Blind Source Separation

Václav Šmídl

April 7, 2025

Recapitulation

1. Model-based development:
   - modelling assumptions $\Longrightarrow$ estimation method
   - Gaussian Mixture $\Longrightarrow$ GMM method (via the EM algorithm)
   - Linear combination with Gaussian noise $\Longrightarrow$ OLS
   - mixture of linear models $\Longrightarrow$ custom algorithm

Recapitulation

1. Model-based development:
   - modelling assumptions $\Longrightarrow$ estimation method
   - Gaussian Mixture $\Longrightarrow$ GMM method (via the EM algorithm)
   - Linear combination with Gaussian noise $\Longrightarrow$ OLS
   - mixture of linear models $\Longrightarrow$ custom algorithm
2. Method-based development
   - PCA $\Longrightarrow$ FA $\Longrightarrow$ BSS

Matrix Least Squares

Consider two least squares problem

$$y_1 = X\theta_1 \qquad y_2 = X\theta_2 \qquad Y = X\Theta$$

with solutions

Matrix Least Squares

Consider two least squares problem

$$y_1 = X\theta_1 \qquad y_2 = X\theta_2 \qquad Y = X\Theta$$

with solutions

$$\hat{\theta}_1 = (X^T X)^{-1} X^T y_1$$
$$\hat{\theta}_2 = (X^T X)^{-1} X^T y_2$$
$$\hat{\Theta} = (X^T X)^{-1} X^T Y$$

Matrix Least Squares

Consider two least squares problem

$$y_1 = X\theta_1 \qquad y_2 = X\theta_2 \qquad Y = X\Theta$$

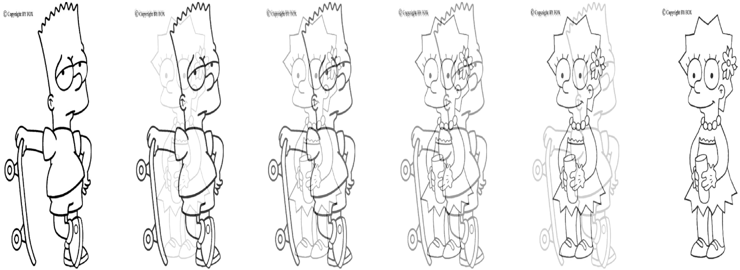
with solutions

$$\hat{\theta}_1 = (X^T X)^{-1} X^T y_1$$
$$\hat{\theta}_2 = (X^T X)^{-1} X^T y_2$$
$$\hat{\Theta} = (X^T X)^{-1} X^T Y$$

In probabilities, matrix Normal distribution

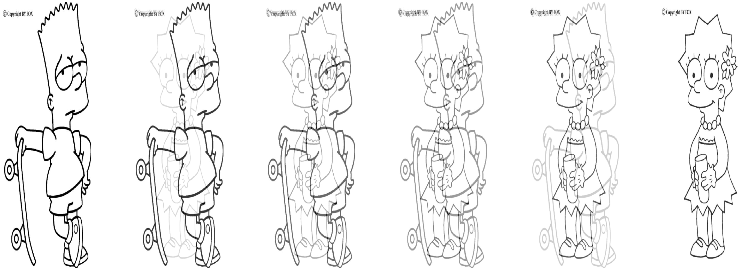
$$\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix}, \begin{bmatrix} (X^T X)^{-1} & 0 \\ 0 & (X^T X)^{-1} \end{bmatrix} \right) <=> \Theta = \mathcal{N}\left(\hat{\Theta}, I \otimes (X^T X)^{-1} \right)$$

Blind source separation – image sequence



- 2 sources

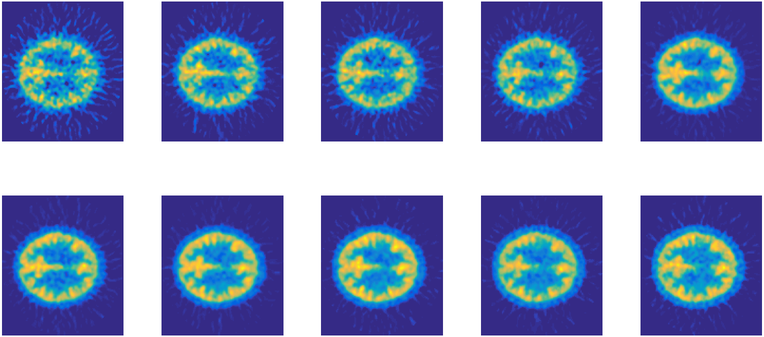
Blind source separation – image sequence



- 2 sources
- weights:

Bart	1	0.8	0.6	0.4	0.2	0
Lisa	0	0.2	0.4	0.6	0.8	1

Medical imaging



- number of sources?
- source images
- time activity of the source

Mathematical description

Linear model of $p \times n$ matrix

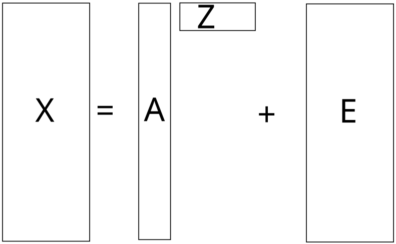
$$x_i = Az_i + e_i, i = 1..n$$

where

A is the $p \times r$ matrix of source images, $r < \min(n, p)$

Z is the $n \times r$ matrix of time activities,

E is the $p \times n$ noise matrix (Gaussian)



The case of known Z: least squares

Consider decomposition $X = AZ^\top$

$$\underline{x}_j = \underline{a}_j Z^\top \quad \Longrightarrow \quad \underline{x}_j^\top = Z \underline{a}_j^\top,$$

for known D, X the estimate of A is

$$\hat{\underline{a}}_j^\top = (Z^\top Z)^{-1} Z^\top \underline{x}_j^\top, \quad \Sigma_{a_j} = (Z^\top Z)^{-1},$$

The case of known Z: least squares

Consider decomposition $X = AZ^\top$

$$\underline{x}_j = \underline{a}_j Z^\top \quad \Longrightarrow \quad \underline{x}_j^\top = Z\underline{a}_j^\top,$$

for known D, X the estimate of A is

$$\underline{\hat{a}}_j^\top = (Z^\top Z)^{-1} Z^\top \underline{x}_j^\top, \quad \Sigma_{a_j} = (Z^\top Z)^{-1},$$

In matrices

$$\hat{A} = XZ(Z^T Z)^{-1}, \quad p(A) = \mathcal{N}(\hat{Z}, I, \Sigma_{a_j}) = \prod_j \mathcal{N}(\underline{\hat{a}}_j, \Sigma_{a_j}),$$

where $\mathcal{N}(\hat{A}, I, \Sigma_{a_j})$, is *matrix normal distribution*, with row and column covariance matrices.
Useful expectation

$$\mathrm{E}(A^T A) = \sum_{j=1}^{p} \underline{a}_j \underline{a}_j^T = \sum_{j=1}^{p} \left(\underline{\hat{a}}_j \underline{\hat{a}}_j^T + \Sigma_{a_j} \right) = \hat{A}^T \hat{A} + p\Sigma_{a_j}$$

Source Separation

Model

$$X = AZ^T + E,$$

Both images A and time activities X are **unknown**!

Source Separation

Model

$$X = AZ^T + E,$$

Both images A and time activities X are **unknown**!

Issue: ambiguity

$$AZ^T = ATT^{-1}Z^T$$
$$(AT)(T^{-1}Z^T) = \overline{AZ}^T$$

matrix T includes both scaling and rotation.

Solution: additional constraints.

Principal component analysis

Consider n, p-dimensional vectors $x_i, i = 1, \dots n$, and their covariance matrix

$$S = \frac{1}{n} \sum (x_i - \overline{x})(x_i - \overline{x})^T.$$

Then r-dimensional vectors z_i,

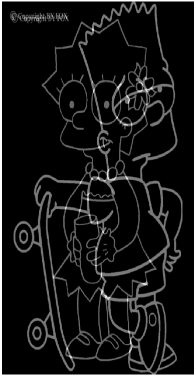
$$z_i = U(:, 1:r)^T x_i, \qquad S = U \Lambda U^T,$$

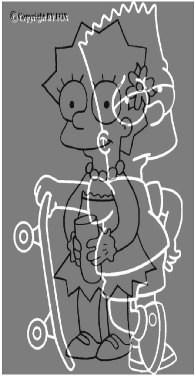
has maximum variance from all possible projections to r dimensions. U are eigenvectors of S sorted with decreasing eigenvalue.

Matrix SVD approach:

$$\begin{aligned} X &\stackrel{\text{svd}}{=} USV, \\ &= \underbrace{(US)}_{\hat{A}} \underbrace{(V)}_{\hat{Z}^T} \end{aligned}$$

Application to sequences of images





- Popular decades ago for speed of computation,
- Requires to find the rotation matrix T,
- Can we do better?
  - Independent Component Analysis (higher order moments)?
  - Structural priors
  - Non-negative matrix factorization

Probabilistic PCA [Tipping, Bishop, 1999]

Consider model:

$$p(x_i|A, z_i, \sigma) = \mathcal{N}(Az_i, \sigma I_p), \qquad p(z_i) = \mathcal{N}(0, I_r),$$

Marginalization of $p(x_i, z_i|A, \sigma)$ over z_i (lesson 1) yields

Probabilistic PCA [Tipping, Bishop, 1999]

Consider model:

$$p(x_i|A, z_i, \sigma) = \mathcal{N}(Az_i, \sigma I_p), \qquad p(z_i) = \mathcal{N}(0, I_r),$$

Marginalization of $p(x_i, z_i|A, \sigma)$ over z_i (lesson 1) yields

$$p(x_i|A, \sigma) = \mathcal{N}(0, C), \qquad p(X|A, \sigma) = \prod_i \mathcal{N}(0, C)$$

$$C = AA^T + \sigma I_p,$$

The likelihood is

$$p(X|A, \sigma) \propto \prod_i |C|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} x_i^\top C^{-1} x_i\right) \propto |C|^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \mathrm{tr}(C^{-1} \sum_i x_i x_i^\top)\right)$$

$$\propto |C|^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \mathrm{tr}(C^{-1} X X^\top)\right)$$

Maximum Likelihood

Maximum likelihood

$$p(X|A,\sigma) \propto \left|AA^T + \sigma I_p\right|^{-\frac{n}{2}} \exp\left(-\frac{1}{2}\mathrm{tr}((AA^T + \sigma I_p)^{-1}XX^\top)\right)$$

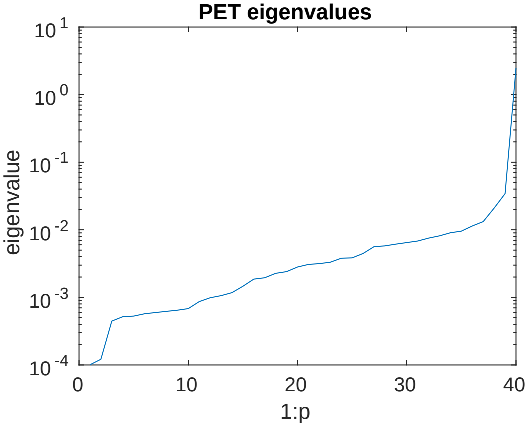
for $\hat{A}$and $\hat{\sigma}$ for given r:

$$\hat{A} = U_{1:r}(\Lambda_{1:r} - \hat{\sigma} I)^{\frac{1}{2}}, \qquad \hat{\sigma} = \frac{1}{d-r}\sum_{i=r+1}^{n} \lambda_i,$$

where $S = U\Lambda U^T$ is eigen-decomposition of $S = \sum x_i x_i^T$.

- distinction from PCA: subtraction of the noise

Image Sequence (PET)



- no plateau

Alternative: EM algorithm

Maximize

$$p(Z|X) \propto \int p(X|A,Z)p(A)dA$$

where

$$p(A) = \mathcal{N}(0, I, I) \propto \exp\left(-\frac{1}{2}\text{tr}\ (AI_rA^TI_p)\right),$$

Joint model

$$\log p(A, X|Z) = -\frac{1}{2}\sigma^{-1}\text{tr}(X - AZ^T)^T(X - AZ^T)$$
$$- \frac{1}{2}\text{tr}\ (AI_rA^TI_p)$$

Conditional $p(A|X, Z, \sigma)$ for known Z and σ

$$p(A|Z, X, \sigma) = \mathcal{N}(\hat{A}, I, \Sigma_A)$$
$$\hat{A} = XZ(Z^TZ + \sigma I)^{-1}, \qquad \Sigma_A = (\sigma^{-1}Z^TZ + I)^{-1}$$

The EM algorithm

Standard form for

E-step (over A): $q(X|X^{(j)}) = \int \log p(X, A|Z) p(A|Z^{(j)}, X) dA$

M-step (of Z): $Z^{(j+1)} = \arg\max_\theta q(Z|Z^{(j)})$

The $q(Z|Z^{(j)})$ factor is:

$$
\begin{aligned}
\log p(A, Z, X) &= -\frac{1}{2}\sigma^{-1}\mathrm{tr}(X - AZ^T)^T(X - AZ^T) \ -\frac{1}{2}\mathrm{tr}\ (AI_rA^TI_p) \\
&= -\frac{1}{2}\sigma^{-1}\mathrm{tr}\left(X^TX - ZA^TX - DAZ^T + ZA^TAZ^T + A^TA\sigma\right) \\
q(Z|Z^{(j)}) &= \mathrm{E}_A\left\{-\frac{1}{2}\sigma^{-1}\mathrm{tr}\left(X^TX - ZA^TX - DAZ^T + ZA^TAZ^T\right)\right\} \\
&= -\frac{1}{2}\sigma^{-1}\mathrm{tr}\left(X^TX - Z\hat{A}^TX - X^T\hat{A}Z^T + Z(\hat{A}^T\hat{A} + p\Sigma_A)Z^T\right) \\
&\propto \mathcal{N}(\hat{Z}, \Sigma_Z), \\
\hat{Z} &= X^T\hat{A}(\hat{A}^T\hat{A} + p\Sigma_A)^{-1},
\end{aligned}
$$

Summary of the EM Algorithm

Iterate two least squares problems:

1. $\hat{A} = XZ(Z^T Z + \sigma I)^{-1}, \Sigma_A = (\sigma^{-1} Z^T Z + I)^{-1}$
2. $\hat{Z} = X^T \hat{A}(\hat{A}^T \hat{A} + p\Sigma_A)^{-1}$

If you initialize at orthogonal solution $\hat{A} = U_{1:r}\Lambda_{1:r}$ it will be operating only on λs.

Summary of the EM Algorithm

Iterate two least squares problems:

1. $\hat{A} = XZ(Z^TZ + \sigma I)^{-1}, \Sigma_A = (\sigma^{-1}Z^TZ + I)^{-1}$
2. $\hat{Z} = X^T\hat{A}(\hat{A}^T\hat{A} + p\Sigma_A)^{-1}$

If you initialize at orthogonal solution $\hat{A} = U_{1:r}\Lambda_{1:r}$ it will be operating only on λs.

- Converges to orthogonal solution even from random start

Summary of the EM Algorithm

Iterate two least squares problems:

1. $\hat{A} = XZ(Z^T Z + \sigma I)^{-1}, \Sigma_A = (\sigma^{-1} Z^T Z + I)^{-1}$
2. $\hat{Z} = X^T \hat{A}(\hat{A}^T \hat{A} + p\Sigma_A)^{-1}$

If you initialize at orthogonal solution $\hat{A} = U_{1:r}\Lambda_{1:r}$ it will be operating only on λs.

- Converges to orthogonal solution even from random start
- Included prior on $p(A) = \mathcal{N}(0, I, I)$,
- Why I do not care about variance of the prior, λI?

Summary of the EM Algorithm

Iterate two least squares problems:

1. $\hat{A} = XZ(Z^T Z + \sigma I)^{-1}, \Sigma_A = (\sigma^{-1} Z^T Z + I)^{-1}$
2. $\hat{Z} = X^T \hat{A}(\hat{A}^T \hat{A} + p\Sigma_A)^{-1}$

If you initialize at orthogonal solution $\hat{A} = U_{1:r}\Lambda_{1:r}$ it will be operating only on λs.

- Converges to orthogonal solution even from random start
- Included prior on $p(A) = \mathcal{N}(0, I, I)$,
- Why I do not care about variance of the prior, λI?
  - ambiguity $AZ^T = ATT^{-1}Z^T$

Summary of the EM Algorithm

Iterate two least squares problems:

1. $\hat{A} = XZ(Z^T Z + \sigma I)^{-1}, \Sigma_A = (\sigma^{-1} Z^T Z + I)^{-1}$
2. $\hat{Z} = X^T \hat{A}(\hat{A}^T \hat{A} + p\Sigma_A)^{-1}$

If you initialize at orthogonal solution $\hat{A} = U_{1:r}\Lambda_{1:r}$ it will be operating only on λs.

- Converges to orthogonal solution even from random start
- Included prior on $p(A) = \mathcal{N}(0, I, I)$,
- Why I do not care about variance of the prior, λI?
  - ambiguity $AZ^T = ATT^{-1}Z^T$
- We can incorporate many more assumptions!
  - positivity
  - $p(A) = \mathcal{N}(0, I, \mathrm{diag}(\alpha))$

Positivity constraint on Z

What if we impose prior $p(Z_{i,j}) = \begin{cases} 1 & \text{if } Z_{i,j} > 0 \\ 0 & \text{otherwise.} \end{cases}$

Positivity constraint on Z

What if we impose prior $p(Z_{i,j}) = \begin{cases} 1 & \text{if } Z_{i,j} > 0 \\ 0 & \text{otherwise.} \end{cases}$

E-step (over A): $q(Z|Z^{(j)}) = \int \log p(X, A, Z) p(A|Z^{(j)}, X) dA$ s.t. $Z_{i,j} > 0, \forall i, j$

M-step (of Z): $Z^{(j+1)} = \arg\max_\theta q(Z|Z^{(j)})$

Positivity constraint on Z

What if we impose prior $p(Z_{i,j}) = \begin{cases} 1 & \text{if } Z_{i,j} > 0 \\ 0 & \text{otherwise.} \end{cases}$

E-step (over A): $q(Z|Z^{(j)}) = \int \log p(X, A, Z) p(A|Z^{(j)}, X) dA$ s.t. $Z_{i,j} > 0, \forall i, j$

M-step (of Z): $Z^{(j+1)} = \arg\max_\theta q(Z|Z^{(j)})$ s.t. $Z_{i,j} > 0, \forall i, j$

Recall the q function:

$$q(Z|Z^{(j)}) = \mathcal{N}(\hat{Z}, \Sigma_Z), \;\; \hat{Z} = X^T \hat{A}(\hat{A}^T \hat{A} + p\Sigma_A)^{-1},$$

With additional constraint we obtain truncated Normal distribution

Extreme $\hat{\hat{Z}}$is $\hat{\hat{Z}}_{i,j} = \begin{cases} \hat{Z}_{i,j} & \text{if } \hat{Z}_{i,j} > 0 \\ 0 & \text{otherwise} \end{cases}$

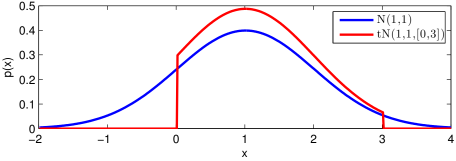
Positivity constraint on A

What if we impose prior $p(A_{i,j}) = t\mathcal{N}(0, I, I, [0, \infty])$

Positivity constraint on A

What if we impose prior $p(A_{i,j}) = t\mathcal{N}(0, I, I, [0, \infty])$

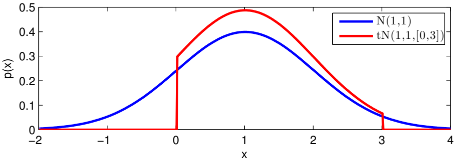
$$q(Z|Z^{(j)}) = \mathrm{E}_A \left\{ -\frac{1}{2}\sigma^{-1}\mathrm{tr}\left(X^T X - ZA^T X - XAZ^T + ZA^T AZ^T \right) \right\}$$



Positivity constraint on A

What if we impose prior $p(A_{i,j}) = t\mathcal{N}(0, I, I, [0, \infty])$

$$q(Z|Z^{(j)}) = \mathrm{E}_A \left\{ -\frac{1}{2}\sigma^{-1}\mathrm{tr}\left(X^TX - ZA^TX - XAZ^T + ZA^TAZ^T\right)\right\}$$



$\mathrm{E}(A) \approx \tilde{A}_{i,j} = \mathcal{N}(\hat{A}, I, \mathrm{diag}(\Sigma_A))$ product of univariate tN. Moments of $t\mathcal{N}(\hat{a}, \sigma_a)$ are

$$\mathrm{E}(a) = \hat{a} + \frac{\phi(\alpha) - \phi(\beta)}{Z_a},$$

$$\mathrm{E}(a^2) = \sigma_a^2 \left\{ 1 + \frac{\alpha\phi(\alpha) - \beta\phi(\beta)}{Z_a} \right\}$$

where $\alpha = \frac{-\hat{a}}{\sqrt{\sigma_a}}$, $\beta = \infty$, $\phi(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/s)$, $Z = \Phi(\beta) - \Phi(\alpha)$, $\Phi(x) = \frac{1}{2}(1 + \mathrm{erf}(x))$.

Variational Bayes for PCA

Bayes rule

$$p(A, Z|X) \propto p(X|A, Z)p(A)p(Z)$$

where

$$p(A) = \mathcal{N}(0, I, I) \propto \exp\left(-\frac{1}{2}\text{tr}\left(AI_rA^TI_p\right)\right),$$
$$p(Z) = \mathcal{N}(0, I, I) \propto \exp\left(-\frac{1}{2}\omega_Z\text{tr}(ZI_rZ^TI_n)\right)$$

we can no longer have both variances one.

Variational Bayes for PCA

Bayes rule

$$p(A, Z|X) \propto p(X|A, Z)p(A)p(Z)$$

where

$$p(A) = \mathcal{N}(0, I, I) \propto \exp\left(-\frac{1}{2}\text{tr}\,(AI_r A^T I_p)\right),$$

$$p(Z) = \mathcal{N}(0, I, I) \propto \exp\left(-\frac{1}{2}\omega_Z \text{tr}(ZI_r Z^T I_n)\right)$$

we can no longer have both variances one.

$$q(A, Z) \approx q(A)q(Z)$$

Yields:

1. $\hat{A} = X\hat{Z}(\hat{Z}^T\hat{Z} + n\Sigma_Z + I)^{-1}, \Sigma_A = (Z^T Z + n\Sigma_Z + I)^{-1}$
2. $\hat{Z} = X^T\hat{A}(\hat{A}^T\hat{A} + p\Sigma_A + \omega_Z I)^{-1}, \Sigma_Z = (\hat{A}^T\hat{A} + p\Sigma_A + \omega_Z I)^{-1}$,

In practice we need to estimate precision of data ω_X, ω_Z needing expectations

$$\text{E}_{A,Z}\left(XA^T AX^T\right)$$

Toy matrix decomposition

Consider 1×1 matrix d, decomposed

$$p(d|a, x) = \mathcal{N}(ax, r_e),$$

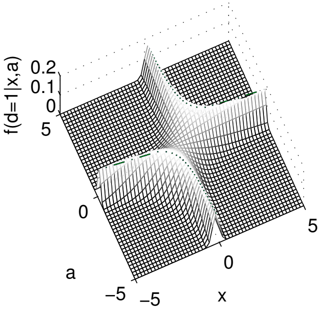
Find a, x.

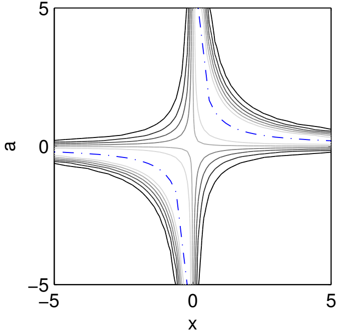
Toy matrix decomposition

Consider 1×1 matrix d, decomposed

$$p(d|a,x) = \mathcal{N}(ax, r_e),$$

Find a, x.

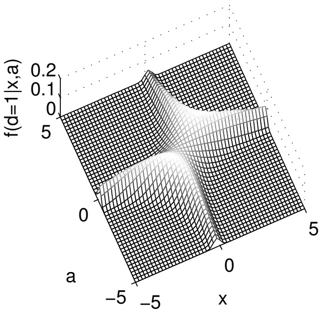


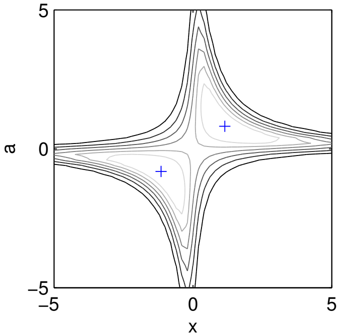


Toy matrix decomposition

Consider 1×1 matrix d, decomposed

$$p(d|a,x) = \mathcal{N}(ax, r_e),$$
$$p(x) = \mathcal{N}(0, r_x),$$
$$p(a) = \mathcal{N}(0, r_a)$$





Toy maximum likelihood

Joint distribution:

$$\log p(d, a, x) \propto -\frac{1}{2r_e}(d - ax)^2 - \frac{1}{2r_a}a^2 - \frac{1}{2r_x}x^2$$

Find

$$\hat{x}, \hat{a} = \arg\max_{a,x}(\log p(d, a, x))$$

For $d < \frac{r_e}{\sqrt{r_a r_x}}$,

$$\hat{x} = 0, \hat{a} = 0,$$

For $d \geq \frac{r_e}{\sqrt{r_a r_x}}$,

$$\hat{x} = \pm\left(d\sqrt{\frac{r_x}{r_a}} - \frac{r_e}{r_a}\right)^{\frac{1}{2}}, \qquad \hat{a} = \pm\left(d\sqrt{\frac{r_a}{r_x}} - \frac{r_e}{r_x}\right)^{\frac{1}{2}}.$$

Note that the product of the maxima is

$$\hat{a}\hat{x} = d - \frac{r_e}{\sqrt{r_a r_x}}.$$

Marginal likelihood (PPCA)

Joint distribution:

$$\log p(d, a, x) \propto -\frac{1}{2r_e}(d - ax)^2 - \frac{1}{2r_a}a^2 - \frac{1}{2r_x}x^2$$

Marginal

$$\begin{aligned} p(a|d) &\propto \int p(d, a, x)dx \\ &\propto \exp(-\frac{1}{2}d^2(a^2 r_x + r_e)^{-1})\sqrt{r_e r_x}(a^2 r_x + r_e)^{-\frac{1}{2}}, \end{aligned}$$

with maximum:

$$\hat{a} = \begin{cases} \frac{\sqrt{d^2 - r_e}}{\sqrt{r_x}} & \text{if } d^2 > r_e, \\ 0 & \text{otherwise}, \end{cases}$$

Variational Bayes

Joint distribution:

$$\log p(d, a, x) \propto -\frac{1}{2r_e}(d - ax)^2 - \frac{1}{2r_a}a^2 - \frac{1}{2r_x}x^2$$

Factor $q(a|d)$

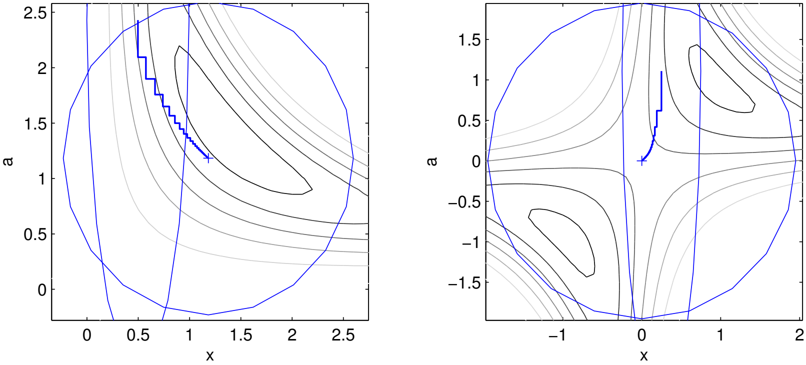
$$\begin{aligned} p(a|d) &\propto \exp\left(\mathrm{E}_x\left[-\frac{1}{2r_e}(d - ax)^2 - \frac{1}{2r_a}a^2\right]\right) \\ &= \mathcal{N}(\hat{a}, \sigma_a), \\ \hat{a} &= \sigma_a d \langle x \rangle r_e^{-1}, \quad \sigma_a^{-1} = \left\langle a^2 \right\rangle r_e^{-1} + r_a^{-1} \end{aligned}$$

Factor $q(x|d)$

$$\begin{aligned} p(a|d) &\propto \exp\left(\mathrm{E}_x\left[-\frac{1}{2r_e}(d - ax)^2 - \frac{1}{2r_a}a^2\right]\right) \\ &= \mathcal{N}(\hat{a}, \sigma_a), \\ \hat{a} &= \sigma_a d \langle x \rangle r_e^{-1}, \quad \sigma_a^{-1} = \left\langle a^2 \right\rangle r_e^{-1} + r_a^{-1} \end{aligned}$$

Convergence of VB

1. compute $\hat{a}, \sigma_a$, and $\langle a \rangle = \hat{a}$, $\left\langle a^2 \right\rangle = \hat{a}^2 + \sigma_a$,
2. compute $\hat{x}, \sigma_x$, and $\langle x \rangle = \hat{x}$, $\left\langle x^2 \right\rangle = \hat{x}^2 + \sigma_x$,



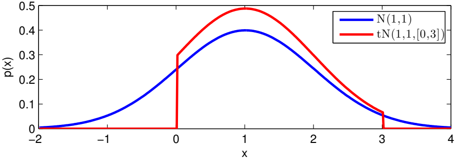
Positive support

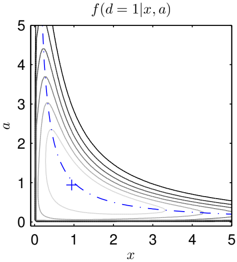
What if we are interested only in the positive solution?

Positive support

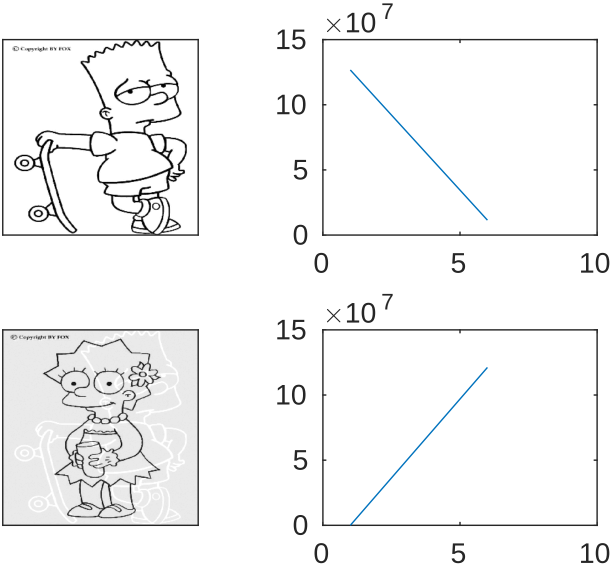
What if we are interested only in the positive solution?

$$
\begin{aligned}
p(d|a,x) &= \mathcal{N}(ax, r_e),\\
p(x) &= t\mathcal{N}(0, r_x, \langle 0, \infty\rangle) \propto \mathcal{N}(0, r_x)\chi(x>0),\\
p(a) &= t\mathcal{N}(0, r_a, \langle 0, \infty\rangle) \propto \mathcal{N}(0, r_a)\chi(a>0),
\end{aligned}
$$





Non-negative Matrix Factorization (NMF)



Extensions: Multivariate PCA

1. Automatic Relevance Determination (#of factors):

$$p(A) = \mathcal{N}(0, I, \text{diag}(\alpha)),$$

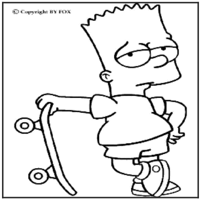
2. Automatic Relevance Determination (#pixels):

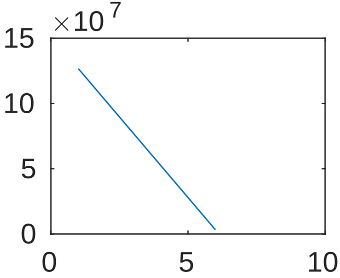
$$p(\text{vec}(A)) = \mathcal{N}(0, \text{diag}(\alpha)),$$

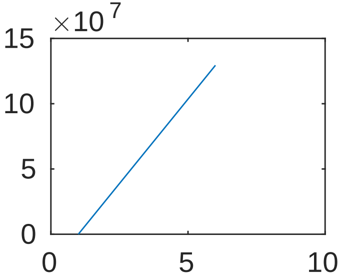
   where α

3. many more

Sparse Non-negative Matrix Factorization (SNMF)





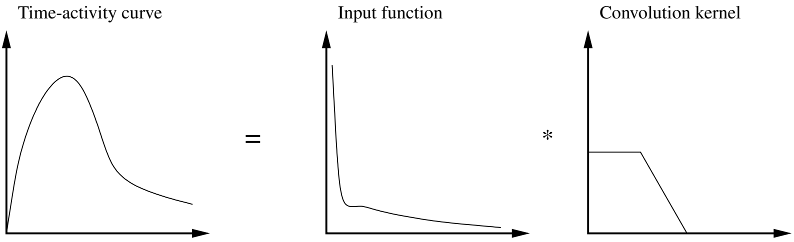


Assignment

Matrix factorization		points
EM or VB, Bart-Lisa		8
+ positive support		+2

BSS with deconvolution

Time-activity curves for brain imaging are results of convolution.



Model of the curve

$$x = b * w = Bw, \qquad B = \begin{pmatrix} b_1 & 0 & 0 & 0 \\ b_2 & b_1 & 0 & 0 \\ \dots & b_2 & b_1 & 0 \\ b_n & \dots & b_2 & b_1 \end{pmatrix}$$

Unknows are b and w. Kernel is sparse.

BSS with deconvolution

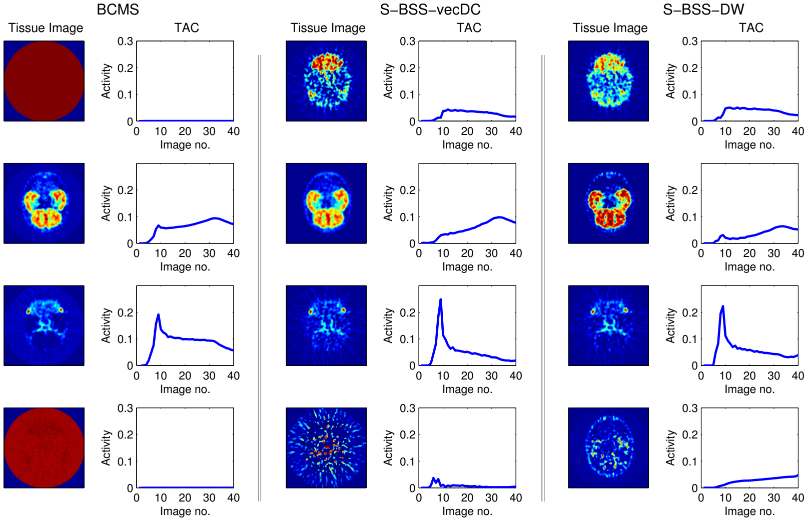


Image deconvolution

(a) out-of-focus blur

Image deconvolution

Mathematical model:

$$d = Ax + e$$

where

d is the observed (blured) image,

x is the original (true) image, unknown

A is the convolution matrix,

e is the measurement (model) error

Find: x, A, $var(e)$

Image deconvolution

Mathematical model:

$$d = Ax + e$$

where

d	is the observed (blured) image,	
x	is the original (true) image,	unknown
A	is the convolution matrix,	unknown kernel
e	is the measurement (model) error	

Find: x, A, $var(e)$

Image deconvolution

Mathematical model:

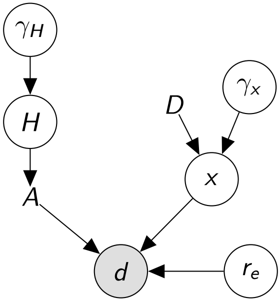
$$d = Ax + e$$

where

d	is the observed (blured) image,	
x	is the original (true) image,	unknown
A	is the convolution matrix,	unknown kernel
e	is the measurement (model) error	unknown variance

Find: x, A, $var(e)$

Model



- number of unknowns $> 3\times$ higher than number of observations
- ARD coefficients γ_H, γ_x, r_e,
- approximations of the covariance matrices by diagonal

Results:

(a) out-of-focus blur

(b) blind deconvolution