



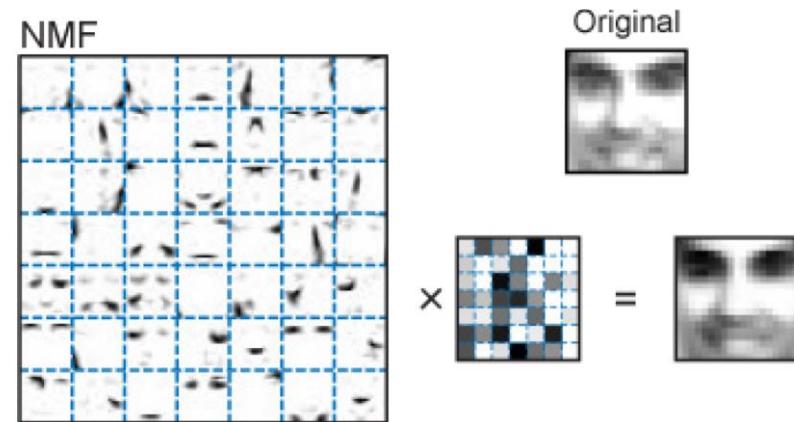
画像データに対する 非負値行列因子分解 (NMF) について

横浜市立大学 植 墓

内容1. いくつかのNMFの紹介

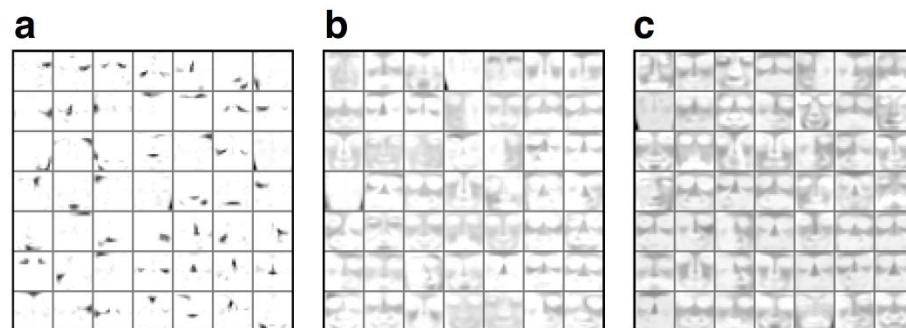
■ Non-negative Matrix Factorization (NMF)

Lee(1999), Lee(2001)



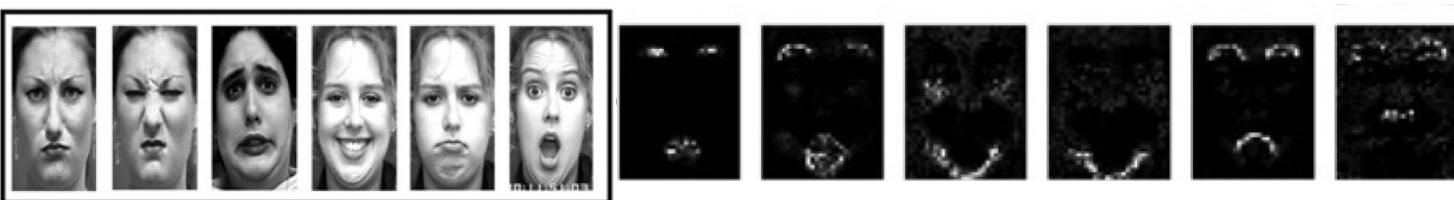
■ NMF with Sparseness Constraints (SNMF)

Hoyer(2004)



■ Discriminant NMF (DNMF)

Zafeiriou(2006), Kotsia(2008)



■ 新規手法の検討

内容2. フェンシング画像データへの適用の検討



PoseNetによる関節推定

スポーツデータ解析コンペでは、
フェンシングの映像データが扱われている。
「得点した(突いた)瞬間」の画像に対する
NMFなどの適用を検討していく。

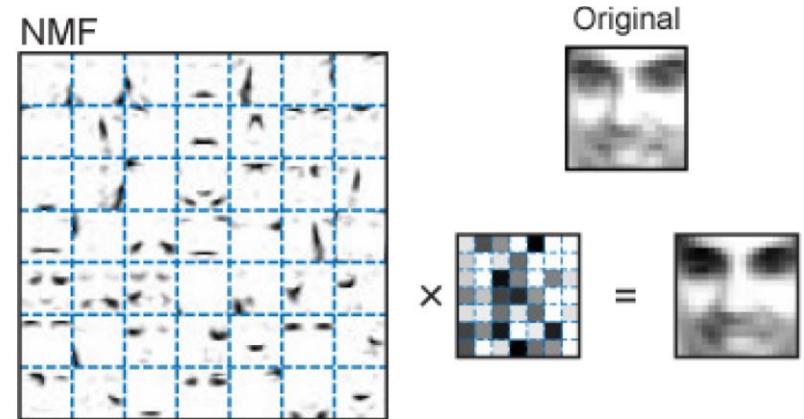


引用：スポーツデータ解析コンペティション発会式チュートリアル

内容1. いくつかのNMFの紹介

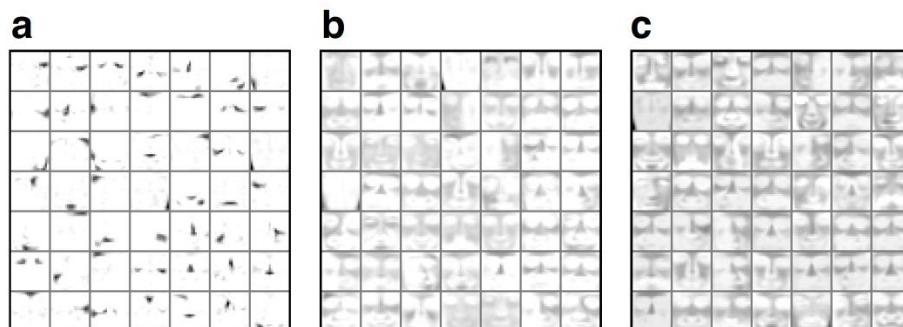
■ Non-negative Matrix Factorization (NMF)

Lee(1999), Lee(2001)



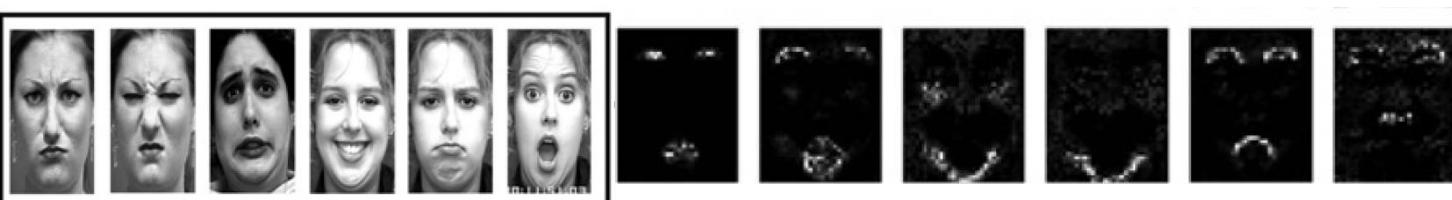
■ NMF with Sparseness Constraints (SNMF)

Hoyer(2004)



■ Discriminant NMF (DNMF)

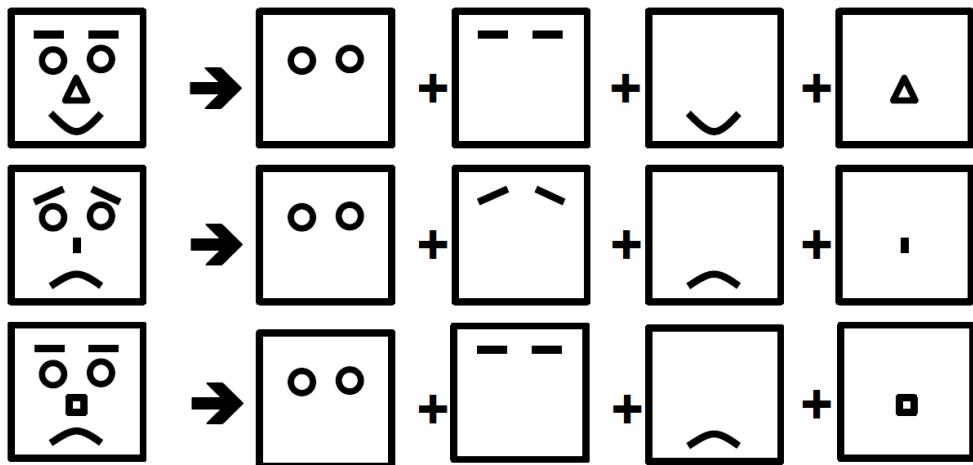
Zafeiriou(2006), Kotsia(2008)



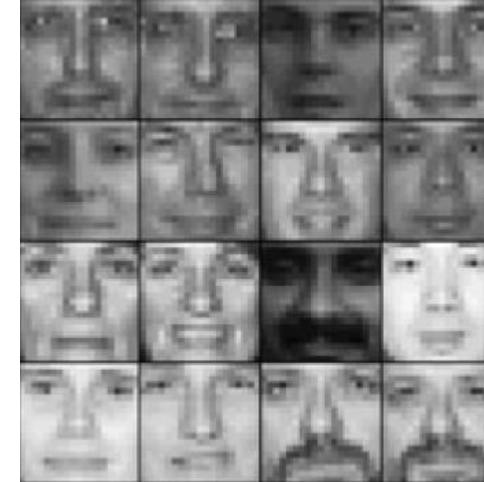
■ 新規手法の検討

NMF (Lee,1999) のモチベーション

- 顔の画像からパーツを抽出したい。
- 顔の画像をパーツの組み合わせで表現したい。 (parts-based representation)



引用：横田達也, テンソル分解の基礎と画像信号処理への応用, (2004)



CBCL顔データセットの一部

- パーツ ≈ クラスターの把握と解釈 (ソフトクラスタリング)
- パーツを特徴量として使って回帰や分類 (次元圧縮)

NMF のモデル

サンプルサイズ (顔画像の枚数) : n

次元 (ピクセルの数) : $p (= p_1 \times p_2)$

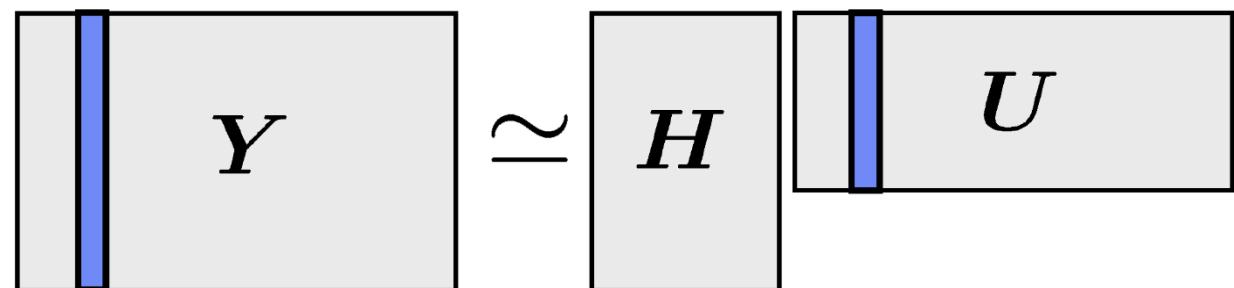
非負の観測値ベクトル (輝度値) : $y_1, y_2, \dots, y_n \in \mathbb{R}^{\geq 0, p}$

非負の基底ベクトル (パーツ) : $h_1, h_2, \dots, h_m \in \mathbb{R}^{\geq 0, p}$ [$m < n, p$]

非負の重み係数 : $u_{1,1}, u_{1,2}, \dots, u_{m,n} \in \mathbb{R}$

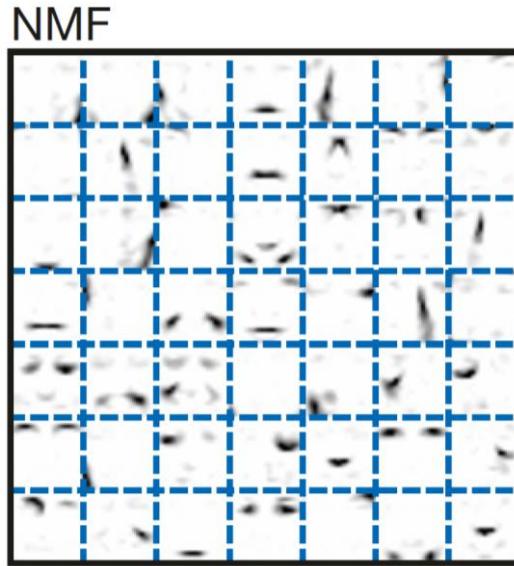
$$y_i \approx \sum_{j=1}^m h_j u_{j,i} \quad \boxed{Y = [y_1, y_2, \dots, y_n] \\ H = [h_1, h_2, \dots, h_m] \\ U = \{u_{ij}\}} \quad \rightarrow \quad Y \approx HU (= X)$$

観測値を
「**非負の基底の足し合わせ**」
で近似表現する。



基底の非負値制約について

「非負値データ（物理量, カウントデータ等）の構成要素（パーツ）である基底も非負であるべき」という考え方による。



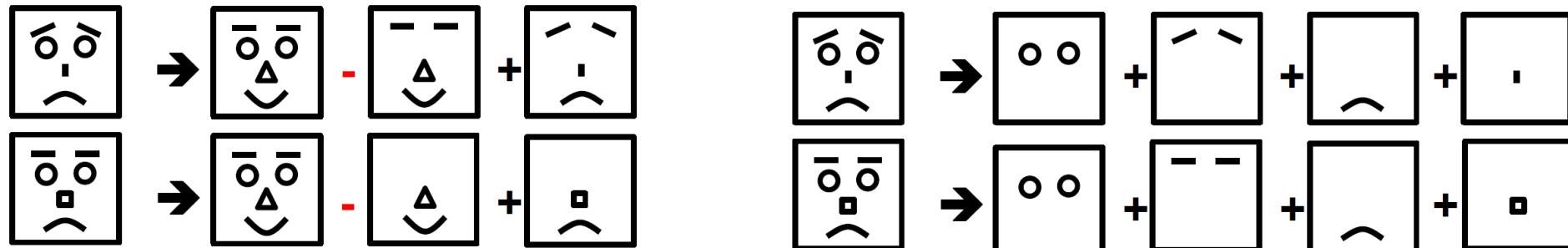
CBCLデータに対するNMFの基底とPCAの主成分
黒：正の値 赤：**負の値**
(詳細は後述)

引用 : Lee(1999)

PCA のように基底（主成分）に非負値を許してしまうと、解釈が難しくなる。

係数の非負値制約について

「観測値(顔)はパーツの足し合わせで表されるべきで、引き算は不自然」という考えによる。



引用：横田達也, テンソル分解の基礎と画像信号処理への応用, (2004)

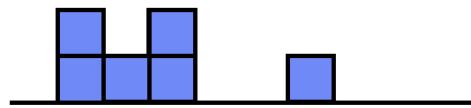
以上の2つの制約により,
「非負の基底ベクトルの非負結合」で観測値ベクトルを表現することになる。

それにより, 係数ベクトルの推定値がスペース(疎)になる傾向が生じる。

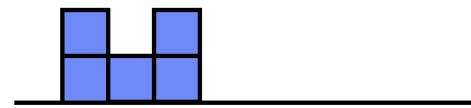
係数ベクトルのスパース性

- 係数が負の値を取っても良い場合

観測データ



基底ベクトル



係数

$\times 1$

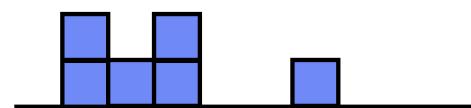


$\times 1$



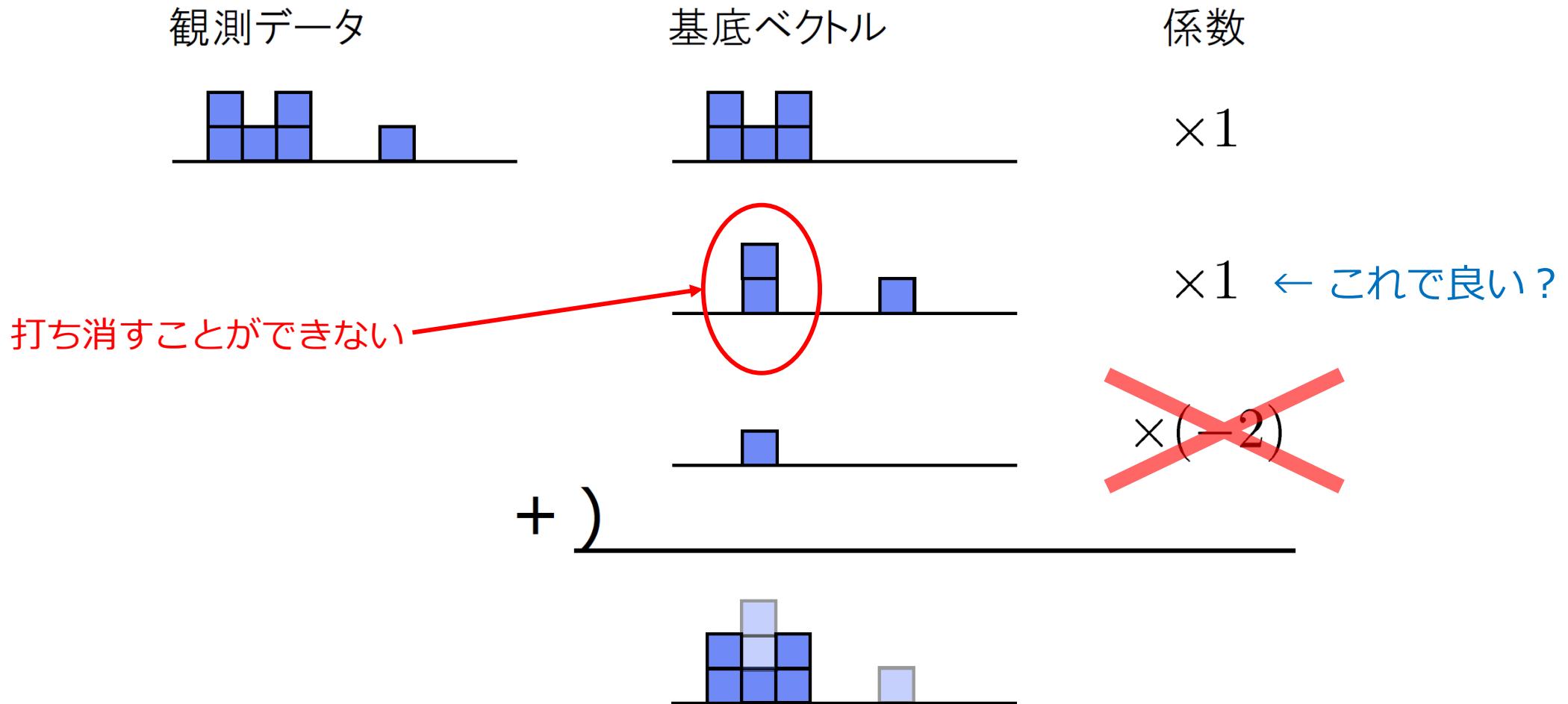
$\times (-2)$

+)



係数ベクトルのスパース性

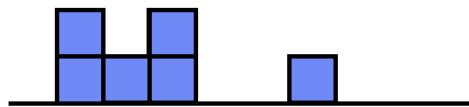
- 係数が負の値を取ってはならない場合



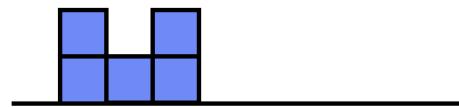
係数ベクトルのスパース性

- 係数が負の値を取ってはならない場合

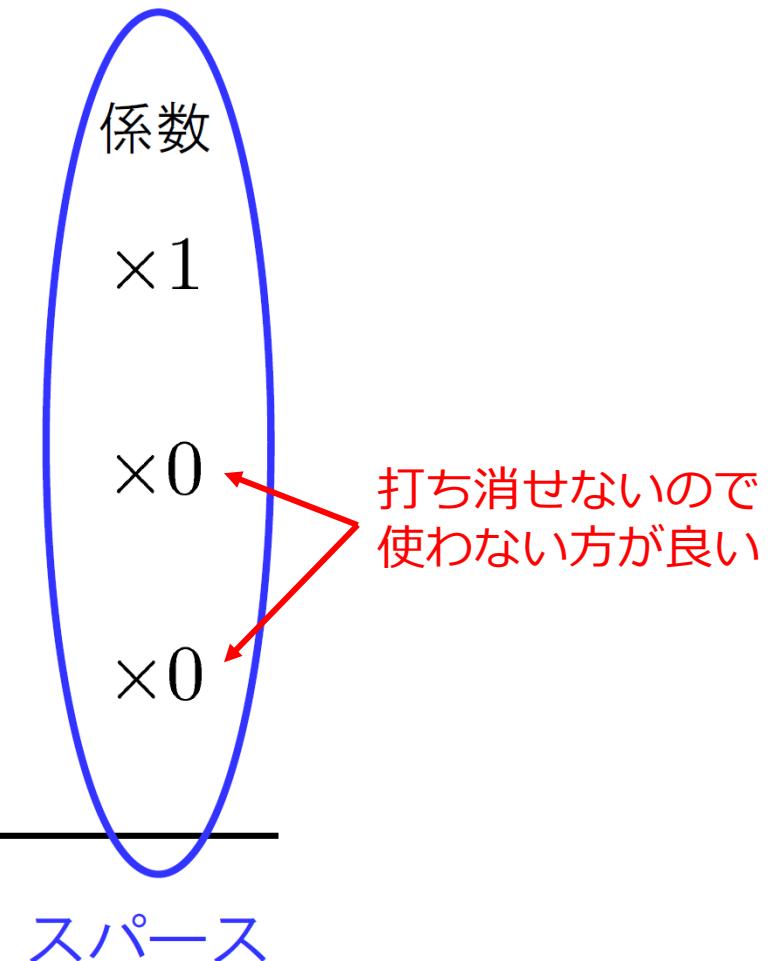
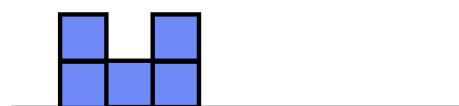
観測データ



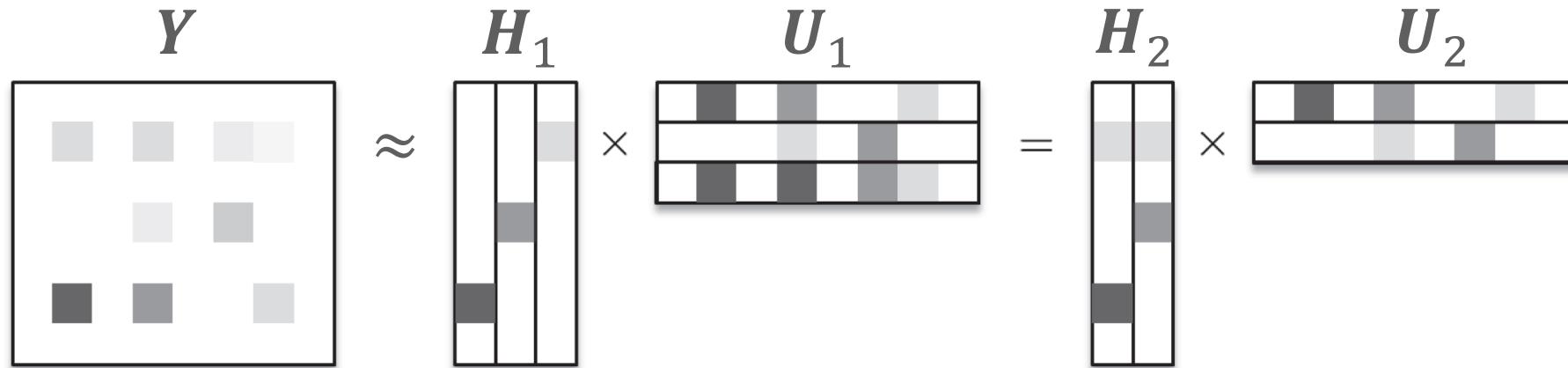
基底ベクトル



+)



共起変数が同じ基底にまとめられる効果



引用：亀岡(2012)

3本の基底ベクトルで Y を近似表現する場合を考える。

共起する変数 (1,3番目と1,2番目) を1つの基底にまとめた H_2 では、
まだ使っていない3本目の基底を用いて、さらに近似精度を高める余地がある。



NMF では、共起変数が同一の基底にまとめられる傾向がある。
(パーツが抽出される根拠とも捉えられる。)

H, U の推定における最適化問題

スカラー x, y の乖離度の基準として例えば…

- 二乗誤差

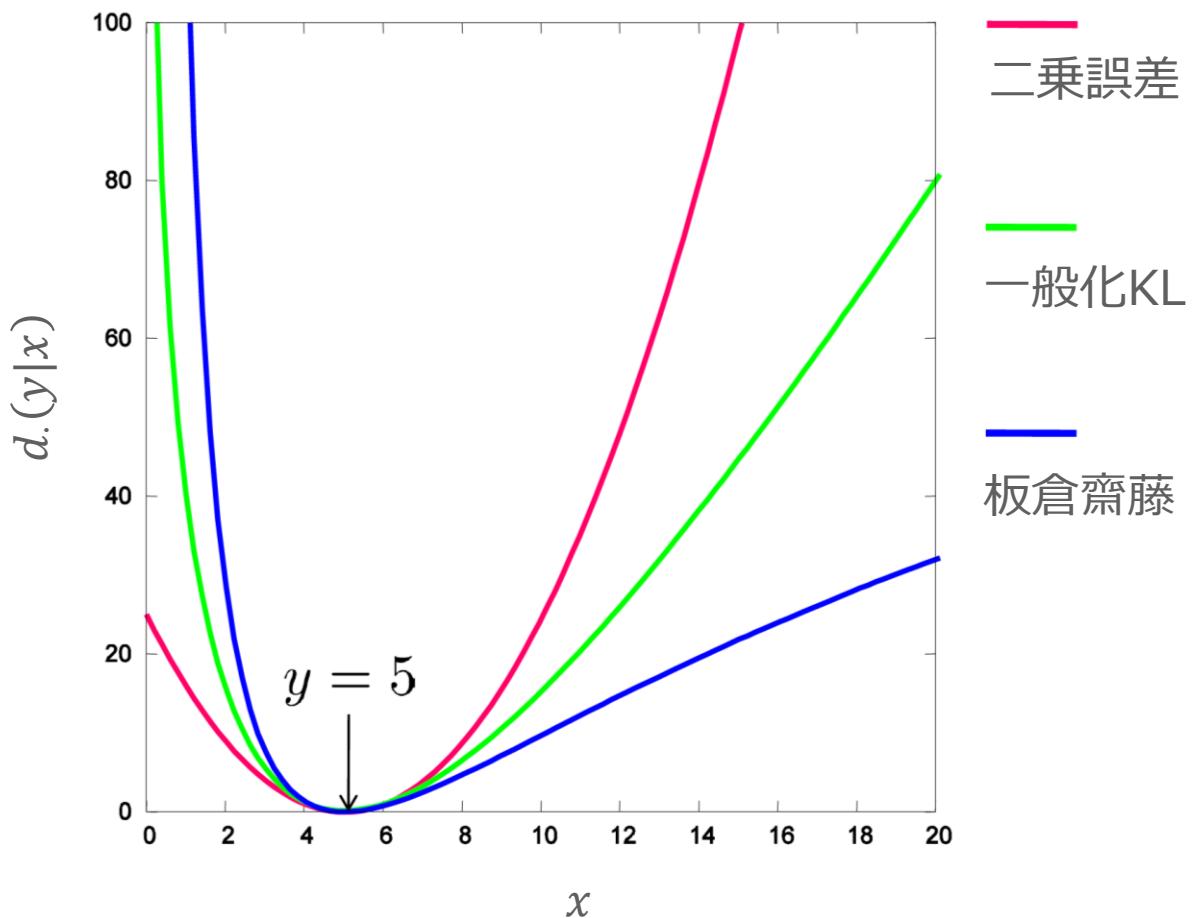
$$d_{SE}(y|x) = (y - x)^2$$

- I (一般化Kullback-Leibler) ダイバージェンス

$$d_{KL}(y|x) = y \log \frac{y}{x} - y + x$$

- 板倉齋藤ダイバージェンス

$$d_{IS}(y|x) = \frac{y}{x} - \log \frac{y}{x} - 1$$



H, U の推定における最適化問題

観測データ行列 Y と HU の乖離度の基準として、「要素単位で見た乖離度の総和」を考える。

$$D.(H, U) = \sum_{i,j} d.(y_{i,j} | \sum_k h_{i,k} u_{k,j}) = \sum_{i,j} d.(y_{i,j} | x_{i,j})$$

H, U の推定値を次の最適化問題を解くことで得る。

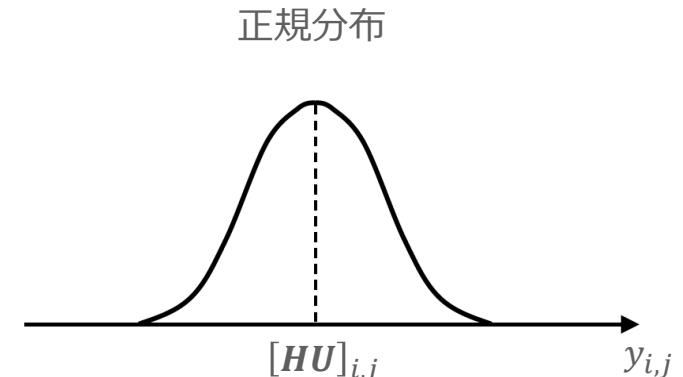
$$\begin{aligned} & \text{minimize} && D.(H, U) \\ & \text{subject to} && h_{i,k} \geq 0, u_{k,j} \geq 0 \end{aligned}$$

最尤推定問題としての解釈

それぞれの乖離度基準を用いた H, U の推定は、次の分布を仮定した最尤推定と等価。

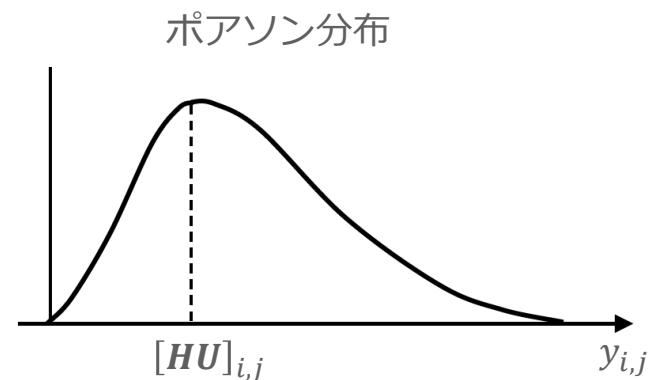
- 二乗誤差 d_{SE}

$$y_{i,j} \sim Normal ([HU]_{i,j}, \sigma^2)$$



- I (一般化Kullback-Leibler) ダイバージェンス d_{KL}

$$y_{i,j} \sim Poisson ([HU]_{i,j})$$



- 板倉齋藤ダイバージェンス d_{IS}

$$y_{ij} \sim Exponential ([HU]_{i,j})$$

推定アルゴリズムの設計

Lee (2001) で提案された推定アルゴリズムを導出する.

例として、二乗誤差 d_{SE} を用いた場合を扱う.

$$\text{minimize} \quad \sum_{i,j} \left(y_{i,j} - \left(\sum_k h_{i,k} u_{k,j} \right) \right)^2 \quad \text{subject to} \quad h_{i,k} \geq 0, \quad u_{k,j} \geq 0$$

複雑でそのまま扱うのは難しい.

→ 補助関数法を利用.

補助関数法

$$f(\theta) = \min_{\bar{\theta}} g(\theta, \bar{\theta})$$

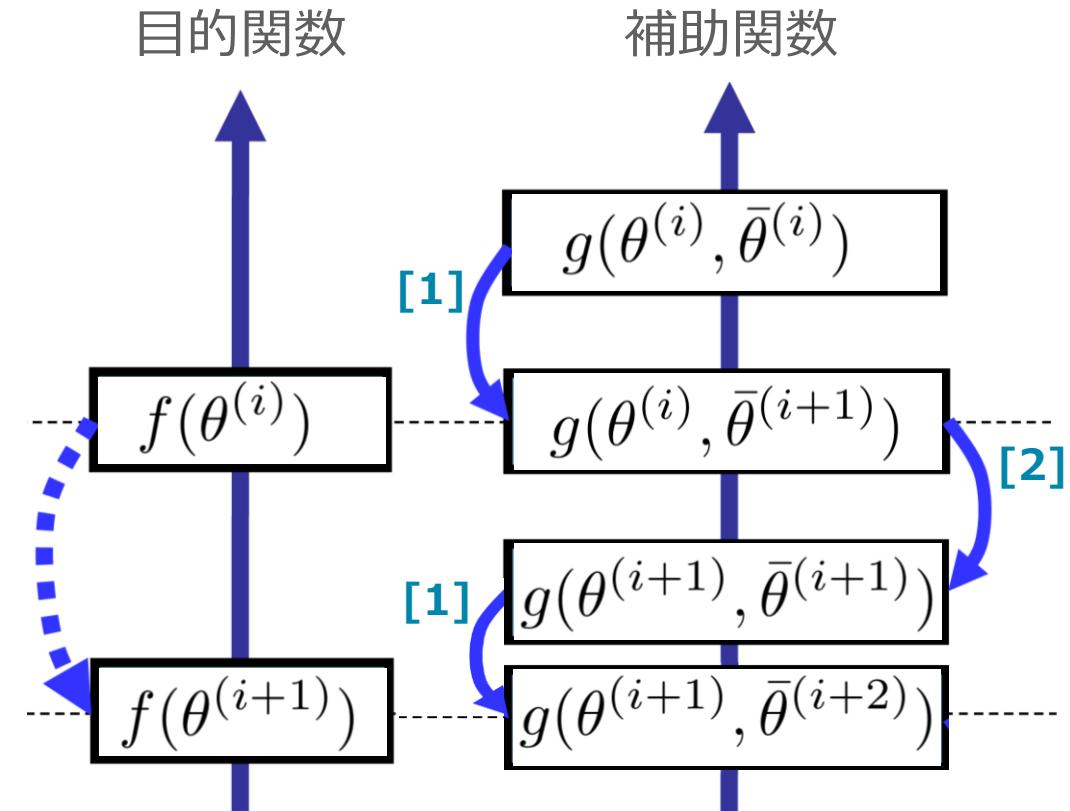
が成り立つとき, $g(\theta, \bar{\theta})$ を $f(\theta)$ の**補助関数**と呼ぶ.

以下の更新を繰り返すことで,

$f(\theta)$ を単調減少させる(θ の列を得る)ことができる.

$$[1] \quad \bar{\theta}^{(i+1)} = \operatorname{argmin}_{\bar{\theta}} g(\theta^{(i)}, \bar{\theta})$$

$$[2] \quad \theta^{(i+1)} = \operatorname{argmin}_{\theta} g(\theta, \bar{\theta}^{(i+1)})$$



補助関数の設計

二乗誤差基準(目的関数)

$$\sum_{i,j} \left(y_{i,j} - \sum_k h_{i,k} u_{k,j} \right)^2 = \sum_{i,j} \left\{ y_{i,j}^2 - 2y_{i,j} \sum_k h_{i,k} u_{k,j} + \left(\sum_k h_{i,k} u_{k,j} \right)^2 \right\}$$

$$\begin{aligned} \left(\sum_k h_{i,k} u_{k,j} \right)^2 &= \left(\sum_k \lambda_{i,k,j} \frac{h_{i,k} u_{k,j}}{\lambda_{i,k,j}} \right)^2 && [\lambda_{i,k,j} \geq 0, \sum_k \lambda_{i,k,j} = 1] \\ &\leq \sum_k \lambda_{i,k,j} \left(\frac{h_{i,k} u_{k,j}}{\lambda_{i,k,j}} \right)^2 && \because \text{Jensen の不等式} \\ &= \sum_k \frac{h_{i,k}^2 u_{k,j}^2}{\lambda_{i,k,j}} \end{aligned}$$



$$\sum_{i,j} \left(y_{i,j} - \sum_k h_{i,k} u_{k,j} \right)^2 \leq \sum_{i,j} \left\{ y_{i,j}^2 - 2y_{i,j} \sum_k h_{i,k} u_{k,j} + \sum_k \frac{h_{i,k}^2 u_{k,j}^2}{\lambda_{i,k,j}} \right\}$$

Jensen の不等式の等号成立条件

$$\lambda_{i,k,j} = \frac{h_{i,k} u_{k,j}}{\sum_{k'} h_{i,k'} u_{k',j}}$$

が成立する時、この等号も成立する。

したがって、

右辺は二乗誤差基準(目的関数)の
補助関数となっている。

h_{ik}, u_{kj} ごとの2次関数の和の形。

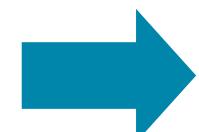
→この補助関数の最小化は容易。

推定アルゴリズム

H, U に関する最小化をしやすい**補助関数**が得られたので、前述の枠組みに当てはめる。

$$[1] \quad \lambda_{i,k,j} \leftarrow \frac{h_{i,k} u_{k,j}}{\sum_{k'} h_{i,k'} u_{k',j}}$$

λ を代入して



まとめる

$$[2.1] \quad h_{i,k} \leftarrow \frac{\sum_j y_{i,j} u_{k,j}}{\sum_j \frac{u_{k,j}^2}{\lambda_{i,k,j}}}$$

$$[2.2] \quad u_{k,j} \leftarrow \frac{\sum_i y_{i,j} h_{i,k}}{\sum_i \frac{h_{i,k}^2}{\lambda_{i,k,j}}}$$

Step.0 $h_{i,k}, u_{k,j}$ を**非負値**で初期化。

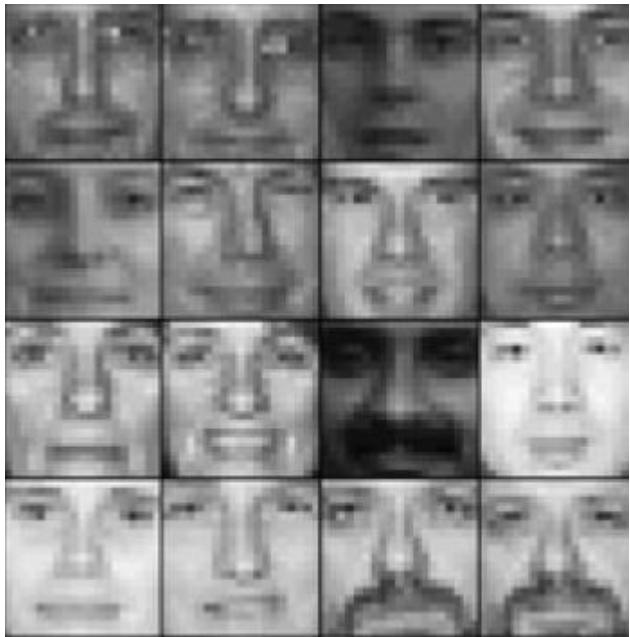
$$\text{Step.1} \quad h_{i,k} \leftarrow h_{i,k} \frac{\sum_j y_{i,j} u_{k,j}}{\sum_j u_{k,j} \sum_{k'} h_{i,k'} u_{k',j}}$$

$$\text{Step.2} \quad u_{k,j} \leftarrow u_{k,j} \frac{\sum_i y_{i,k} h_{i,k}}{\sum_i h_{i,k} \sum_{k'} h_{i,k'} u_{k',j}}$$

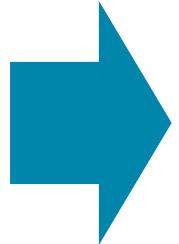
Step.3 収束判定。Step.1,2 の繰り返し。

※ 非負値制約は常に満たされ続ける。

顔画像データへの適用例 1 (Lee, 2001)



CBCL 顔画像データセット
 $n = 2429, p = 19 \times 19$



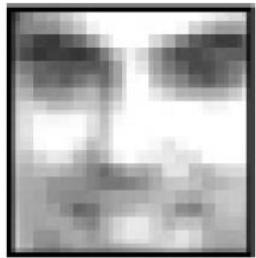
- ベクトル量子化 (VQ)
- 主成分分析 (PCA)
- 非負値行列因子分解 (NMF)

抽出する基底(クラスタ,主成分)の数は 49 に設定.

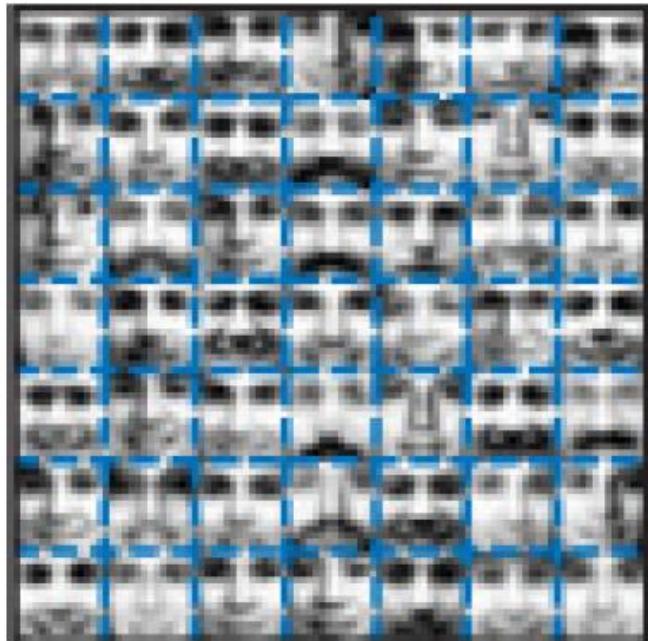
VQ では K-means 法が用いられ, NMF では一般化KLダイバージェンスが用いられた.

顔画像データへの適用例 1 (Lee, 2001)

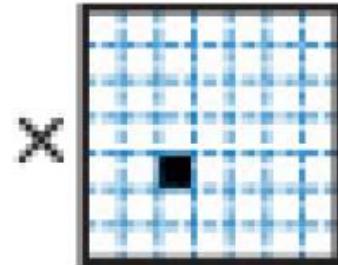
Original



VQ

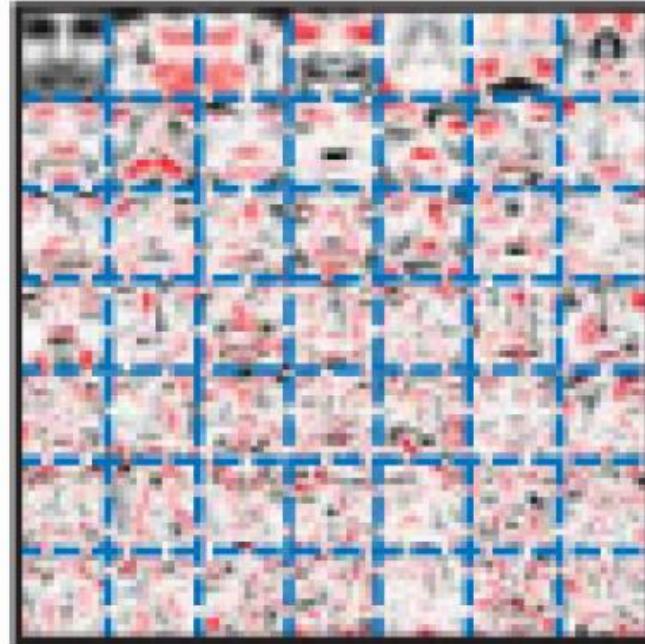


II

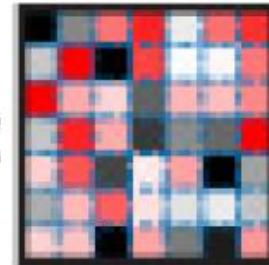


×

PCA



II



×

VQ では「典型的な顔」が基底(クラスター重心)として抽出されている.
PCAで抽出された基底(主成分)には負の値が含まれ、解釈しづらい。



「パーツ」の抽出は、
なされていない。

NMFの顔画像データへの適用例 1 (Lee, 2001)

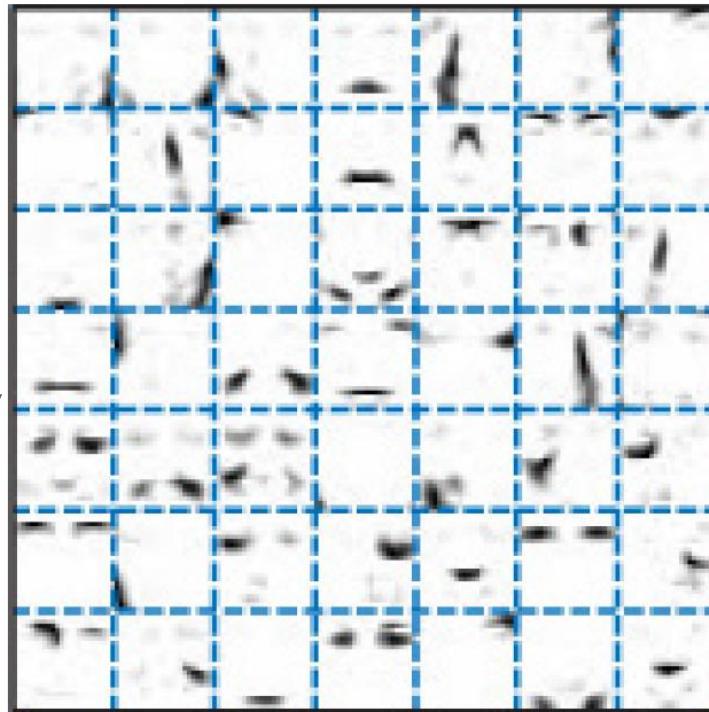
基底と係数への**非負値制約**によって、「**パーツ**」が抽出されている。

ハードクラスタリングであるVQに対し、NMFはソフトクラスタリングとしても捉えられる。

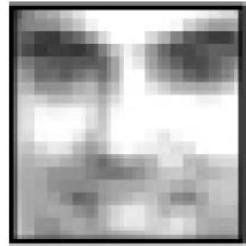
(今回は触れないが、論文中では文書クラスタリングへの適用例も示されている。)

引用 : Lee(1999)

NMF



Original



$$\text{Original} \times \begin{matrix} \text{NMF Basis} \\ \text{Matrix} \end{matrix} = \text{Reconstructed Face}$$

VQ

$$\text{Original} \times \begin{matrix} \text{VQ Basis} \\ \text{Matrix} \end{matrix} = \text{Reconstructed Face}$$

PCA

$$\text{Original} \times \begin{matrix} \text{PCA Basis} \\ \text{Matrix} \end{matrix} = \text{Reconstructed Face}$$

NMFの顔画像データへの適用例 2 (Hoyer, 2004)

ORL 顔画像データセット

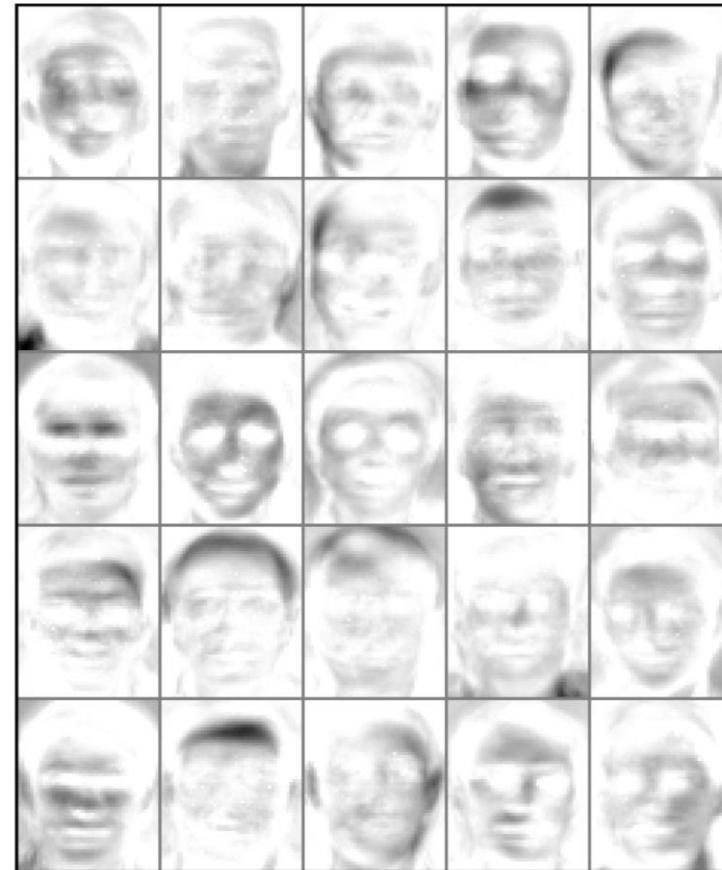
$n = 400$ (40人×角度等の違う10枚)



NMF



基底数 : 25

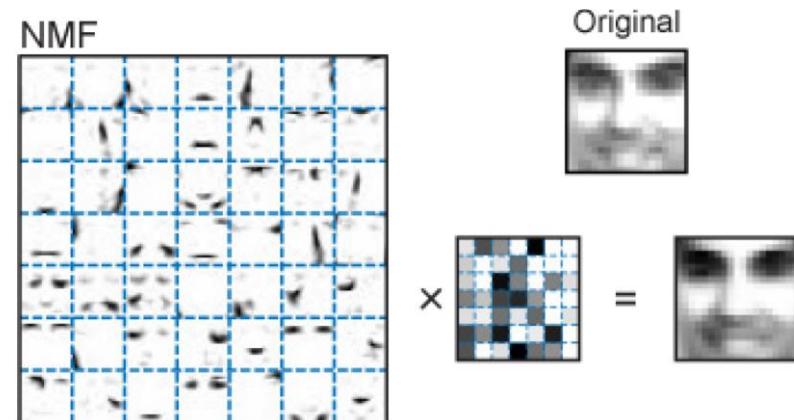


パーツを上手く抽出できず、各基底が顔全体を表現してしまっている。

内容1. いくつかのNMFの紹介

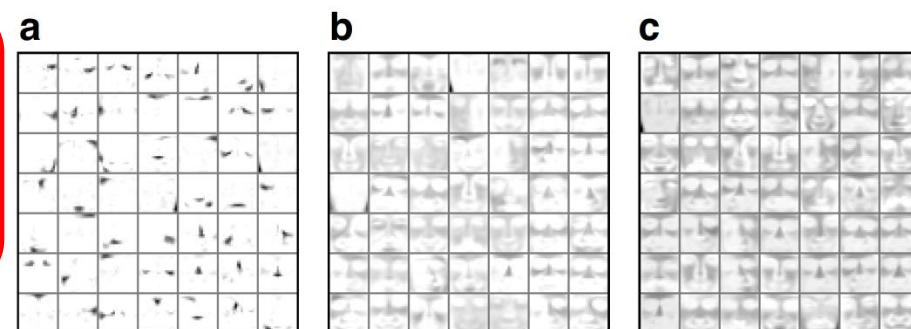
■ Non-negative Matrix Factorization (NMF)

Lee(1999), Lee(2001)



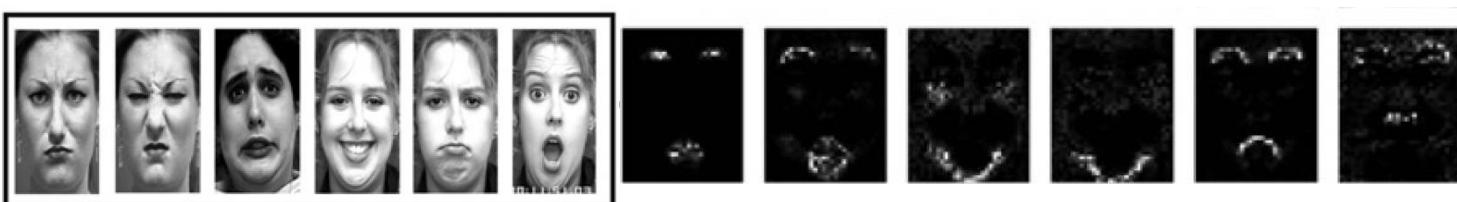
■ NMF with Sparseness Constraints (NMFS)

Hoyer(2004)



■ Discriminant NMF (DNMF)

Zafeiriou(2006), Kotsia(2008)



■ 新規手法の検討

NMF with Sparseness Constraints のモチベーション

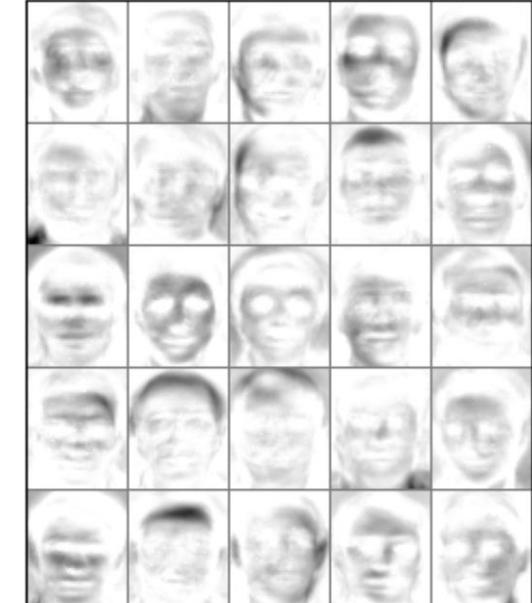
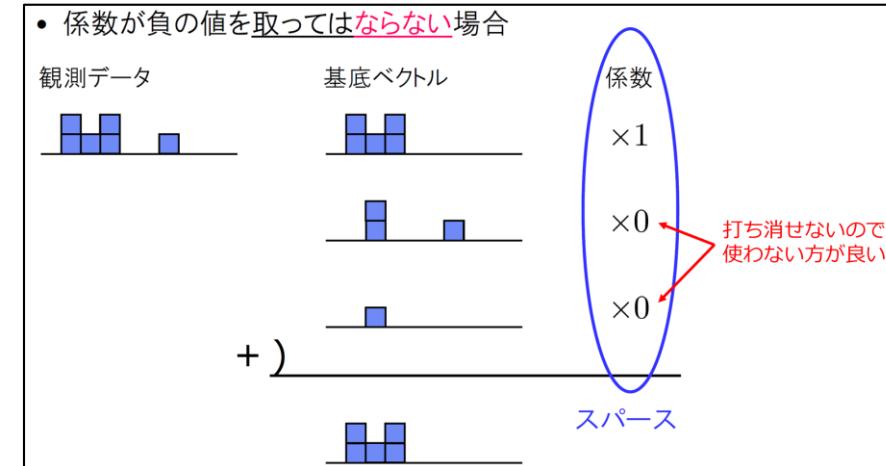
NMFでは、最適化問題にスパース性を促すような明示的な記述は無いが、
非負値制約によって係数ベクトルがスパースになる傾向がある。

Lee(1999)では "Naturally Sparse" と言われている。

→ 明示的にスパース性を促す記述を組み込み、
係数ベクトルのスパース性の度合いをコントロールしたい。

うまくパーツに別れて基底ベクトルもスパースになることがあるが、
データによってはそうならない場合もある。

→ 基底ベクトルのスパース性を促す記述も組みみたい。



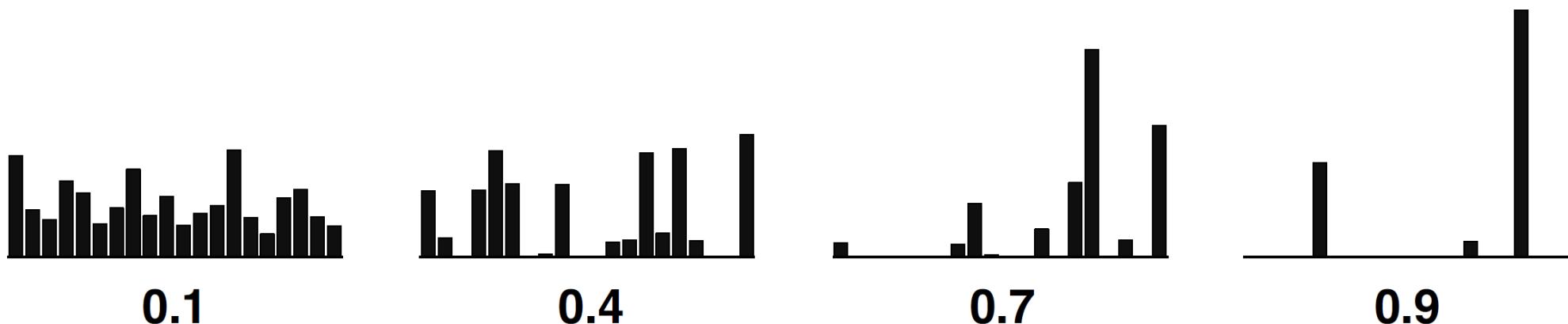
引用 : Hoyer(2004)

スパースネスの指標

n 次元ベクトルの x のスパースネス(スパース性の程度)を以下で定義する.

$$\text{sparseness}(\boldsymbol{x}) = \frac{\sqrt{n} - \frac{\sum |x_i|}{\sqrt{\sum x_i^2}}}{\sqrt{n} - 1}$$

0に近いほど非スパース, 1に近いほどスパース.



引用 : Hoyer(2004)

NMF with Sparseness Constraints の最適化問題

モデルや乖離度基準は NMF と同じだが、最適化問題にスパースネスに関する制約を課す。

$$\begin{aligned} & \text{minimize} && D_{SE}(\mathbf{H}, \mathbf{U}) \\ & \text{subject to} && h_{i,k} \geq 0, u_{k,j} \geq 0, \\ & && \text{sparseness}(\mathbf{h}_l) = S_h, \\ & && \text{sparseness}(\mathbf{u}_m) = S_u. \end{aligned}$$

※ 片方のみのスパースネス制約を課すことも可能

\mathbf{h}_l : \mathbf{H} の第 l 列ベクトル

\mathbf{u}_m : \mathbf{U} の第 m 行ベクトル

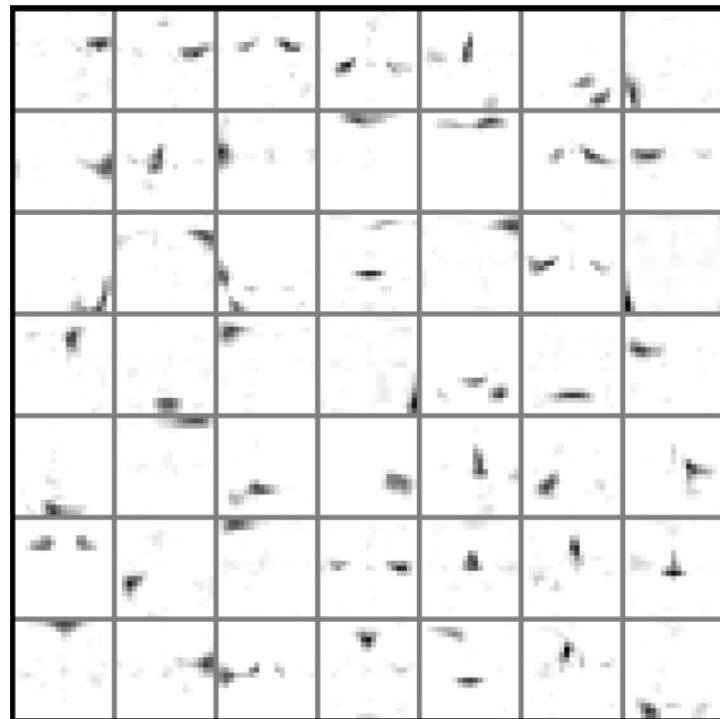
S_h, S_u はチューニングパラメータであり、この値を通して基底と係数のスパースネスを調整する。

推定には射影勾配アルゴリズムが用いられるが、ここでは省略する。

NMF-Sの顔画像データへの適用例

前述のORL顔データセットに、 NMF with Sparse Constraints を適用。スパースネス値 S_h, S_u を変化させている。

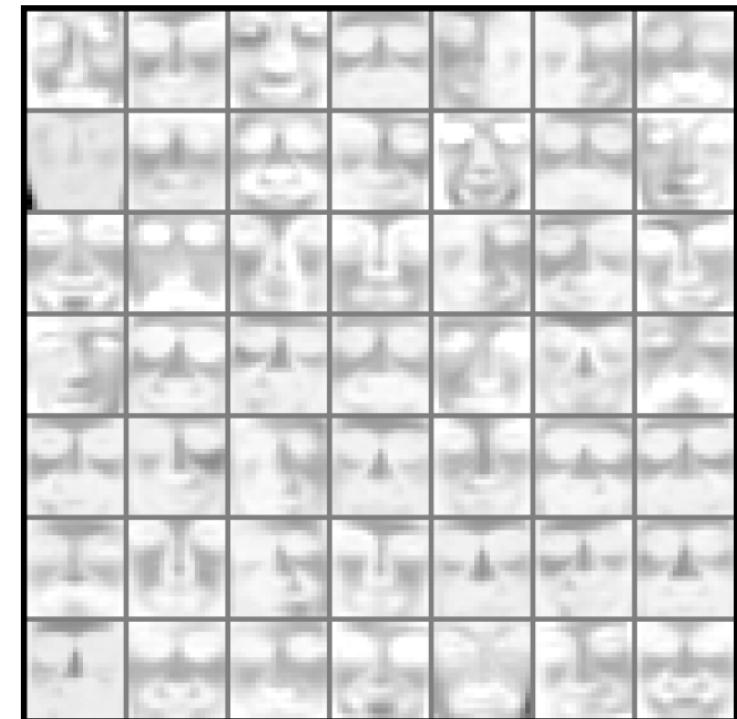
$$S_h = 0.8$$



$$S_u = 0.8$$



$$S_h = 0.2$$



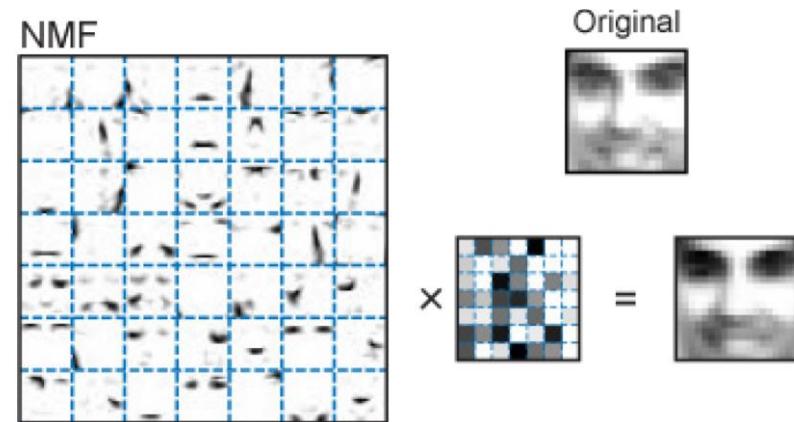
基底ベクトルに対して高いスパースネスを設定したとき、うまくパートを抽出できている。
(通常のNMFでは基底が全体を表してしまっていた。)

引用 : Hoyer(2004)

内容1. いくつかのNMFの紹介

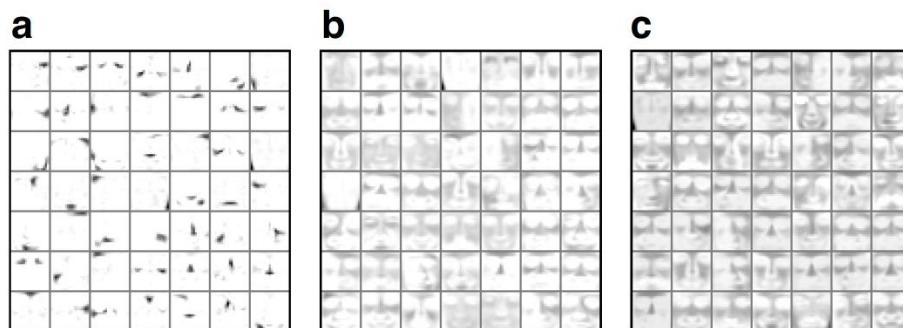
■ Non-negative Matrix Factorization (NMF)

Lee(1999), Lee(2001)



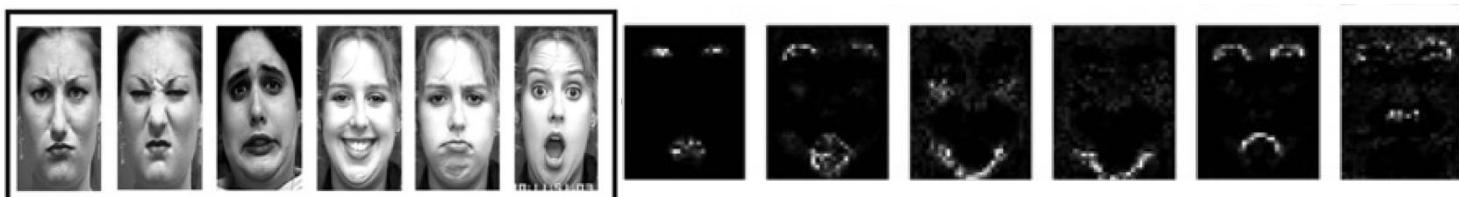
■ NMF with Sparseness Constraints (SNMF)

Hoyer(2004)



■ Discriminant NMF (DNMF)

Zafeiriou(2006), Kotsia(2008)



■ 新規手法の検討

Discriminant NMF (DNMF) のモチベーション

クラスラベルの付いたデータを扱う.
例：表情（怒り，悲しみ，喜び，驚き…）



Kotsia(2008) で扱われたデータの一部

- クラスの情報を付加的に使って、パート抽出の精度を高めたい。
- クラス間の違いをパートで表現して、解釈や予測に役立てたい。

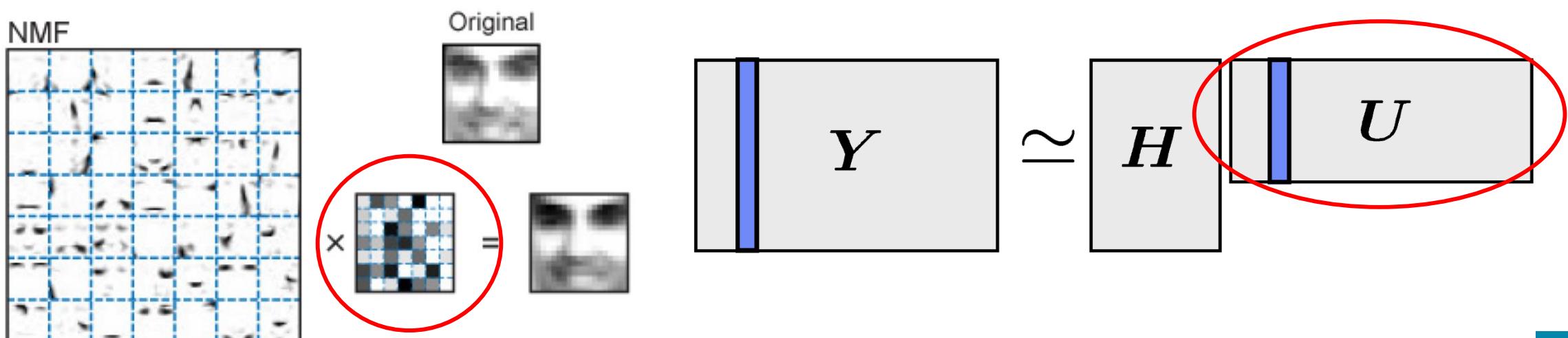
Fisher LDA と同様の考え方に基づき、クラス間変動とクラス内変動を用いる。

係数ベクトルのクラス間分散とクラス内分散

- 異なるクラスでは「基底(パーツ)の使い方」「どの基底を使うことが多いか」が異なるはず。
例：笑顔では細い目，驚いた顔では見開いた目が表れやすい。
- クラス間で異なる使われ方をするような基底(パーツ)を抽出したい。特に判別のため。



係数ベクトルのクラス間分散＆クラス内分散を見る。



引用 : Lee(1999)

Discriminant NMF の最適化問題

$$\text{minimize} \quad D_{KL}(\mathbf{H}, \mathbf{U}) + \gamma \operatorname{tr}[\mathbf{S}_w] - \delta \operatorname{tr}[\mathbf{S}_b]$$

$$\text{subject to} \quad h_{i,k} \geq 0, \quad u_{k,j} \geq 0$$

\mathbf{S}_w : 係数ベクトル \mathbf{u} のクラス内分散行列

\mathbf{S}_b : 係数ベクトル \mathbf{u} のクラス間分散行列

γ, δ は正のチューニングパラメータである.

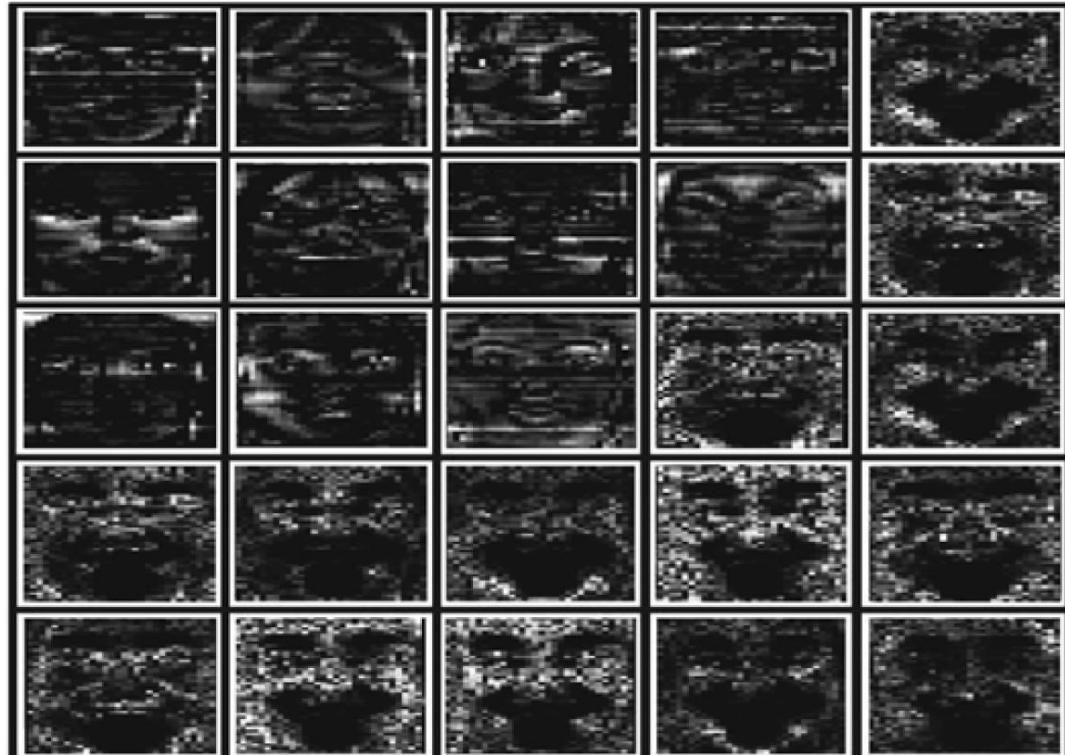
推定には補助関数法によるアルゴリズムが用いられるが、ここでは省略する.

DNMF の顔画像データへの適用例 (Kotsia, 2008)

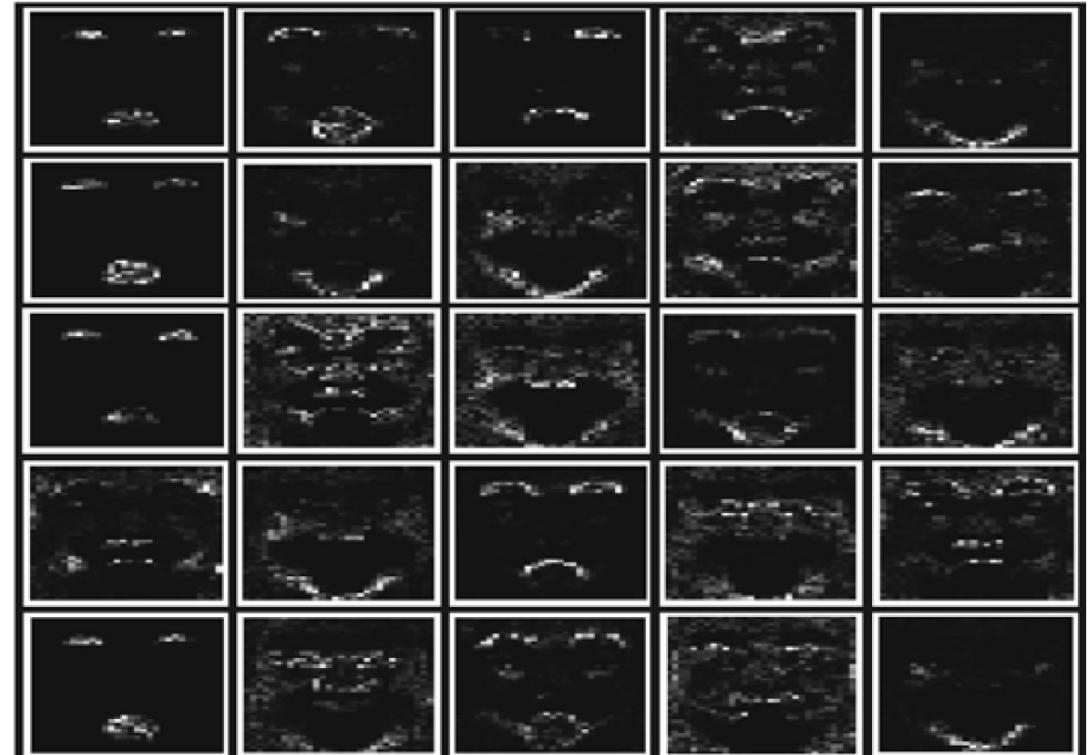


{怒り, 嫌悪, 恐怖, 幸福, 悲しみ, 驚き}
の6クラスのラベルが付いた顔画像データ.

NMF



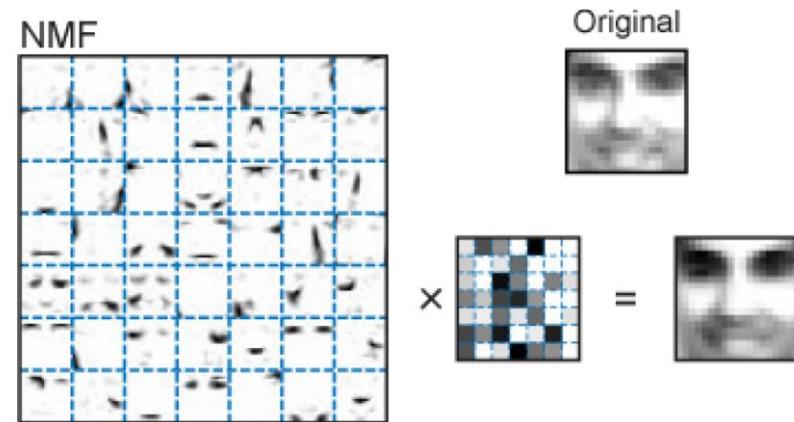
DNMF



内容1. いくつかのNMFの紹介

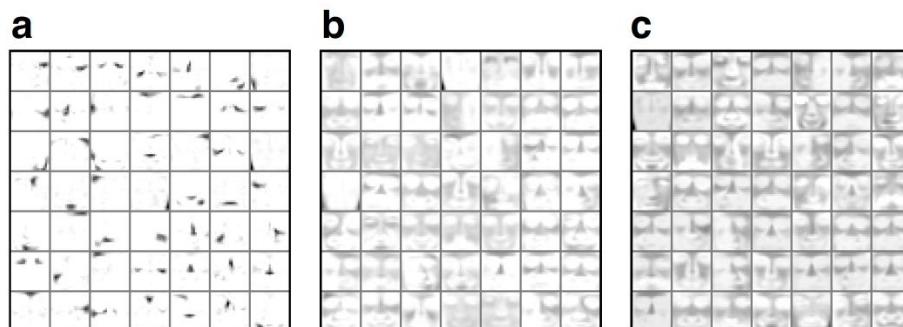
■ Non-negative Matrix Factorization (NMF)

Lee(1999), Lee(2001)



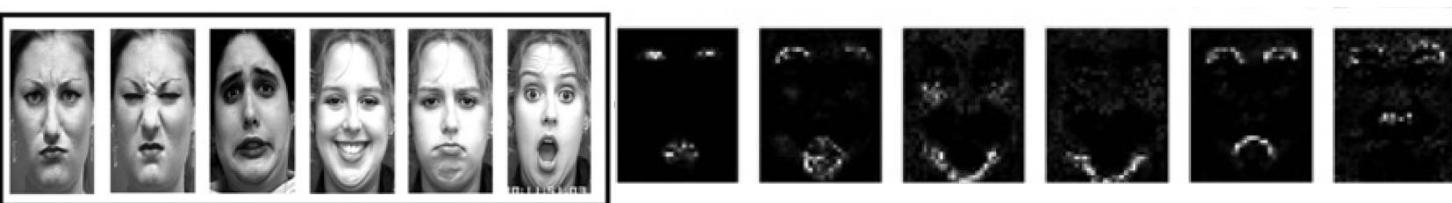
■ NMF with Sparseness Constraints (SNMF)

Hoyer(2004)



■ Discriminant NMF (DNMF)

Zafeiriou(2006), Kotsia(2008)



■ 新規手法の検討

新規手法の検討 NMFS + DNMF

DNMFのようにクラス情報を使う場合においても，

基底や係数のスパースネスの調整ができれば，有益であろう.

DNMFとNMFSを愚直に組み合わせて，例えば次の最適化問題を考えられる.

$$\begin{aligned} & \text{minimize} && D.(\mathbf{H}, \mathbf{U}) + \gamma \operatorname{tr}[\mathbf{S}_w] - \delta \operatorname{tr}[\mathbf{S}_b] \\ & \text{subject to} && h_{i,k} \geq 0, u_{k,j} \geq 0, \\ & && \text{sparseness}(\mathbf{h}_l) = S_h, \\ & && \text{sparseness}(\mathbf{u}_m) = S_u. \end{aligned}$$

今後，アルゴリズムの設計や最適化問題の改変などを検討している. (その前に先行研究調査)

内容2. フェンシング画像データへの適用の検討



PoseNetによる関節推定

スポーツデータ解析コンペでは、
フェンシングの映像データが扱われている。
「得点した(突いた)瞬間」の画像に対する
NMFなどの適用を検討していく。



引用：スポーツデータ解析コンペティション発会式チュートリアル

フェンシング エペ とは



引用 : Yuki Ota Fencing Visualized Project – MORE ENJOY FENCING

様々な得点パターン (突きのパターン)

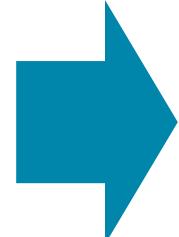


NMF, NMFS の適用イメージ

動画から抽出した
得点(突き)の瞬間の画像



前処理



アングル等の
画像内の人配置を
統一化

見たい選手以外の
オブジェクトを
除去

PoseNetを用いて
関節だけの
棒人間画像に変換

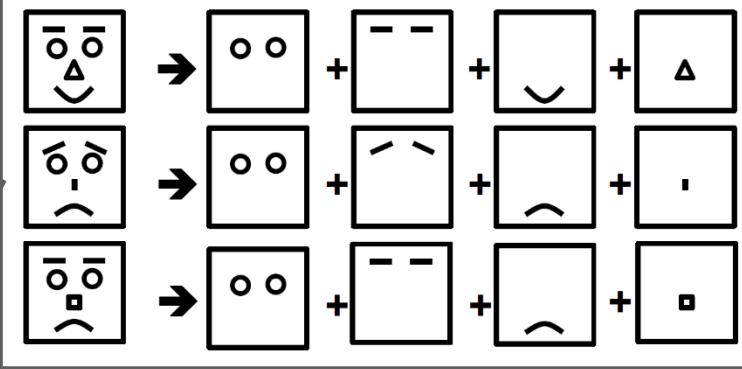
類推

NMF
NMFS
(VQ)



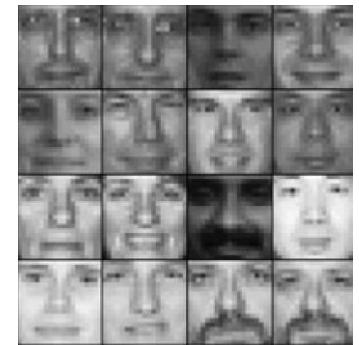
脚,腕,上体などの
パートを抽出
できるか?

- ・個人の得失点パターンの抽出
- ・選手タイプの分類
- などに繋がる可能性。



前処理の例 カメラのパンによる歪みを補正

パノラマ画像を生成し各フレームをマッチングさせることができれば、
パンによる歪みを補正できる。
→ 同じアングルからの「得失点(突き)シーン画像」が得られる。



CBCL顔データセットの一部

補正前



補正後



引用：スポーツデータ解析コンペティション発会式チュートリアル

前処理の例 PoseNet による関節推定 (棒人間化)

PoseNet を用いて関節を推定できれば、

関節とそれを結ぶ線(棒人間)だけの画像に変換する事も可能。

→ よりシンプルな画像にNMF等を適用することで、解釈性が高まるかもしれない。

得点シーン



関節の推定結果



引用：スポーツデータ解析コンペティション発会式チュートリアル

DMNF の適用イメージ / エペ特有の「同時得点」

同時得点



左の選手だけ得点



引用 : Yuki Ota Fencing Visualized Project – MORE ENJOY FENCING

「同時得点を避けて、いかに自分だけが得点できるか」が勝敗を分ける. (by試合実況)



「同時得点」と「自分だけ得点」の 2 クラスのラベルを画像に割り当て、DNMFを適用したい.
→ 自分だけ得点できるのはどういう姿勢の時か、体のパーソベースで解釈できる可能性.

今後の課題

- NMFS と DNF を組み合わせた新規手法の検討について
 - ・関連手法の調査
 - ・アルゴリズムの設計とそれに伴う最適化問題の改変
- フェンシング画像データへの NMF の適用について
 - ・得点シーンの抽出をはじめとした前処理
 - ・得点(突き)直前の動作も含めた動画解析

参考文献

- D. Lee and H. Seung, Learning the parts of objects by non-negative matrix factorization, Nature, 401, (1999).
- D. Lee and H. Seung, Algorithms for non-negative matrix factorization, Advances in Neural Information Processing 13, MIT Press, (2001).
- P. Hoyer, Non-negative Matrix Factorization with Sparseness Constraints, Journal of Machine Learning Research, 5, (2004).
- S. Zafeiriou, A. Tefas, I. Buciu, I. Pitas, Exploiting Discriminant Information in Nonnegative Matrix Factorization With Application to Frontal Face Verification, IEEE Transactions on Neural Networks, 17, (2006).
- Kotsia, I. Buciu, I. Pitas, An analysis of facial expression recognition under partial facial image occlusion, Image and Vision Computing, 26, (2008).
- 亀岡 弘和, 非負値行列因子分解, 計測と制御, 51, (2012).