{x}_{\ell^Q+1}^0 = \mathbf{x}_1^0, \mathbf{x}_{\ell^Q+2}^0 = \mathbf{x}_2^0, \dots$ となるよう繰り返す.
- (2) 行列 $\mathbf{B}^d \in \mathbb{R}^{p_d \times Q}$, ($d \in \{1, \dots, D\}$) の各要素を独立に $N[0, 1]$ から生成する. \mathbf{B}^d を使って, 各ドメインのデータ $\mathbf{x}_i^d = \mathbf{B}^d \mathbf{x}_i^0 + \boldsymbol{\varepsilon}_i^d$, ($i = 1, \dots, n_d$) を生成する. ただし $\boldsymbol{\varepsilon}_i^d$ は独立に $N(0, \sigma^2 \mathbf{I}_{p_d})$ に従う. 生成された \mathbf{X}^d の各列を平均 0, 分散 1 になるよう標準化する.
- (3) 各格子点について, 各ドメイン $d \in \{1, \dots, D\}$ から生成されたデータベクトルの数は m_1, \dots, m_D である. これら m_1, \dots, m_D 個のデータベクトルは全て同じ格子点から生成されているので, マッチングウェイトを $w_{ij}^{de} = 1$ とする. 同様の操作を全ての格子点について同様にマッチングウェイトを指定する. なお, 同一ドメイン内の 2 つのデータベクトルのマッチングウェイト w_{ij}^{dd} は常に 0 であるとする. こうして得られたマッチング行列を, $\bar{\mathbf{W}}$ で書き, ドメイン d とドメイン e の間のマッチング行列を特に $\bar{\mathbf{W}}^{de}$ で書く.
- (4) $\bar{\mathbf{W}}^{de}$ を割合 α でランダムにリサンプリングし, \mathbf{W}_0^{de} とする. $d \in \{2, \dots, D\}, e \in \{1, \dots, (d-1)\}$ について, 同様のリサンプリングを行う. ただし, 同一ドメイン内のマッチングウェイトは $\mathbf{W}_0^{dd} = \mathbf{0}$ とし, かつ $\mathbf{W}_0^{de} = (\mathbf{W}_0^{ed})^\top$ とする. \mathbf{W}_0^{de} を併せて $\mathbf{W} (= [\mathbf{W}_0^{de}])$ と書く.

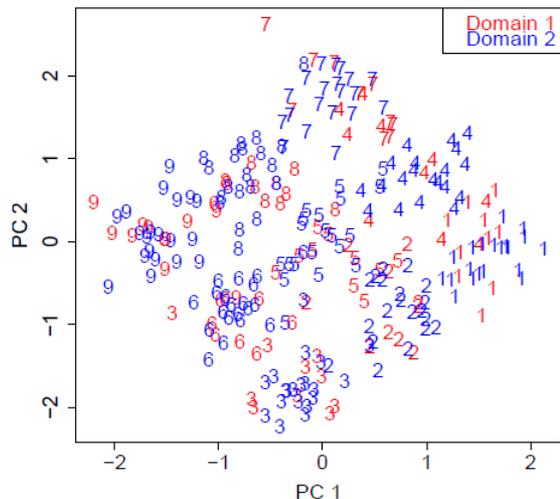
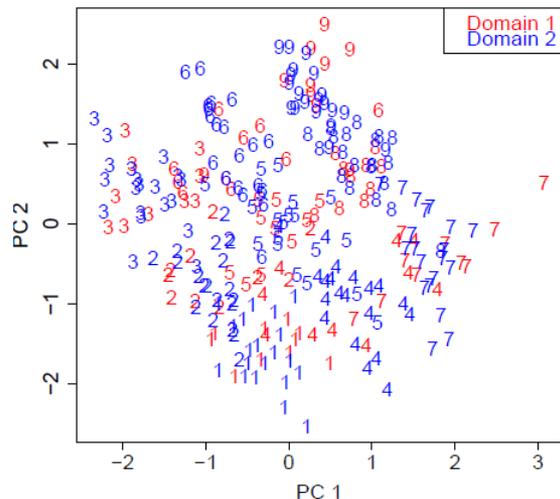
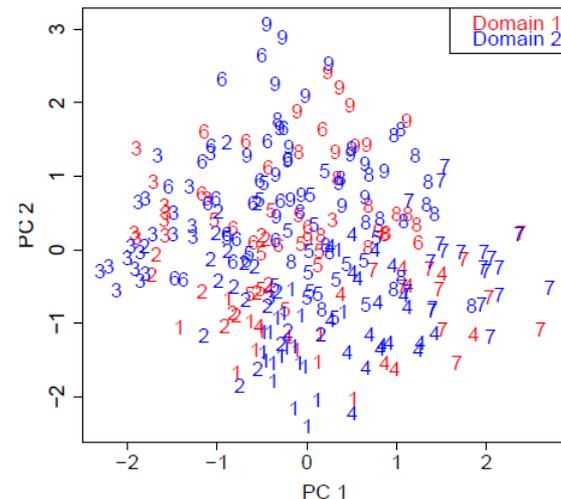
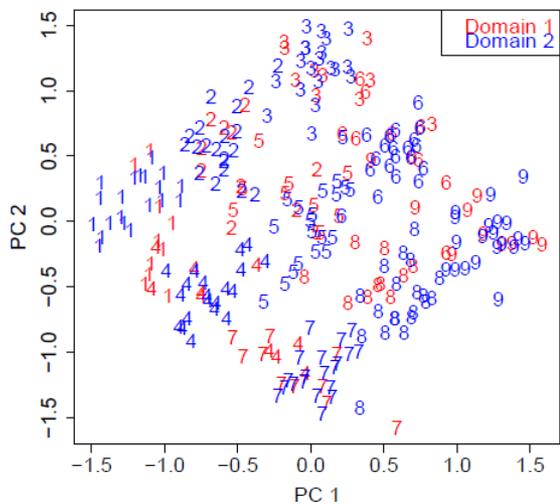
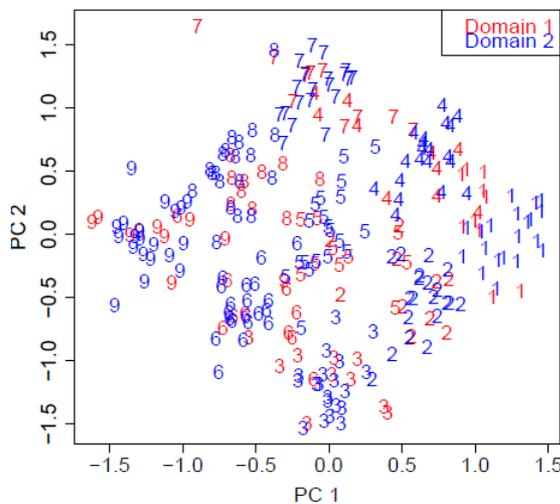
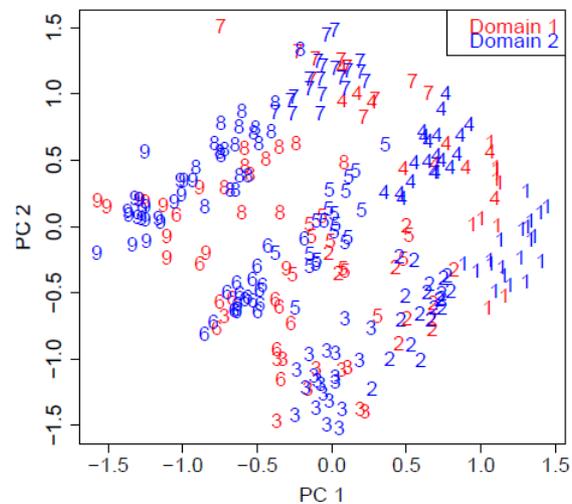


$$\mathbf{x}_i^0 \in \{-1, 0, 1\}^2$$

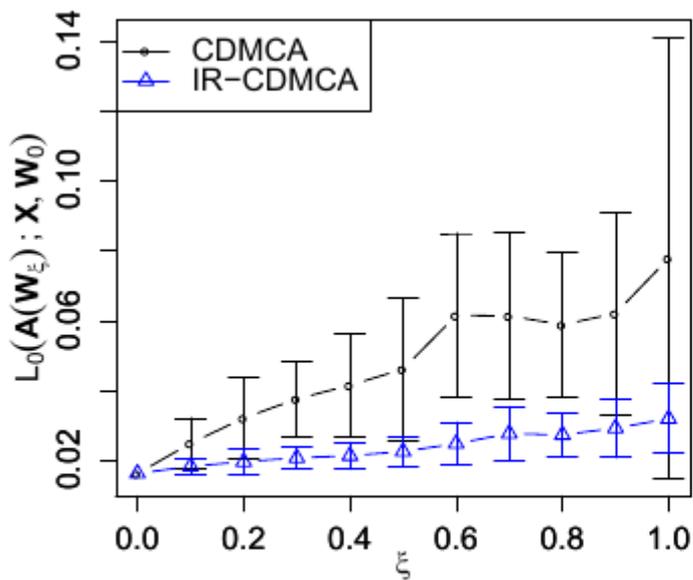
$$\mathbf{x}_i^d = \mathbf{B}^d \mathbf{x}_i^0 + \boldsymbol{\varepsilon}_i^d, \quad \boldsymbol{\varepsilon}_i^d \sim N[\mathbf{0}, 0.2\mathbf{I}_{p_d}]$$

$$\alpha = 0.02$$

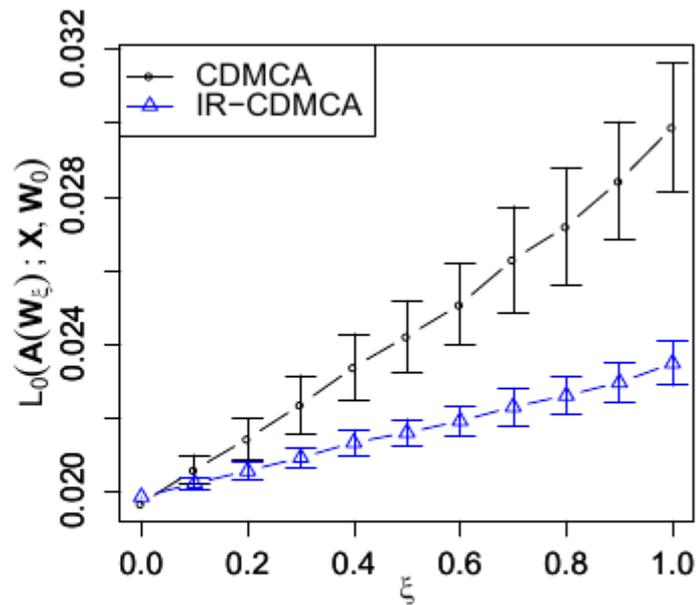
$$D = 2, \sigma^2 = 0.5$$

 CDMCA, ($\xi = 0$)

 CDMCA, ($\xi = 0.5$)

 CDMCA, ($\xi = 1$)

 IR-CDMCA ($\gamma = 3$), ($\xi = 0$)

 IR-CDMCA ($\gamma = 3$), ($\xi = 0.5$)

 IR-CDMCA ($\gamma = 3$), ($\xi = 1$)


$$D = 2, \sigma^2 = 0.2$$

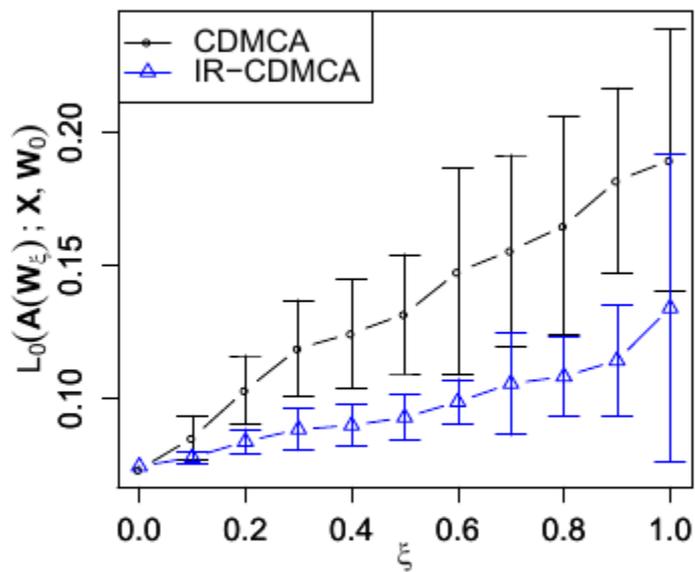


(a) $\alpha = 0.02$

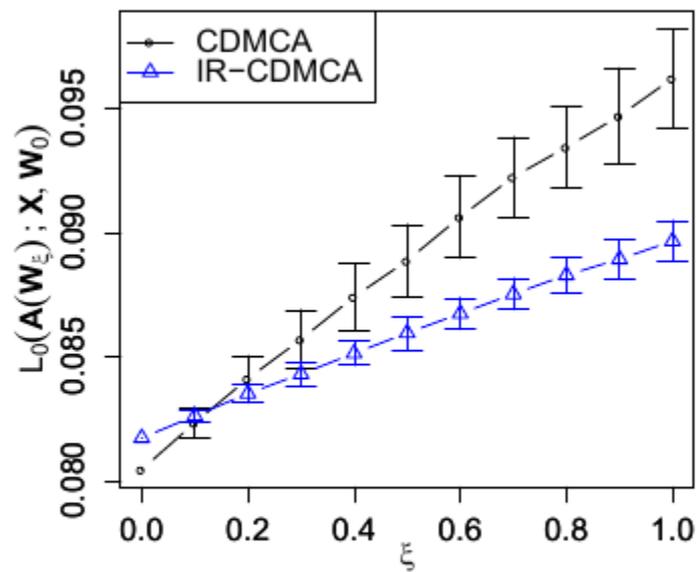


(b) $\alpha = 1$

$$D = 2, \sigma^2 = 0.5$$

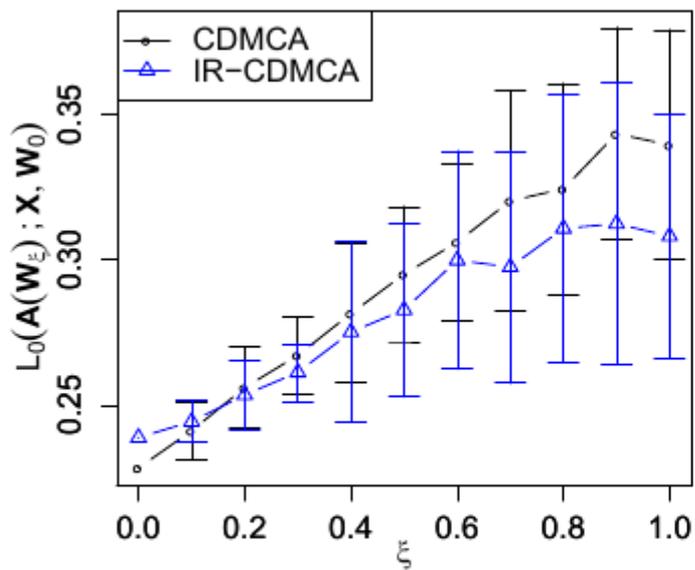


(a) $\alpha = 0.02$

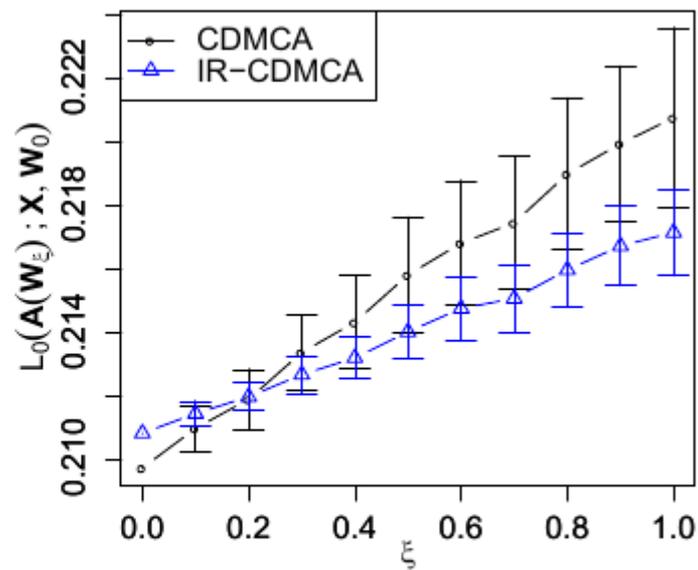


(b) $\alpha = 1$

$$D = 2, \sigma^2 = 1$$



(a) $\alpha = 0.02$



(b) $\alpha = 1$

γ - 損失関数に関連した規準

γ -Divergence (Robust divergence)[10]

$$d_\gamma(Q, P) = -\frac{1}{\gamma} \log \int p(x)^\gamma dQ(x) + \frac{1}{1+\gamma} \log \int p(x)^{1+\gamma} dx$$

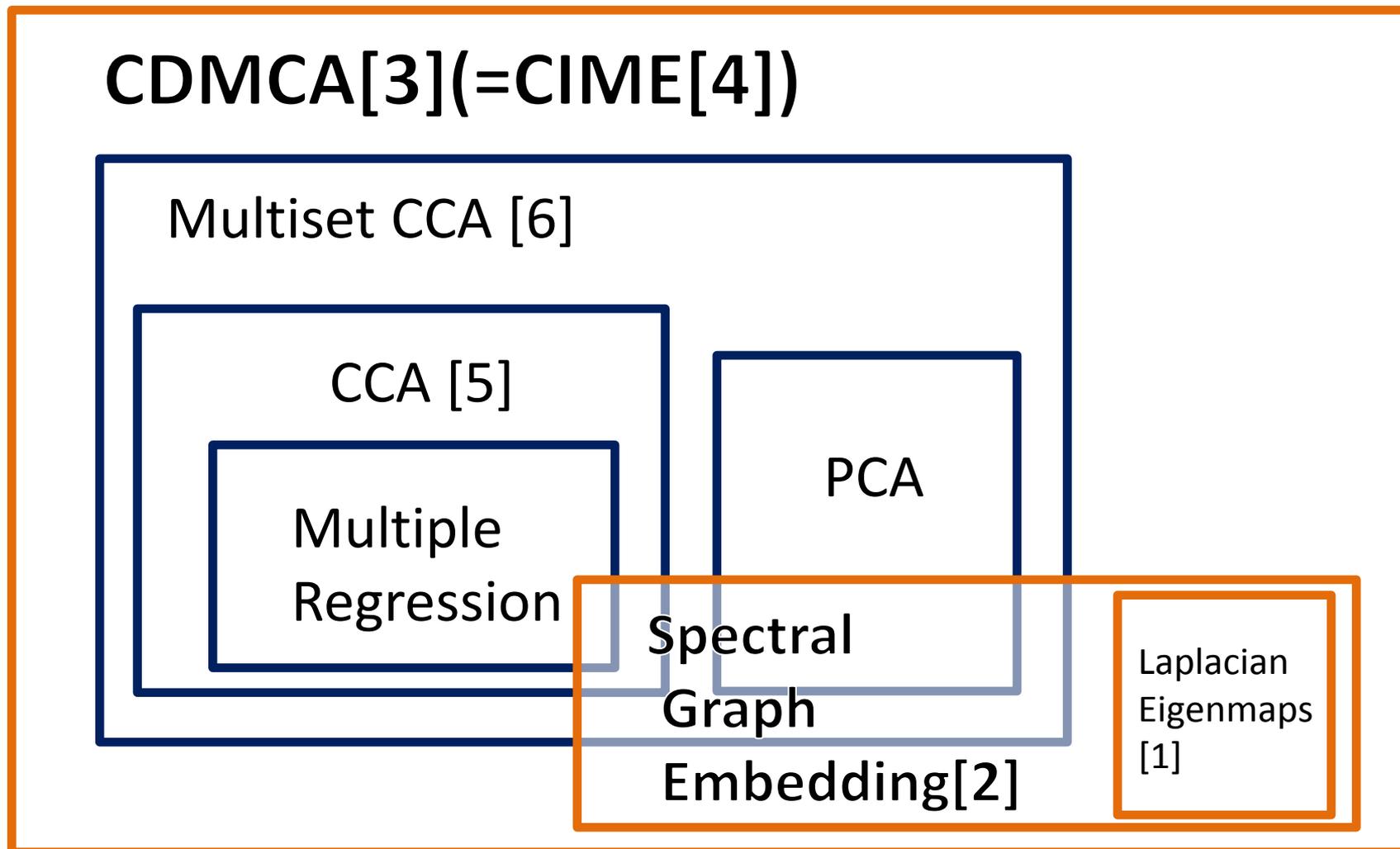
$$P(x) = N[\mu; 1/2]$$

疑似平均 (Quasilinear Mean)[11]

$$\phi^{-1} \left(\frac{1}{n} \sum_{i=1}^n \phi(x_i) \right) \text{ に非線形変換 } \phi(x) = \exp(-\gamma x)$$

など.

多変量解析手法の包含関係



関連研究:

Laplacian Eigenmaps (Belkin and Niyogi, (2001)[1])

既知 $\mathbf{y}_1, \dots, \mathbf{y}_n \in \mathbb{R}^p$

$$\text{接続行列を生成: } w_{ij} = \begin{cases} 1 & (\|\mathbf{y}_i - \mathbf{y}_j\|_2^2 < \varepsilon) \\ 0 & (\text{Otherwise}) \end{cases}$$

接続行列の重み付け: $w_{ij} \leftarrow w_{ij} r_{ij}$

$$(r_{ij} = \exp(-\gamma \|\mathbf{y}_i - \mathbf{y}_j\|_2^2))$$

$$\hat{\mathbf{Y}} = \arg \min_{\mathbf{Y}^\top \mathbf{M} \mathbf{Y} = \mathbf{I}} L_0(\mathbf{Y}; \mathbf{W})$$