Klasifikace a rozpoznávání

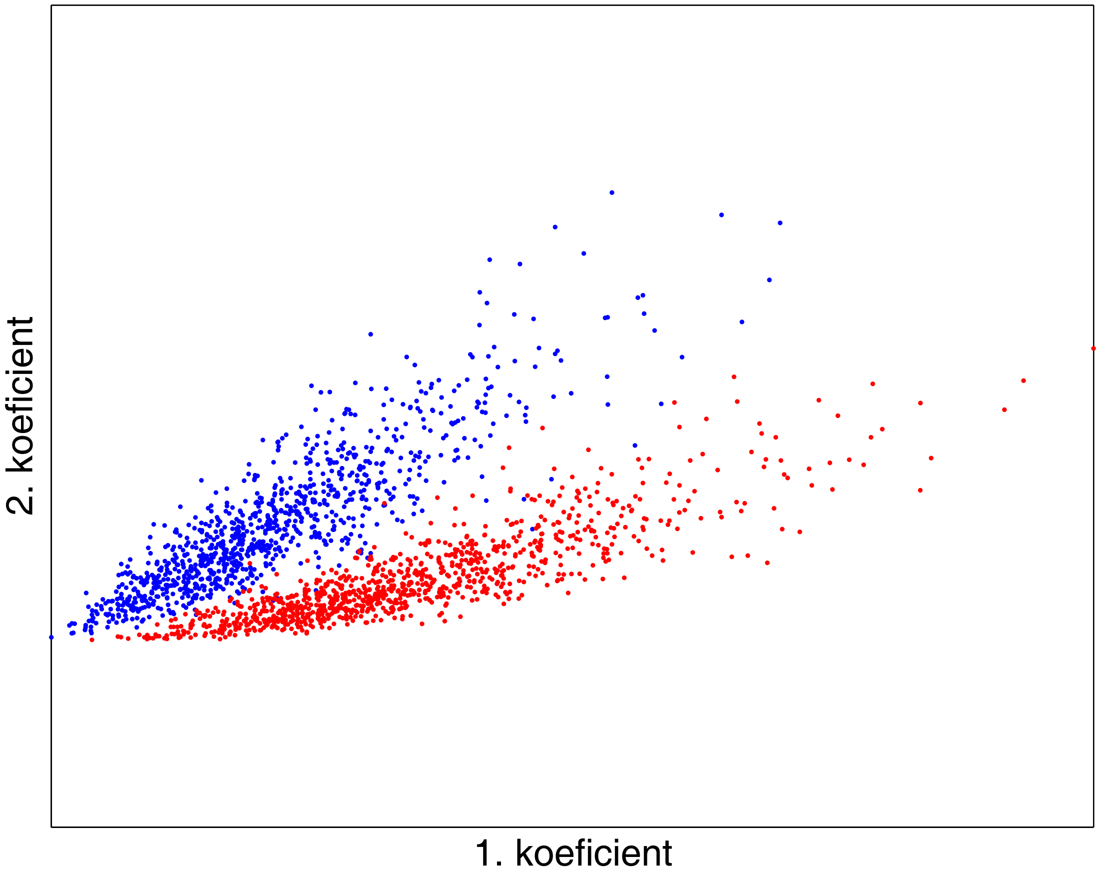
Extrakce příznaků

Extrakce příznaků - parametrizace

- Poté co jsme ze snímače obdržely data která jsou relevantní pro naši klasifikační úlohu, je potřeba je přizpůsobit potřebám rozpoznávače
- Klasifikátory mají rády parametry které jsou:
  - Gaussovského rozložení (většinou vícerozměrného)
  - Nekorelované
  - Nízkodimenzionální

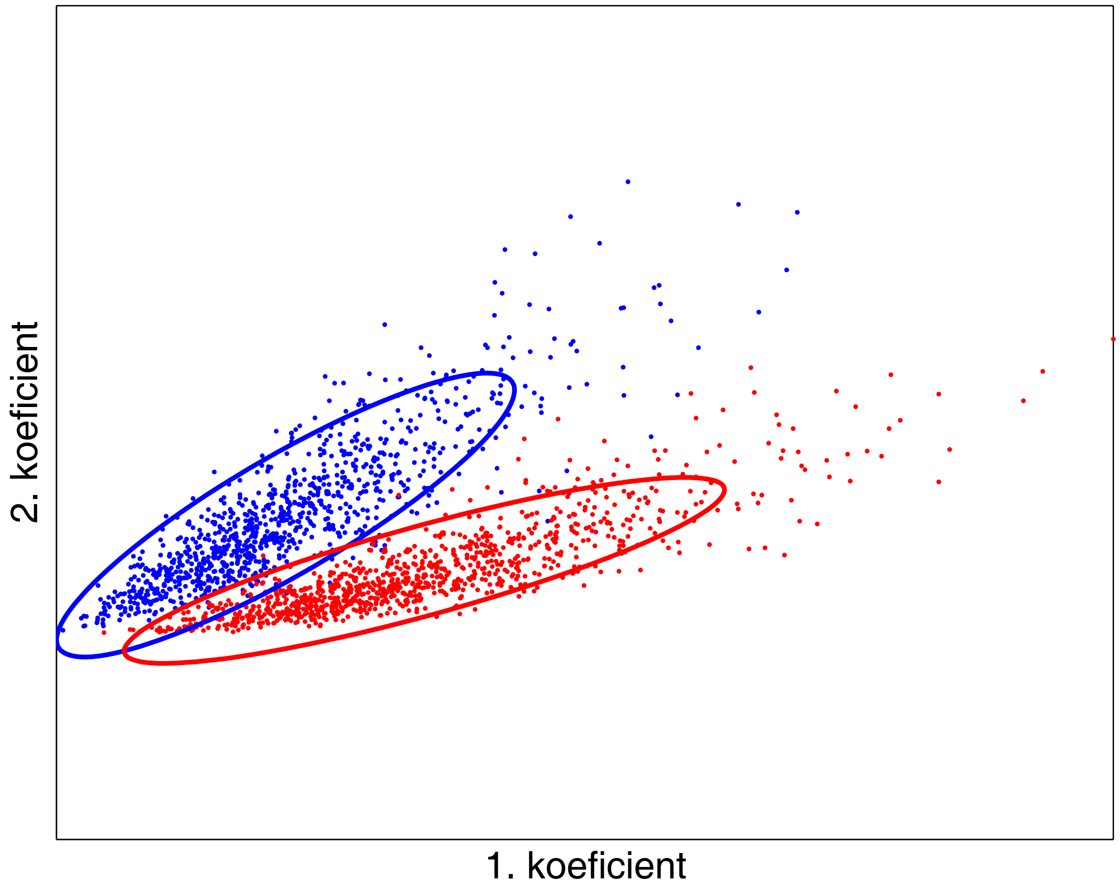
Příklad parametrizace pro 2D vstupní vektory

- Mějme vzorky (příklady) 2D rozložení pro dvě třídy.



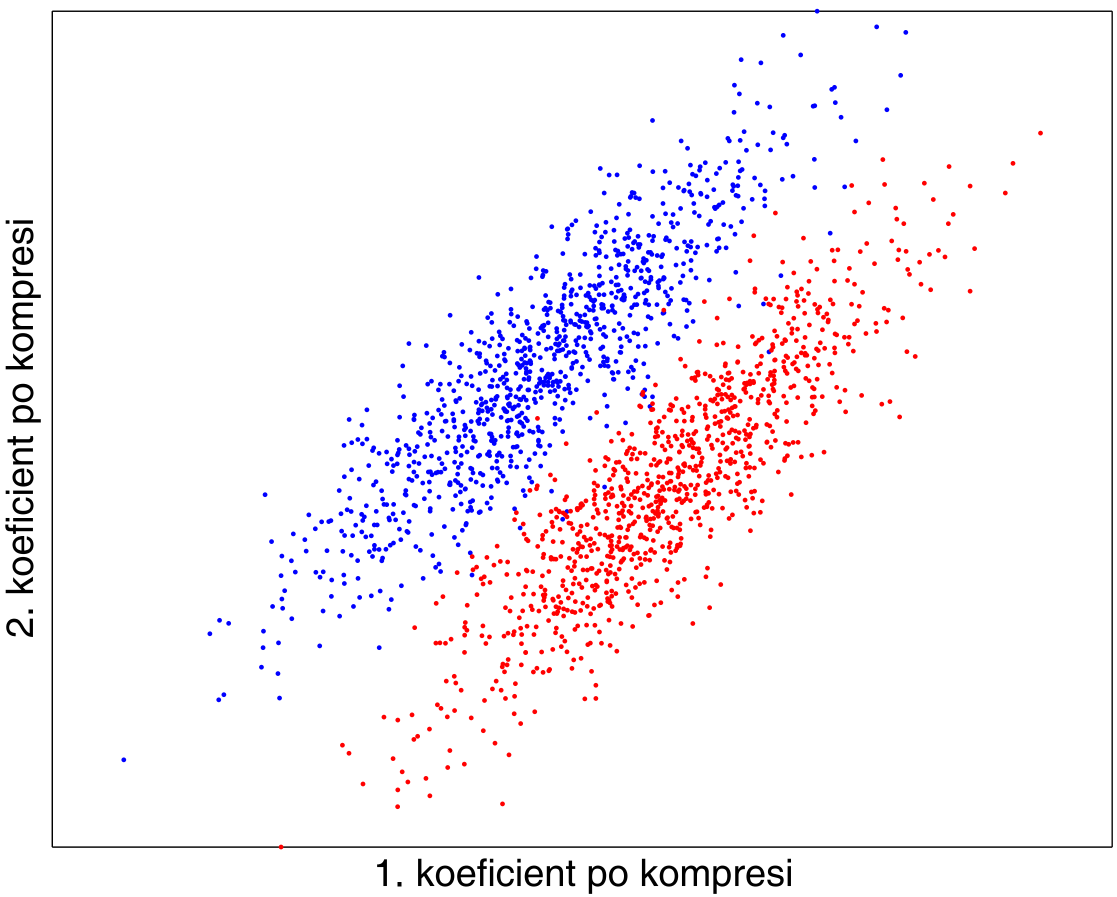
Příklad parametrizace pro 2D vstupní vektory

- Rozložení není příliš gaussovské.
- Provedeme třetí odmocninou obou koeficientů.



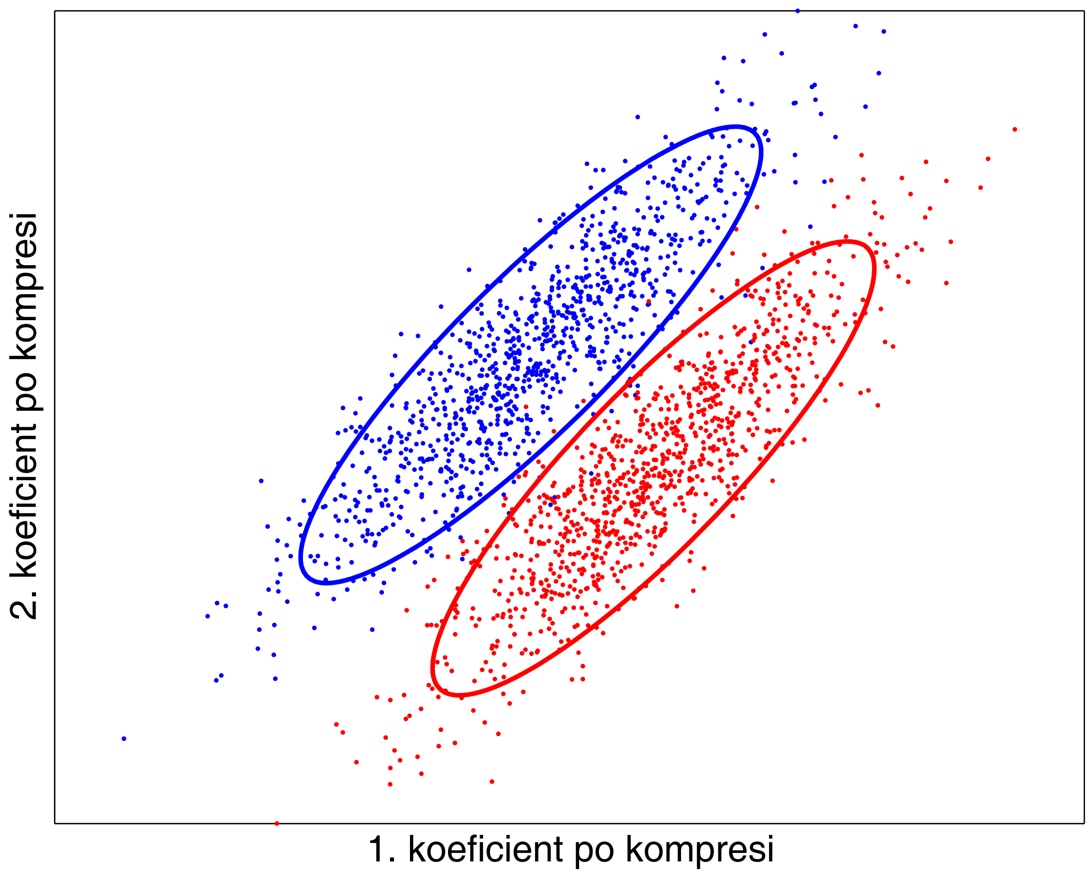
Příklad parametrizace pro 2D vstupní vektory

- Prostor se komprimuje – nelineárně deformuje...



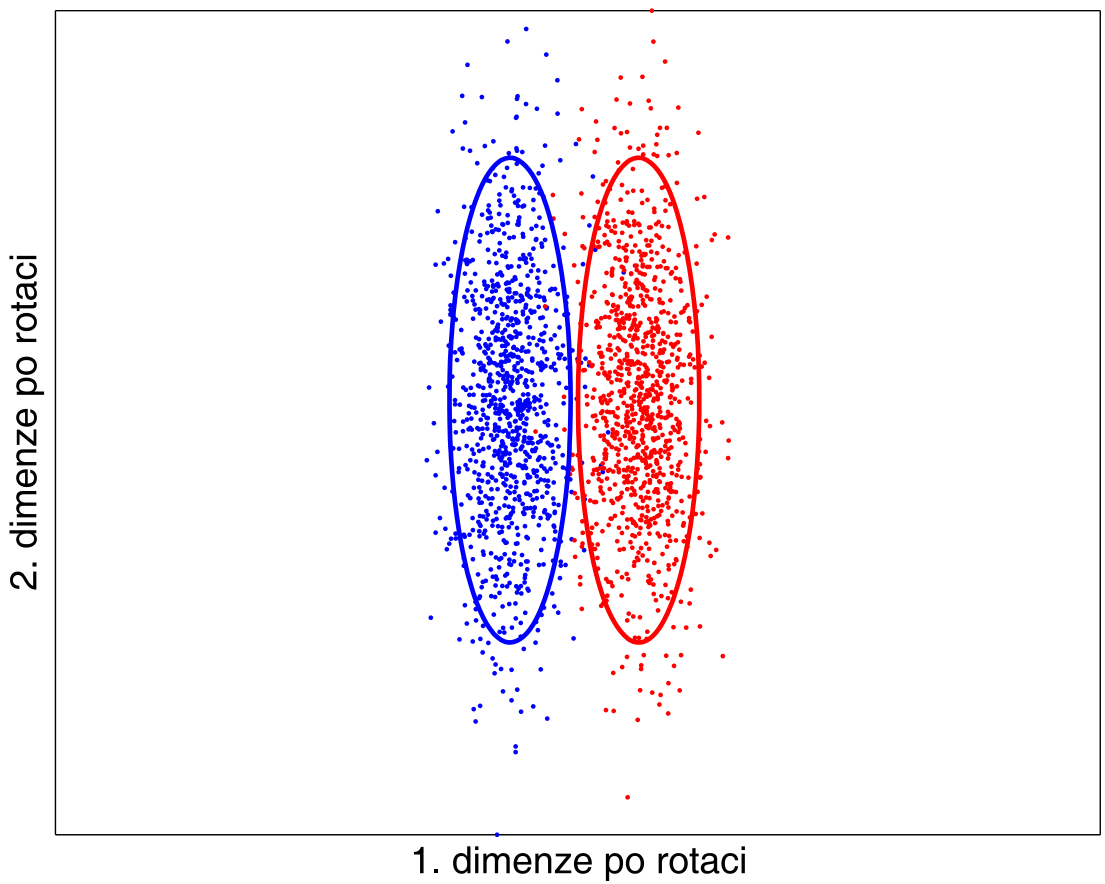
Příklad parametrizace pro 2D vstupní vektory

- ... a rozložení pro každou třídu je nyní gaussovské.
- Koeficienty jsou ale korelované.
- Je vhodné prostor otočit tak aby se koeficienty dekorelovaly.



Příklad parametrizace pro 2D vstupní vektory

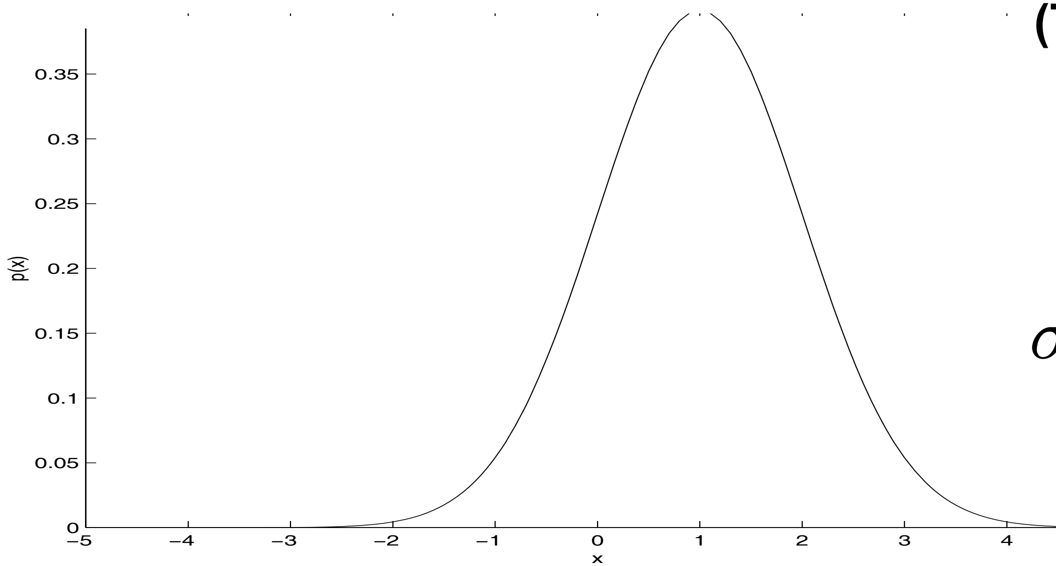
- Nyní jsou koeficienty dekorelovány.
- Svislá dimenze je navíc zbytečná, protože třídy se v ní zcela překrývají.



Gaussovské rozložení (jednorozměrné)

Evaluation:

$$\mathcal{N}(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



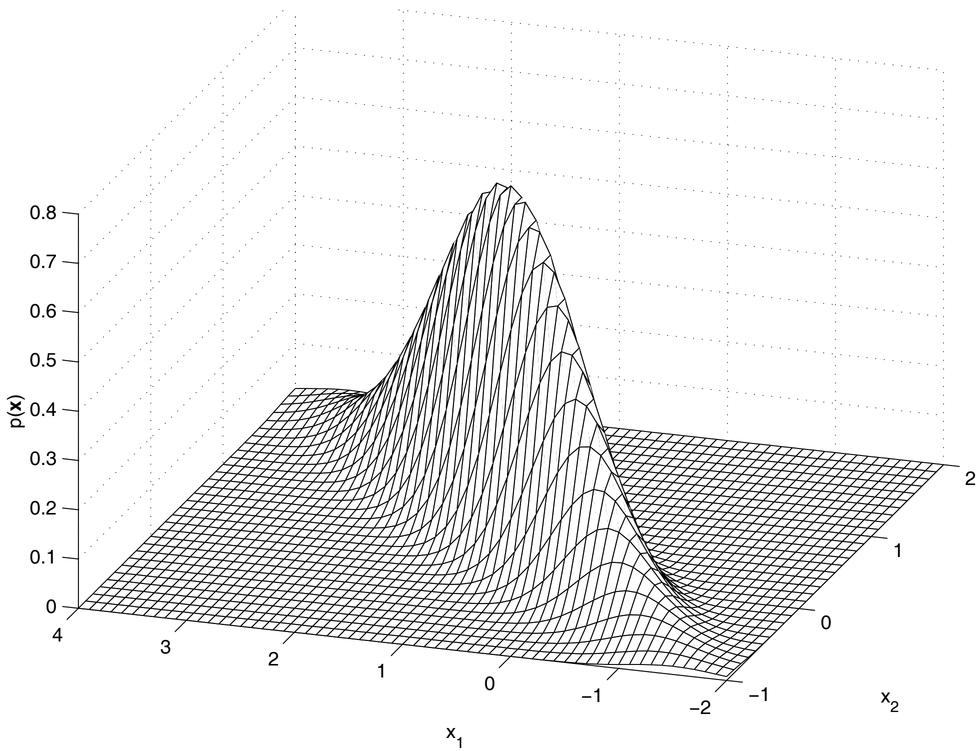
ML odhad parametrů (Trénování):

$$\mu = \frac{1}{T}\sum_{t=1}^{T} x(t)$$

$$\sigma^2 = \frac{1}{T}\sum_{t=1}^{T}(x(t)-\mu)^2$$

Gaussian distribution (2 dimensions)

$$\mathcal{N}(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^P|\boldsymbol{\Sigma}|}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}$$

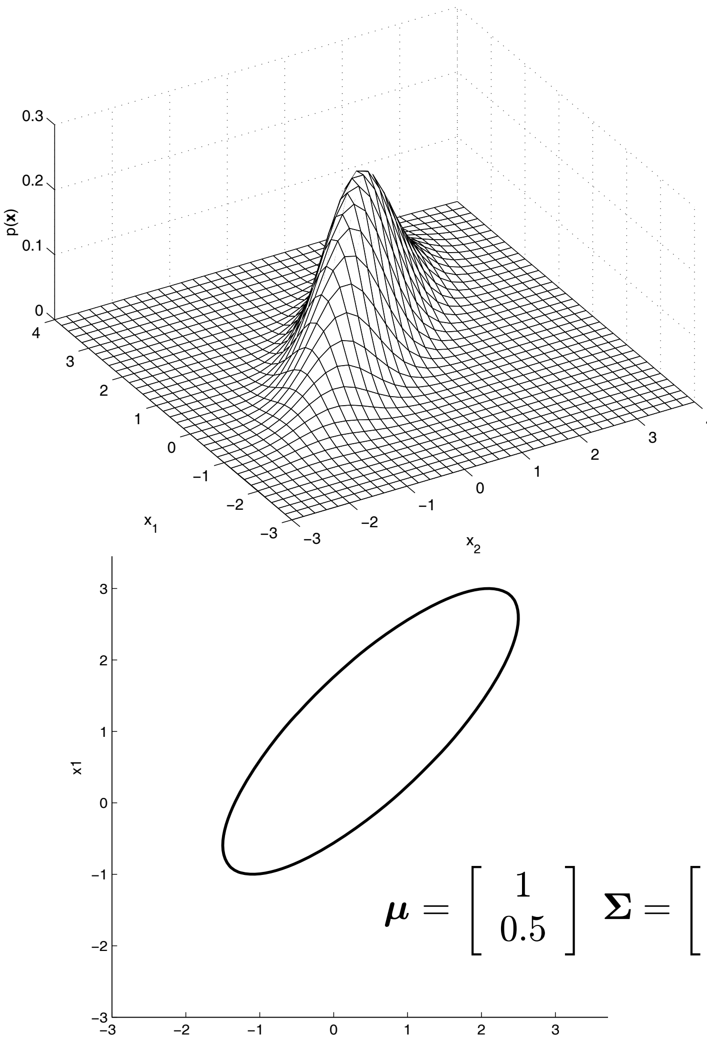


ML odhad of parametrů (Trénování):

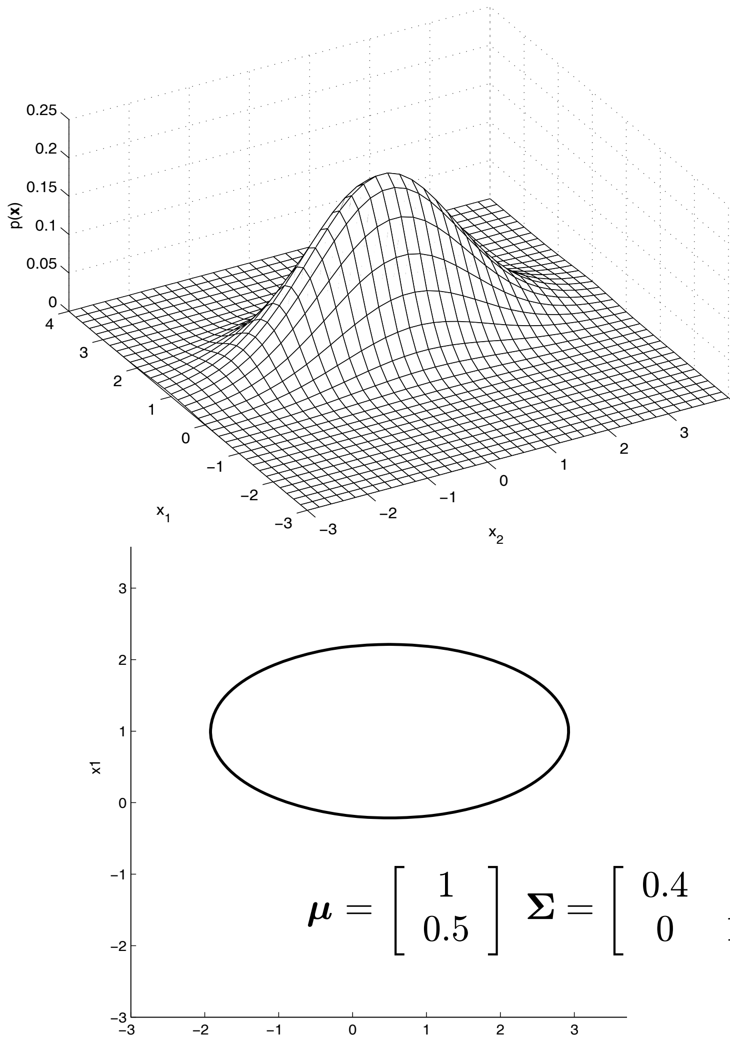
$$\boldsymbol{\mu} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{x}(t)$$

$$\boldsymbol{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{x}(t)-\boldsymbol{\mu})(\boldsymbol{x}(t)-\boldsymbol{\mu})^T$$

Plná a diagonální kovarianční matice



$$\boldsymbol{\mu} = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix} \; \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}$$



$$\boldsymbol{\mu} = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix} \; \boldsymbol{\Sigma} = \begin{bmatrix} 0.4 & 0 \\ 0 & 1.5 \end{bmatrix}$$

Diagonální kovarianční matice

$$\mathcal{N}(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^P|\boldsymbol{\Sigma}|}}e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}$$

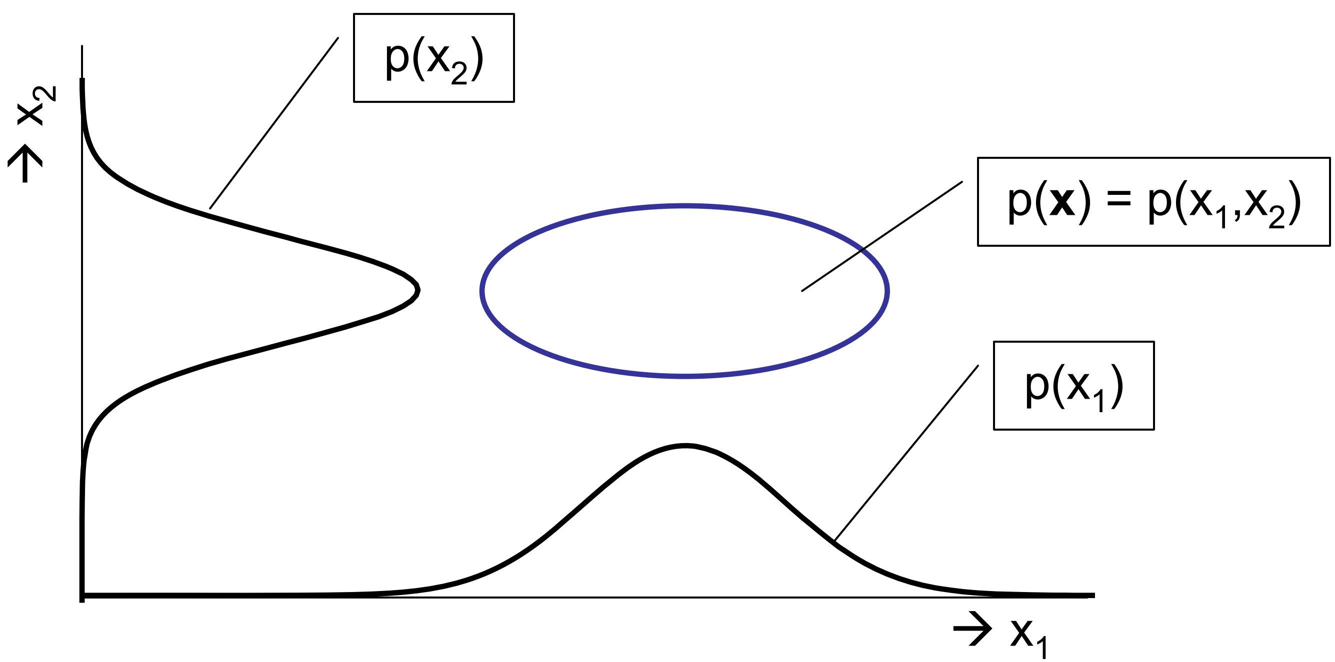
Pokud je **Σ** diagonální matice s koeficienty v diagonále σ^2_i ➔

$$\begin{aligned}
\mathcal{N}(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma}) &= \frac{1}{\sqrt{(2\pi)^P\prod_{i=1}^P\sigma_i^2}}\exp(-\frac{1}{2}\sum_{i=1}^P\frac{(x_i-\mu_i)^2}{\sigma_i^2}) \\
&= \prod_{i=1}^P\frac{1}{\sqrt{(2\pi)\sigma_i^2}}\exp(-\frac{1}{2}\frac{(x_i-\mu_i)^2}{\sigma_i^2}) \\
&= \prod_{i=1}^P\mathcal{N}(x_i;\mu_i,\sigma_i^2)
\end{aligned}$$

Diagonální kovarianční matice

$P(A,B) = P(A)P(B) \Rightarrow$ Jevy A a B jsou statisticky nezávislé

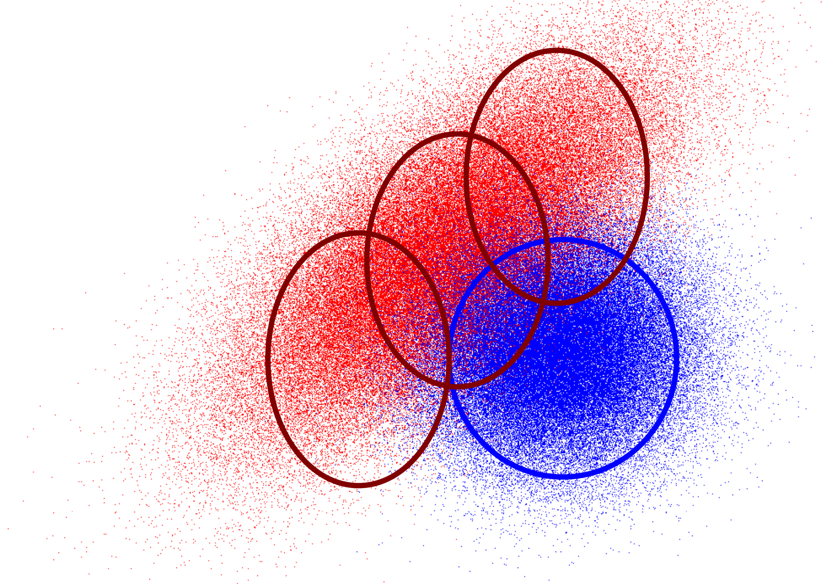
$\mathcal{N}(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma}) = \prod_{i=1}^{P} \mathcal{N}(x_i;\mu_i,\sigma_i^2) \Rightarrow$ Koeficienty x_i příznakového vektoru **x** jsou statisticky nezávislé.

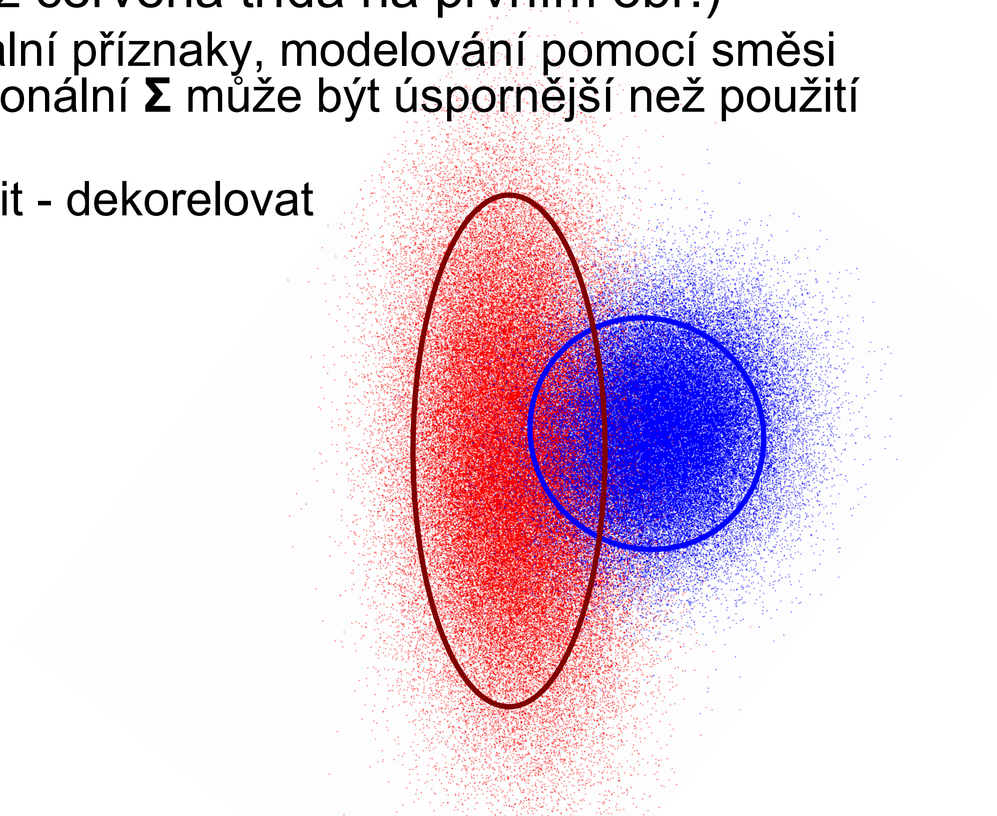


Diagonální kovarianční matice

Proč nás zajímá?

- Pomůže nám pochopit význam plné kovarianční matice v gaussovském rozložení
- Úspora parametrů při modelování dat
- Pokud jsou data korelována (viz červená třída na prvním obr.)
  - Zvláště pro vysoce dimenzionální příznaky, modelování pomocí směsi gaussovských rozložení s diagonální **Σ** může být úspornější než použití jedné gaussovky s plnou **Σ**
  - Můžeme se pokusit data natočit - dekorelovat

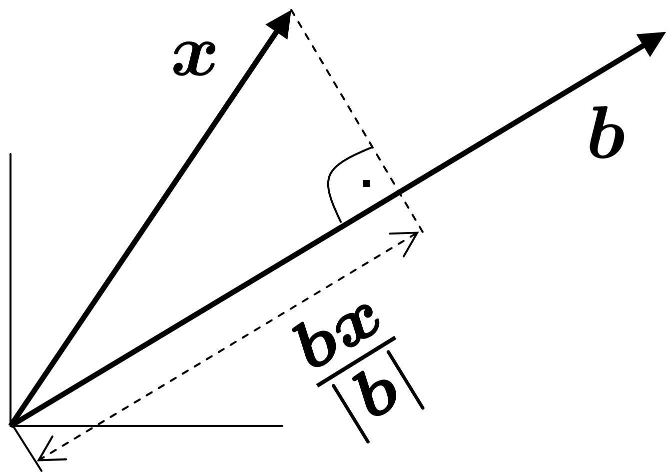




Skalární součin

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \boldsymbol{b} = \begin{bmatrix} b_1 & b_2 \end{bmatrix}$$

$$\boldsymbol{bx} = \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = b_1 x_1 + b_2 x_2$$



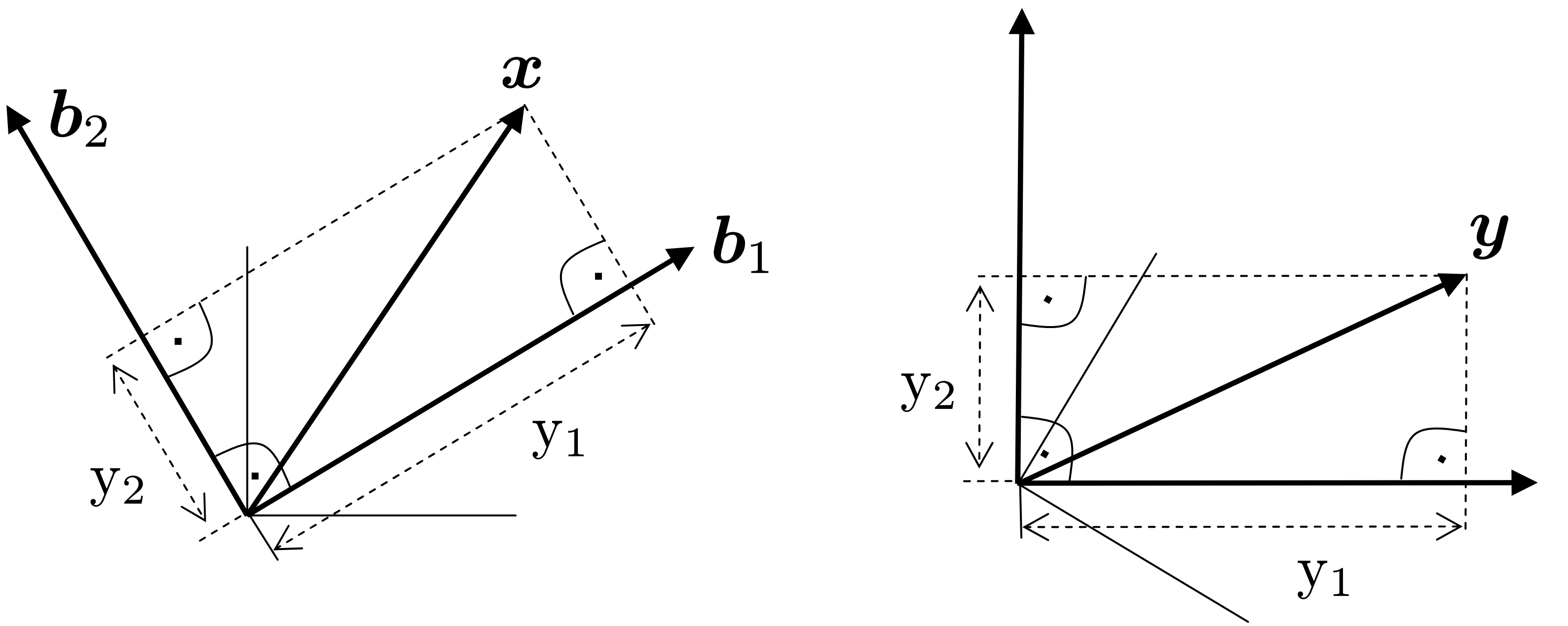
Rotace vektoru

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \boldsymbol{B} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

Nechť $\mathbf{b}_1$ a $\mathbf{b}_2$ jsou ortonormální baze

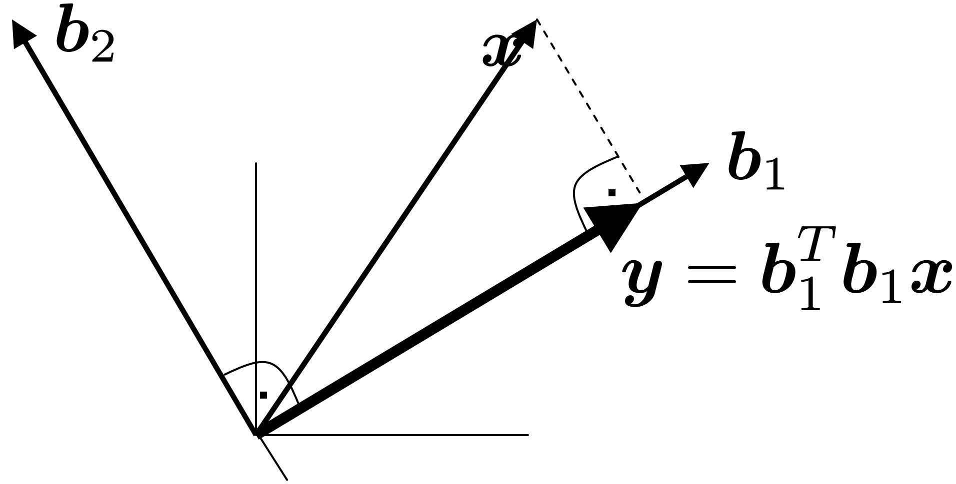
- Vektory jsou na sebe kolmé
- Mají délku $|\mathbf{b}_1| = |\mathbf{b}_2| = 1$

Potom $\mathbf{y} = \mathbf{B}\,\mathbf{x}$ je otočený vektor $\mathbf{x}$, kde $\mathbf{b}_1$ a $\mathbf{b}_2$ ukazují v původním prostoru směry nových os



Projekce vektoru

- Nechť **B** je matice ortonormálních bází a **B**’ matice tvořena pouze několika řádky (bázemi) matice **B.**
- Potom **y** = **B**’^{T}**B**’**x** je projekce vektoru **x** do bází **B**’.



Vlastní čísla a vektory

λ je vlastní číslo a **e** je odpovídající vlastni vektor čtvercové matice **Σ**, pokud platí:

$$\boldsymbol{\Sigma e} = \boldsymbol{e}\lambda$$

PxP matice má (nanejvýš) P různých vlastních čísel. Nechť je **Λ** diagonální matice všech vlastních čísel a matice **E** obsahuje ve sloupcích odpovídající vlastní vektory.

$$\boldsymbol{\Sigma E} = \boldsymbol{E \Lambda}$$

Nás bude zajímat speciální případ kdy matice **Σ** je symetrická. Potom budou sloupce matice **E** tvořit ortonormální báze. Pro takovou matici potom platí: $\mathbf{E}^T\mathbf{E} = \mathbf{E}^{-1}\mathbf{E} = \mathbf{I}$, kde **I** je jednotková matice. Tedy platí následující rozklady matic:

$$\boldsymbol{E}^T \boldsymbol{\Sigma E} = \boldsymbol{\Lambda} \qquad \boldsymbol{\Sigma} = \boldsymbol{E \Lambda E}^T$$

μ transformovaných dat

- Jak se změní odhady střední hodnoty a kovarianční matice pokud původní data transformujeme: **y = Ax**

$$
\begin{aligned}
\boldsymbol{\mu}_{\boldsymbol{y}} &= \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{A}\boldsymbol{x}(t) \\
&= \boldsymbol{A}\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{x}(t) \\
&= \boldsymbol{A}\boldsymbol{\mu}_{\boldsymbol{x}}
\end{aligned}
$$

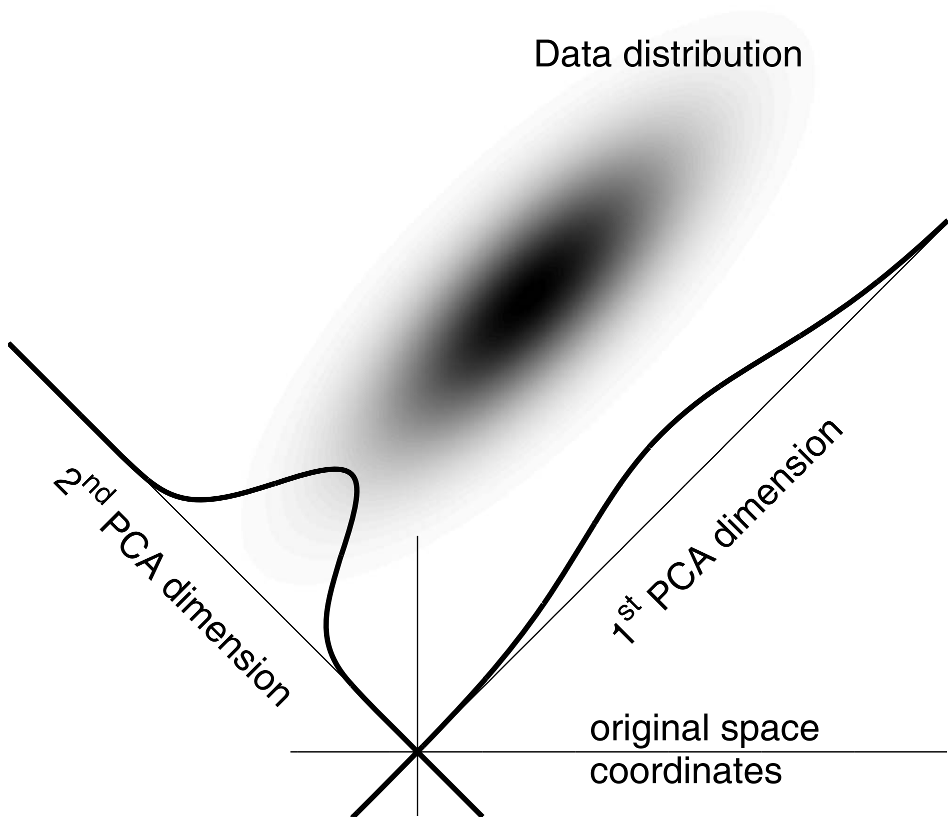
$\boldsymbol{\Sigma}$ transformovaných dat

$$\begin{aligned}
\boldsymbol{\Sigma_y} &= \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{A}\boldsymbol{x}(t)-\boldsymbol{\mu}_y)(\boldsymbol{A}\boldsymbol{x}(t)-\boldsymbol{\mu}_y)^T \\
&= \boldsymbol{A}\frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{x}(t)-\boldsymbol{\mu}_x)(\boldsymbol{x}(t)-\boldsymbol{\mu}_x)^T\boldsymbol{A}^T \\
&= \boldsymbol{A}\boldsymbol{\Sigma_x}\boldsymbol{A}^T
\end{aligned}$$

- Co se stane když jako **A** použijeme transponovanou matici vlastních vektoru kovarianční matice $\boldsymbol{\Sigma}_x$? (Proč transponovanou? Protože vlastní vektory máme ve sloupcích a ne v řádcích).
- Jaký význam mají vlastní čísla?

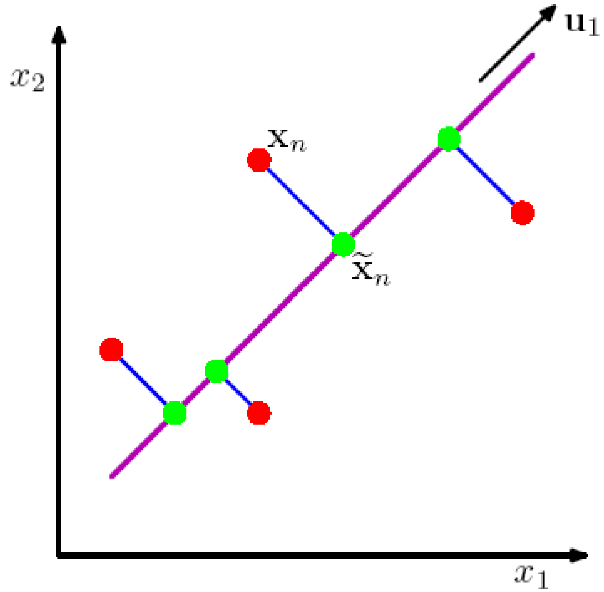
Analýza hlavních komponent

(Principal Component Analysis - PCA)



Analýza hlavních komponent

- Umožňuje:
  - Dekorelaci – vlastní vektory kovarianční matice definuji souřadný systém ve kterých jsou data dekorelována – mají diagonální kovarianční matici
  - Redukci dimenzí – promítnutí dat do pouze několika vlastních vektorů odpovídajících největším vlastním číslům (směry s nevětší variancí) umožní optimální rekonstrukci dat s nejmenší kvadratickou chybou (mean square error - MSE)
  - Redukce dimenzí provádíme pokud věříme, že v některých směrech není užitečná informace ale pouze (gaussovský) šum s nízkou variabilitou.

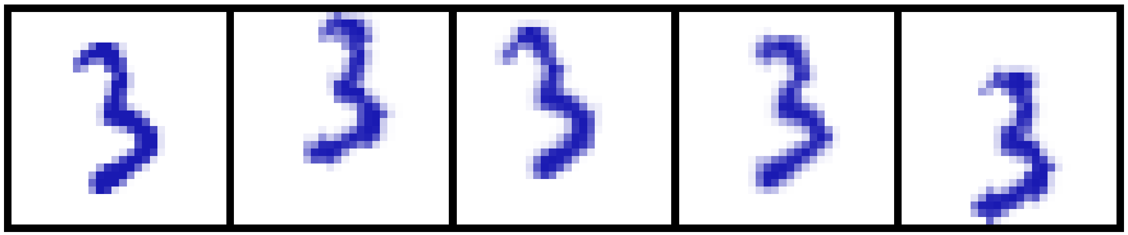


Interpretace $\boldsymbol{\Sigma}$ v gaussovském rozložení

$$
\begin{aligned}
\mathcal{N}(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma}) &= \frac{1}{\sqrt{(2\pi)^P|\boldsymbol{\Sigma}|}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})} \\
&= \frac{1}{\sqrt{(2\pi)^P|\boldsymbol{\Sigma}|}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{E}\boldsymbol{\Lambda}^{-1}\boldsymbol{E}^T(\boldsymbol{x}-\boldsymbol{\mu})} \\
&= \frac{1}{\sqrt{(2\pi)^P|\boldsymbol{\Lambda}|}} e^{-\frac{1}{2}(\boldsymbol{E}^T\boldsymbol{x}-\boldsymbol{E}^T\boldsymbol{\mu})^T\boldsymbol{\Lambda}^{-1}(\boldsymbol{E}^T\boldsymbol{x}-\boldsymbol{E}^T\boldsymbol{\mu})}
\end{aligned}
$$

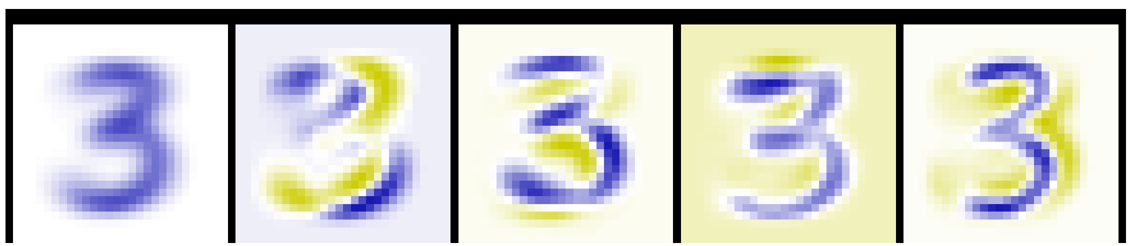
PCA - Příklad

- Obrázky 100x100 pixelů – 10000 dimensionální vektory



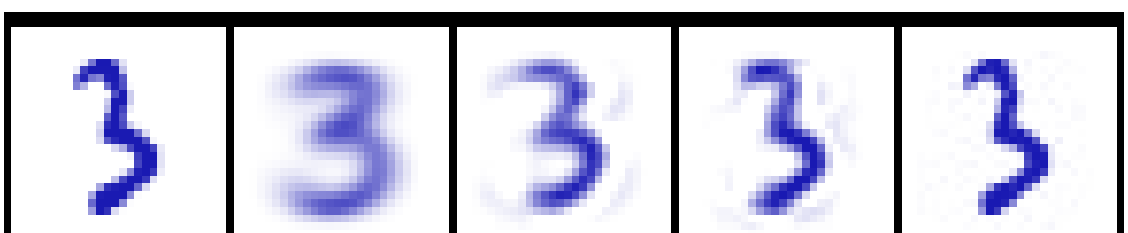
- Střední hodnota, vlastní čísla a vlastní vektory

μ $\lambda_1=3.4\cdot10^5$ $\lambda_2=2.8\cdot10^5$ $\lambda_3=2.4\cdot10^5$ $\lambda_3=1.6\cdot10^5$



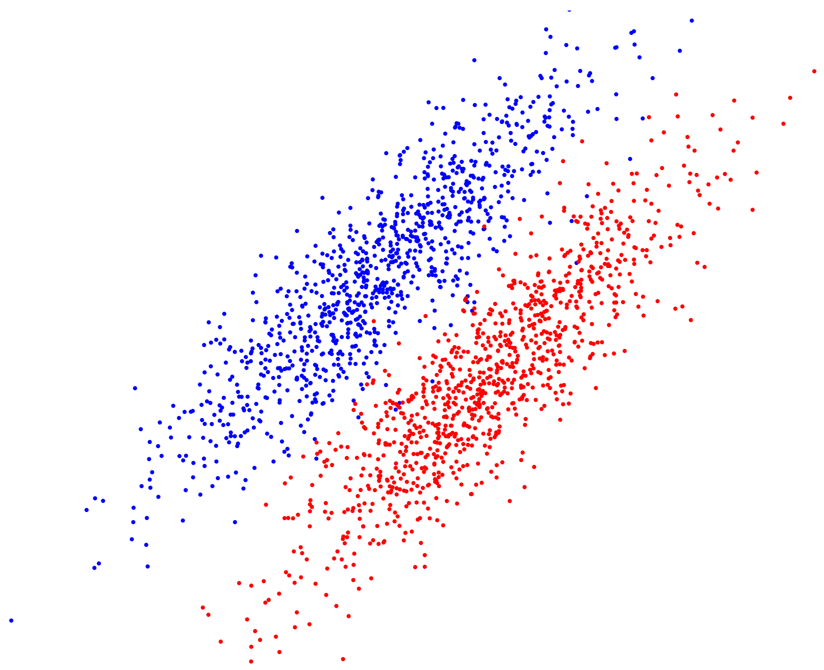
- Střední hodnota, vlastní čísla a vlastní vektory

Originál M = 1 M=10 M=50 M=250



PCA - Příklad

- Jakou dimenzi si PCA vybere na tomto příkladě?
- Bude to výhodné pro klasifikaci tříd?

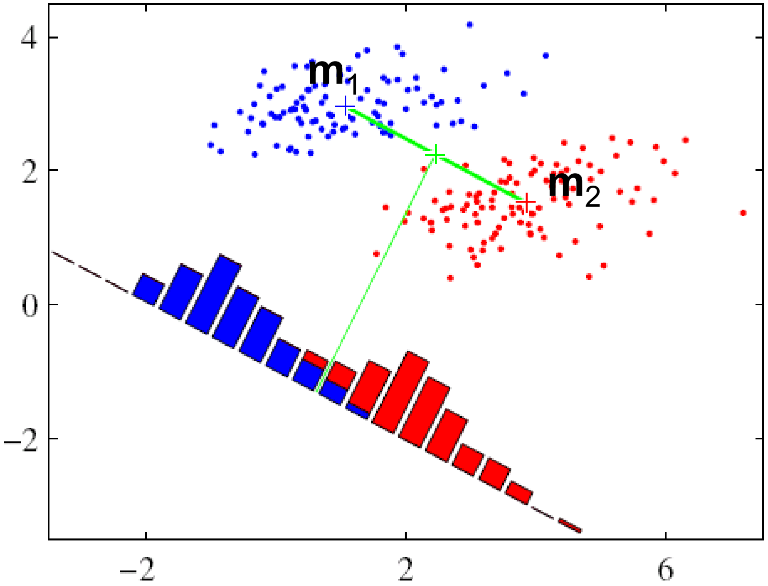


Lineární diskriminační analýza

- Opět se pokusíme promítnout data pouze do určitého směru: $y = \mathbf{w}^{\mathrm{T}}\mathbf{x}$
- Tentokrát ale budeme chtít aby v tomto směru byly separovány třídy.
- Intuitivně by nás mohlo napadnout vybrat směr ve kterém jsou nejlépe odděleny průměty středních hodnot tříd $\mathbf{m}_1$ a $\mathbf{m}_2$. Hledáme tedy $\mathbf{w}$, které maximalizuje:

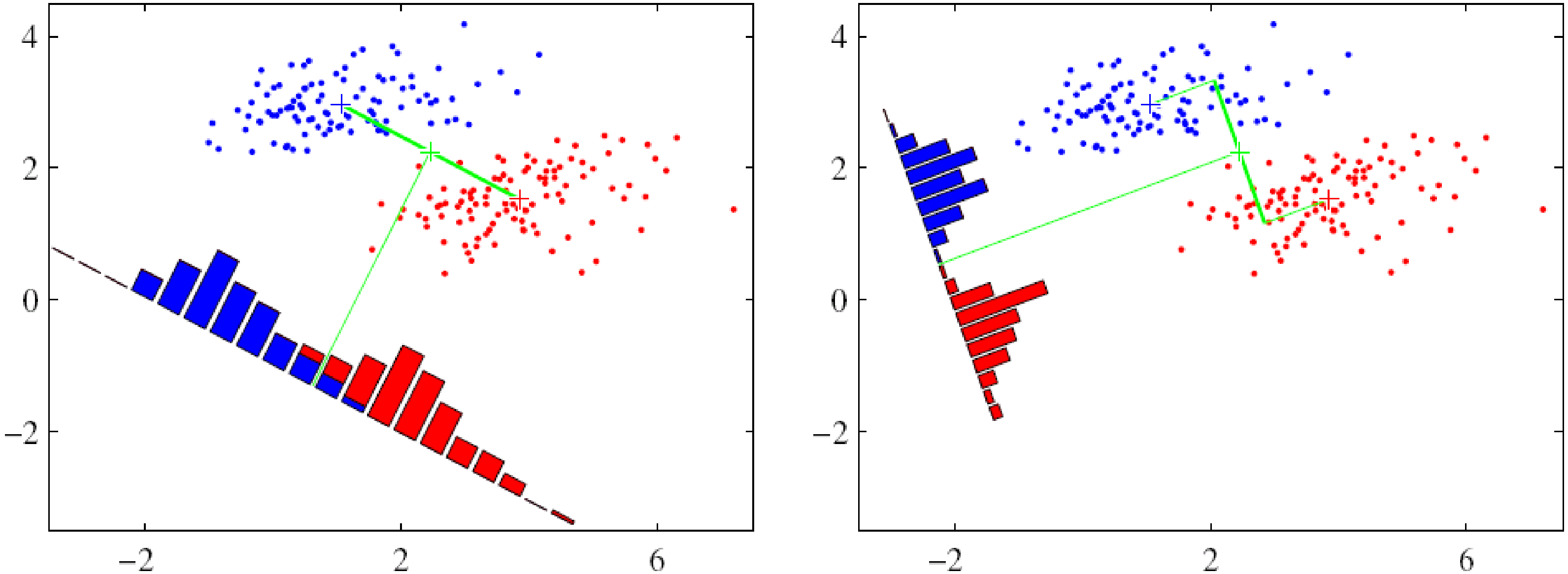
$$m_2 - m_1 = \mathbf{w}^{\mathrm{T}}(\mathbf{m}_2 - \mathbf{m}_1)$$

$$m_k = \mathbf{w}^{\mathrm{T}}\mathbf{m}_k$$



Lineární diskriminační analýza

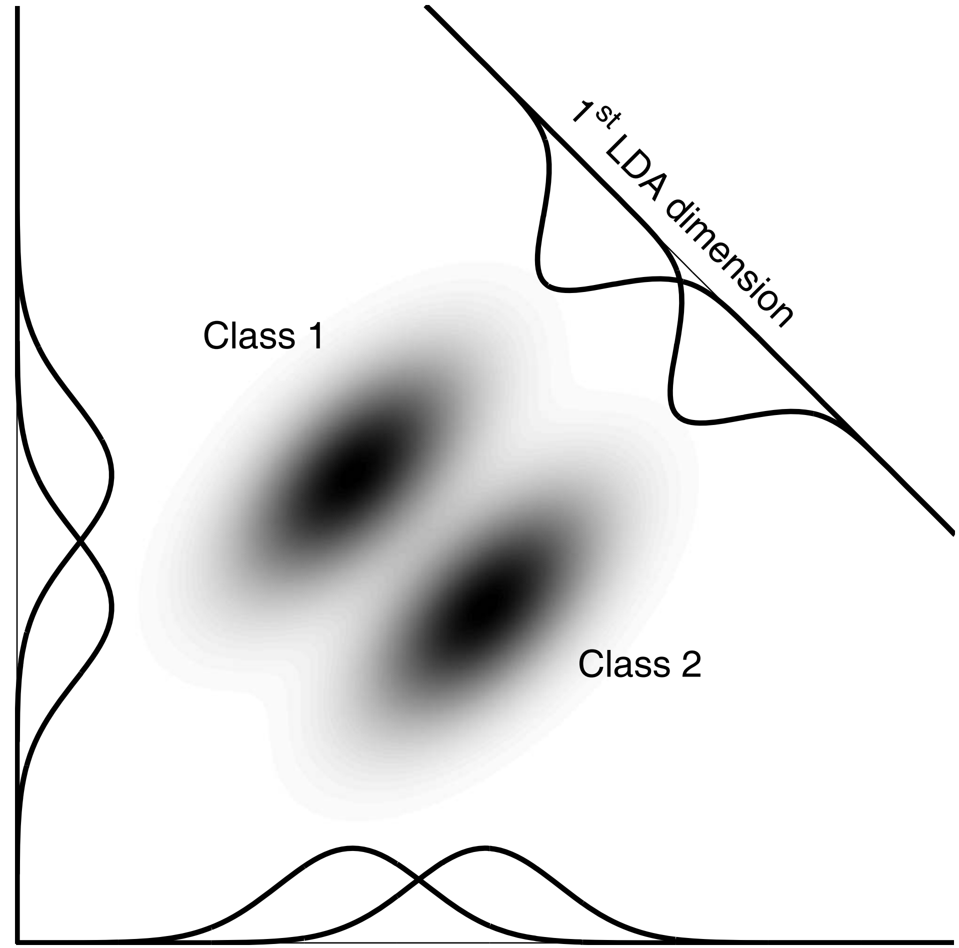
- Lze však najít i lepší směr:



- Snažíme se data promítnout do takového směru, kde
  - Maximalizujeme vzdálenost mezi středními hodnotami tříd
  - Minimalizujeme průměrnou varianci tříd
- Maximalizujeme tedy

$$J(\mathbf{w}) = \frac{(m_2 - m_1)^2}{s_1^2 + s_2^2} \qquad s_k^2 = \sum_{n \in \mathcal{C}_k} (y_n - m_k)^2$$

Lineární diskriminační analýza



Lineární diskriminační analýza

- LDA dimenze dány vlastními vektory matice $\mathbf{\Sigma}_{ac}\mathbf{\Sigma}_{wc}^{-1}$
- Σ_{ac} – kovarianční matice spočítaná se středních hodnot tříd
- Σ_{wc} – průměrná kovarianční matice tříd
- Lze zobecnit pro více tříd – vlastní vektory s největšími vlastními čísly odpovídají směrům ve kterých jsou třídy nelépe separovány
- Pro J tříd bude pouze J-1 vlastních čísel nenulových
- Pokud mají všechny třídy gaussovské rozložení se stejnou kovarianční maticí, LDA transformace transformuje prostor tak, že mohou byt třídy optimálně modelovány gaussovským rozložení s diagonální kovarianční maticí

$$\hat{\mathbf{\Sigma}}_{wc} = \frac{1}{N}\sum_{j=1}^{J} N_j \hat{\mathbf{\Sigma}}^{(j)}$$

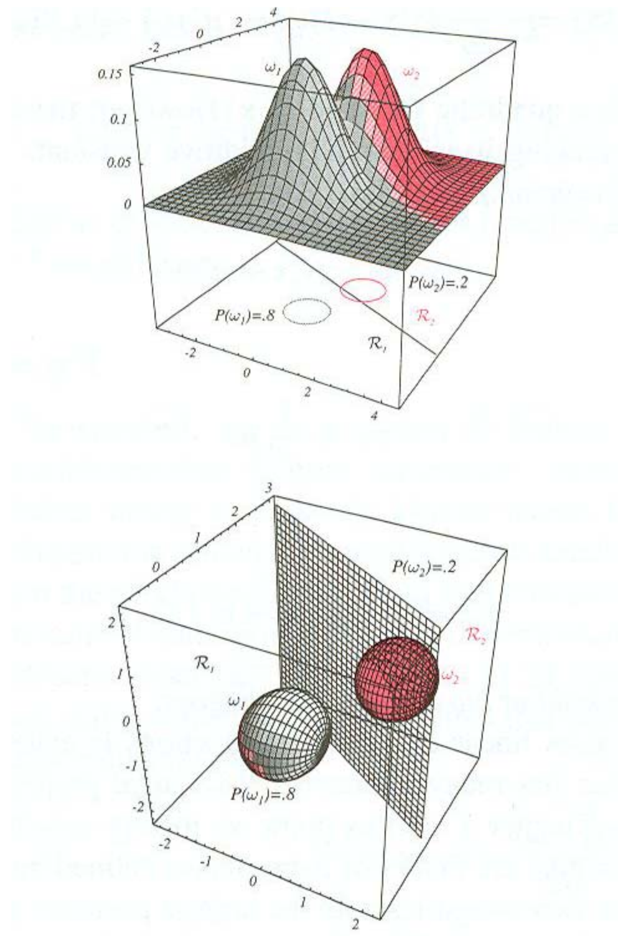
$$\hat{\mathbf{\Sigma}}^{(j)} = \frac{1}{N_j}\sum_{i=1}^{N_j} (\mathbf{x}_i^{(j)} - \hat{\boldsymbol{\mu}}^{(j)})(\mathbf{x}_i^{(j)} - \hat{\boldsymbol{\mu}}^{(j)})^T$$

$$\hat{\mathbf{\Sigma}}_{ac} = \frac{1}{N}\sum_{j=1}^{J} N_j (\hat{\boldsymbol{\mu}}^{(j)} - \hat{\boldsymbol{\mu}})(\hat{\boldsymbol{\mu}}^{(j)} - \hat{\boldsymbol{\mu}})^T$$

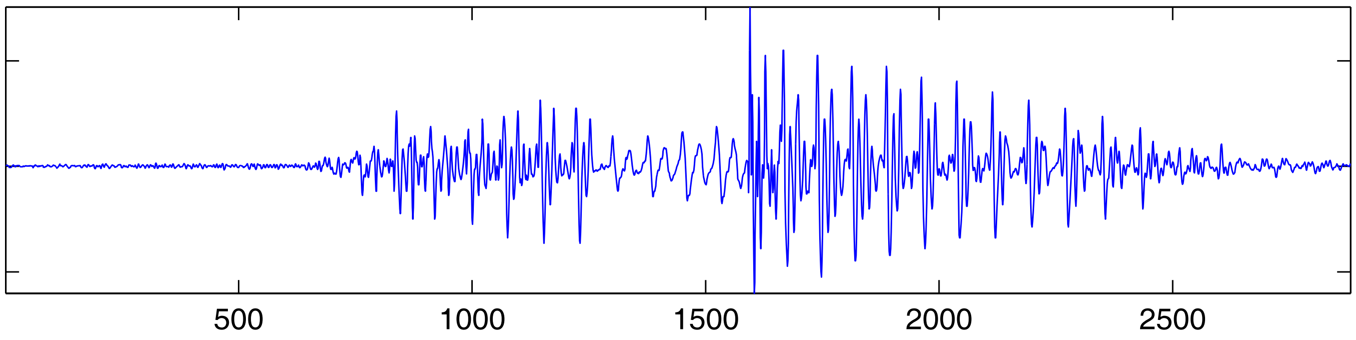
$$\hat{\boldsymbol{\mu}}^{(j)} = \frac{1}{N_j}\sum_{i=1}^{N_j} \mathbf{x}_i^{(j)}$$

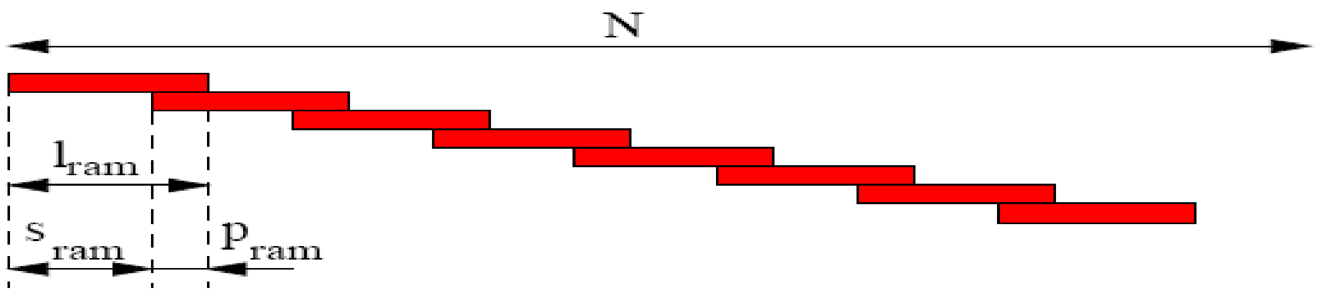
LDA a lineární klasifikátor

Dvě třídy s gaussovským rozložením se stejnou kovarianční matici jsou opravdu optimálně oddělitelné lineárním klasifikátorem (přímkou, rovinou, hyper-rovinou)

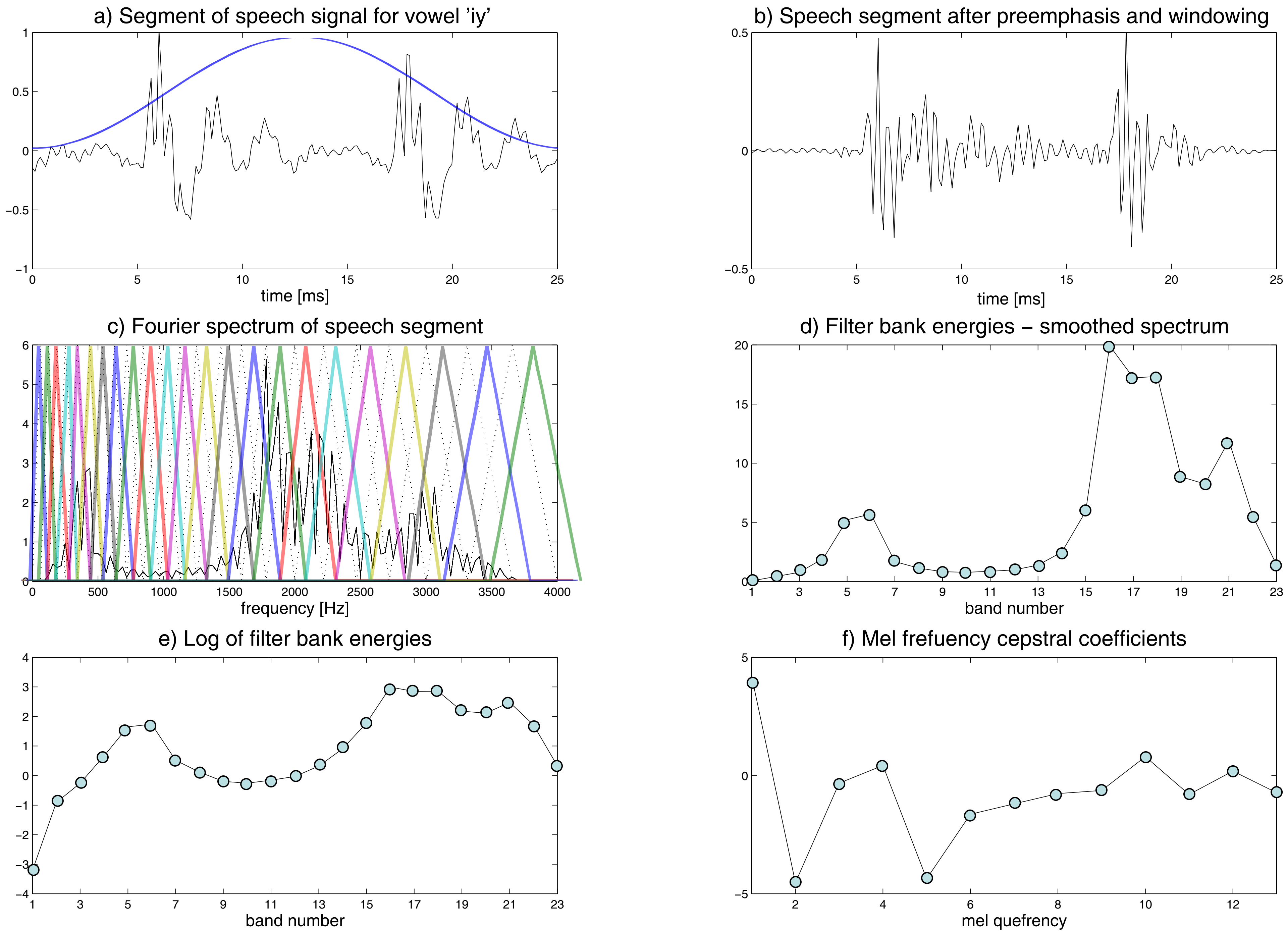


Extrakce příznaku pro řeč - MFCC (Mel frequency cepstral coefficients)

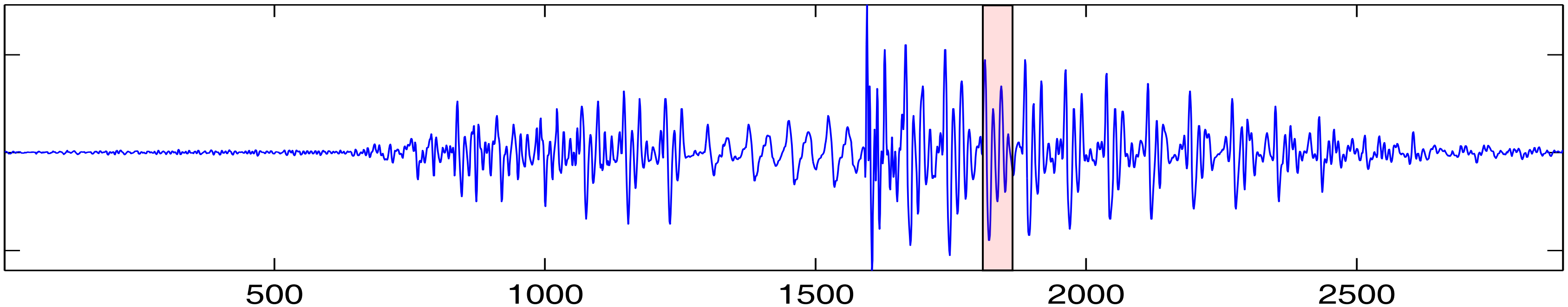




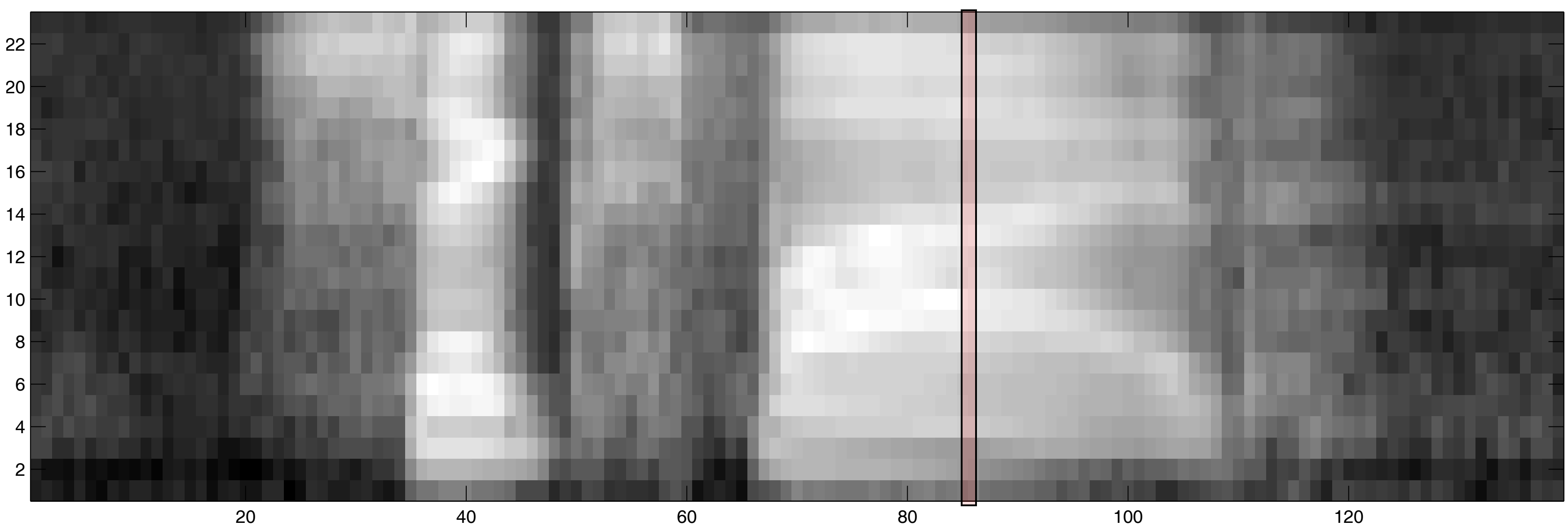
- Nejprve řečový signál rozdělíme do asi 20ms překrývajících se segmentů

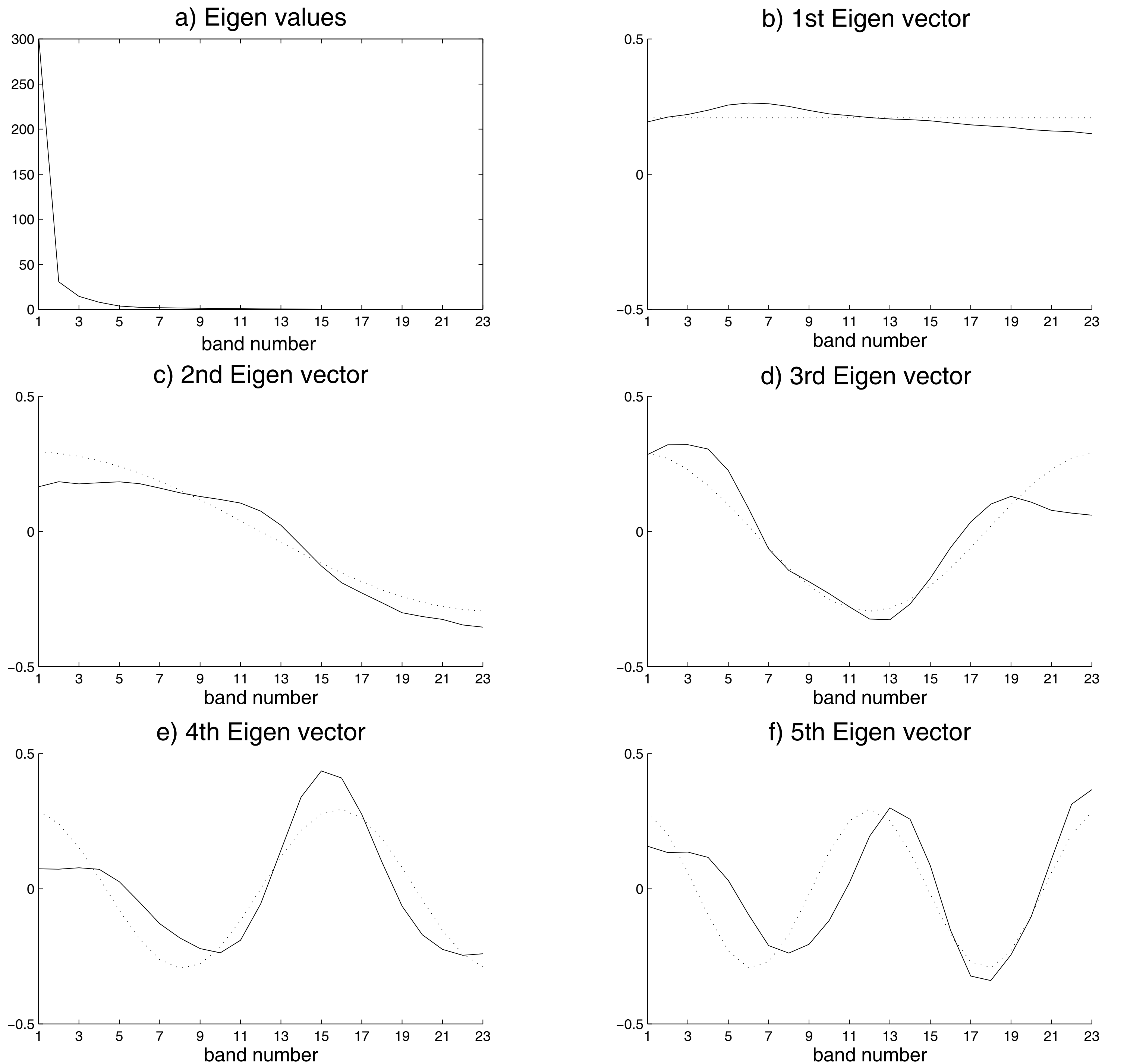

a) Segment of speech signal for vowel 'iy'
time [ms]
b) Speech segment after preemphasis and windowing
time [ms]
c) Fourier spectrum of speech segment
frequency [Hz]
d) Filter bank energies – smoothed spectrum
band number
e) Log of filter bank energies
band number
f) Mel frefuency cepstral coefficients
mel quefrency

Původní signál



Logaritmický vystup z banky filtru – je třeba již jen dekorelovat





a) Eigen values
300
250
200
150
100
50
0
1 3 5 7 9 11 13 15 17 19 21 23
band number
b) 1st Eigen vector
0.5
0
−0.5
1 3 5 7 9 11 13 15 17 19 21 23
band number
c) 2nd Eigen vector
0.5
0
−0.5
1 3 5 7 9 11 13 15 17 19 21 23
band number
d) 3rd Eigen vector
0.5
0
−0.5
1 3 5 7 9 11 13 15 17 19 21 23
band number
e) 4th Eigen vector
0.5
0
−0.5
1 3 5 7 9 11 13 15 17 19 21 23
band number
f) 5th Eigen vector
0.5
0
−0.5
1 3 5 7 9 11 13 15 17 19 21 23
band number

Singular Value Decomposition - SVD

$$A = UDV^T$$

- **A** je jakákoli mxn matice
- **U** je mxn matice kde sloupce jsou ortonormální báze
- **V** je nxn matice kde sloupce jsou ortonormální báze
- **D** je nxn je diagonální matice

- Předpokládejme, že matice A je matice s příznakovými vektory v řádcích s již odečtenou střední hodnotou ➔ **Σ** = **A**^{T}**A**
- Potom z následujících vztahů vyplývá, ze:
  - **V** jsou vlastní vektory **Σ**
  - Diagonála **D** obsahuje odmocniny z vlastních čísel **Σ** (variance ve směrech vlastních vektorů)

$$A^T A = VDU^T UDV^T = VD^2 V^T$$

$$AA^T = UDV^T VDU^T = UD^2 U^T$$