



# ANALÝZA A KLASIFIKACE DAT



**prof. Ing. Jiří Holčík, CSc.**



# IX. METODA FUKUNAGY - KOONTZE



# PROBLÉMY A PODMÍNKY

PCA algoritmus dokáže najít popis obrazů s optimálně redukovaným počtem příznaků s hlediska střední kvadratické odchylky aproximace

- ☑ **disperzní matice**  $\Rightarrow$  preference příznaků s největším rozptylem
- ☑ **autokorelační matice**  $\Rightarrow$  sice lepší situace, ale může být i tak dost bezcenná z hlediska klasifikace

# PROBLÉMY A PODMÍNKY

PCA algoritmus dokáže najít popis obrazů s optimálně redukovaným počtem příznaků s hlediska střední kvadratické odchylky aproximace

- ☑ **disperzní matice**  $\Rightarrow$  preference příznaků s největším rozptylem
- ☑ **autokorelační matice**  $\Rightarrow$  sice lepší situace, ale může být i tak dost bezcenná z hlediska klasifikace

**JAK NA TO?**

# PROBLÉMY A PODMÍNKY

PCA algoritmus dokáže najít popis obrazů s optimálně redukovaným počtem příznaků s hlediska střední kvadratické odchylky aproximace

- ☑ **disperzní matice**  $\Rightarrow$  preference příznaků s největším rozptylem
- ☑ **autokorelační matice**  $\Rightarrow$  sice lepší situace, ale může být i tak dost bezcenná z hlediska klasifikace

## JAK NA TO?

- ☑ výběr příznaků podle charakteristických čísel uspořádaných vzestupně

# PROBLÉMY A PODMÍNKY

PCA algoritmus dokáže najít popis obrazů s optimálně redukovaným počtem příznaků s hlediska střední kvadratické odchylky aproximace

- ☑ **disperzní matice**  $\Rightarrow$  preference příznaků s největším rozptylem
- ☑ **autokorelační matice**  $\Rightarrow$  sice lepší situace, ale může být i tak dost bezcenná z hlediska klasifikace

## JAK NA TO?

- ☑ výběr příznaků podle charakteristických čísel uspořádaných vzestupně
- ☑ v dichotomickém případě – třeba **rozklad podle Fukunagy a Koontze**

# PRINCIP

- ✓ vychází z normalizace autokorelační funkce;
- ✓ výstupem normalizace situace popsaná vztahem

$$\kappa(\mathbf{y}') = \mathbf{E},$$

$\mathbf{E}$  je jednotková matice a  $\mathbf{y}'$  reprezentuje obraz, pro který platí

$$\mathbf{y}' = \mathbf{U} \cdot \mathbf{y},$$

kde  $\mathbf{U}$  je matice normalizační transformace

# PRINCIP

- ☑ pro autokorelační matici transformovaných příznaků platí

$$\mathbf{K}(\mathbf{y}') = \sum_{k=1}^K \mathbf{y}_k \mathbf{y}_k^T \quad \mathbf{y}_k = \sum_{i=1}^K \mathbf{y}_i^T \mathbf{U}_i \quad \mathbf{y}_k^T \mathbf{U}_k = \mathbf{y}_k^T \mathbf{U}_k$$

- ☑ s tím můžeme psát

$$\mathbf{U} \cdot \mathbf{K}(\mathbf{y}') \cdot \mathbf{U}^T = \mathbf{E}$$



# PRINCIP

- ✓ připomínka:

$$\kappa(\mathbf{y}) = \sum_{r=1}^R P(\omega^r) \cdot \int_{\mathcal{Y}^m} \mathbf{y} \cdot \mathbf{y}^T p(\mathbf{y} | \omega^r) d\mathbf{y} = \int_{\mathcal{Y}^m} \mathbf{y} \cdot \mathbf{y}^T p(\mathbf{y}) d\mathbf{y}$$

- ✓ tedy pro dichotomickou situací je

$$\kappa(\mathbf{y}) = P(\omega_1) \cdot \kappa_{\omega_1}(\mathbf{y}) + P(\omega_2) \cdot \kappa_{\omega_2}(\mathbf{y}),$$

kde

$$\kappa_{\omega}(\mathbf{y}) = \int_{\mathcal{Y}^m} \mathbf{y} \cdot \mathbf{y}^T p(\mathbf{y} | \omega) d\mathbf{y}, \quad r = 1, 2$$

je autokorelační matice pro prvky z r-té třídy

# PRINCIP

☑ rovnici  $\mathbf{U} \cdot \kappa(\mathbf{y}) \cdot \mathbf{U}^T = \mathbf{E}$  s tím můžeme psát ve tvaru

$$\mathbf{S}_1 + \mathbf{S}_2 = \mathbf{E},$$

kde

$$\mathbf{S}_r = P(\omega_r) \cdot \mathbf{U} \cdot \kappa_{\omega_r}(\mathbf{y}) \cdot \mathbf{U}^T, r = 1, 2.$$

# PRINCIP

- ✓ pro charakteristická čísla  $\lambda_i^{(1)}$  a charakteristické vektory  $\mathbf{v}_i^{(1)}$  matice  $\mathbf{S}_1$  z definice platí

$$\mathbf{S}_1 \cdot \mathbf{v}_i^{(1)} = \lambda_i^{(1)} \cdot \mathbf{v}_i^{(1)}, \quad i = 1, 2, \dots, m.$$

- ✓ obdobně pro matici  $\mathbf{S}_2$

$$\mathbf{S}_2 \cdot \mathbf{v}_i^{(2)} = (\mathbf{E} - \mathbf{S}_1) \cdot \mathbf{v}_i^{(2)} = \lambda_i^{(2)} \cdot \mathbf{v}_i^{(2)},$$
$$i = 1, 2, \dots, m;$$

odkud po úpravách

$$\mathbf{S}_1 \cdot \mathbf{v}_i^{(1)} = (1 - \lambda_i^{(2)}) \cdot \mathbf{v}_i^{(2)},$$
$$i = 1, 2, \dots, m.$$

# PRINCIP

☑ z toho pak srovnáním je

$$\mathbf{v}_i^{(1)} = \mathbf{v}_i^{(2)}, i = 1, 2, \dots, m \text{ a } \lambda_i^{(1)} = 1 - \lambda_i^{(2)}.$$

Protože z vlastností matic jsou jejich vlastní čísla  $\lambda_i^{(r)} \in \langle 0, 1 \rangle$ ,  $r=1,2$ ;  $i=1,\dots,m$ , jsou vlastní čísla matice  $\mathbf{S}_1$  podle indexu  $i$  uspořádána vzestupně a matice  $\mathbf{S}_2$  sestupně. Tedy nejdůležitější příznaky pro popis jedné třídy jsou současně nejméně důležité pro popis druhé třídy.

☑ báze souřadnicový systém vybíráme z vektorů  $\mathbf{v}_1^{(1)}, \mathbf{v}_2^{(1)}, \dots$  pro třídu  $\omega_1$  a  $\mathbf{v}_m^{(1)}, \mathbf{v}_{m-1}^{(1)}, \dots$  pro třídu  $\omega_2$ .

# PRINCIP

## MATICE **U** NORMALIZAČNÍ TRANSFORMACE

- ✓ bez důkazů  $\mathbf{U} = \mathbf{U}_1 \cdot \mathbf{U}_2,$
- ✓ kde  $\mathbf{U}_1$  představuje matici transformace autokorelační matice  $\kappa(\mathbf{y})$  na matici diagonální  $\kappa(\mathbf{U}_1 \cdot \mathbf{y})$ . To lze provést, když

$$\mathbf{U}_1 = \begin{bmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_m^T \end{bmatrix}$$

kde  $\mathbf{v}_i, i=1, \dots, m$  jsou vlastní vektory autokorelační matice  $\kappa(\mathbf{y})$ .

# PRINCIP

## MATICE U NORMALIZAČNÍ TRANSFORMACE

- ☑ transformovaná matice  $\kappa(\mathbf{U}_1 \cdot \mathbf{y})$  má tvar

$$\kappa(\mathbf{U}_1 \cdot \mathbf{y}) = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_m \end{bmatrix}$$

# PRINCIP

## MATICE $\mathbf{U}$ NORMALIZAČNÍ TRANSFORMACE

- ☑  $\mathbf{U}_2$  převádí výše uvedenou diagonální matici na jednotkovou

$$\mathbf{U}_2 = \begin{bmatrix} 1/\lambda_1 & 0 & \dots & 0 \\ 0 & 1/\lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/\lambda_m \end{bmatrix}$$

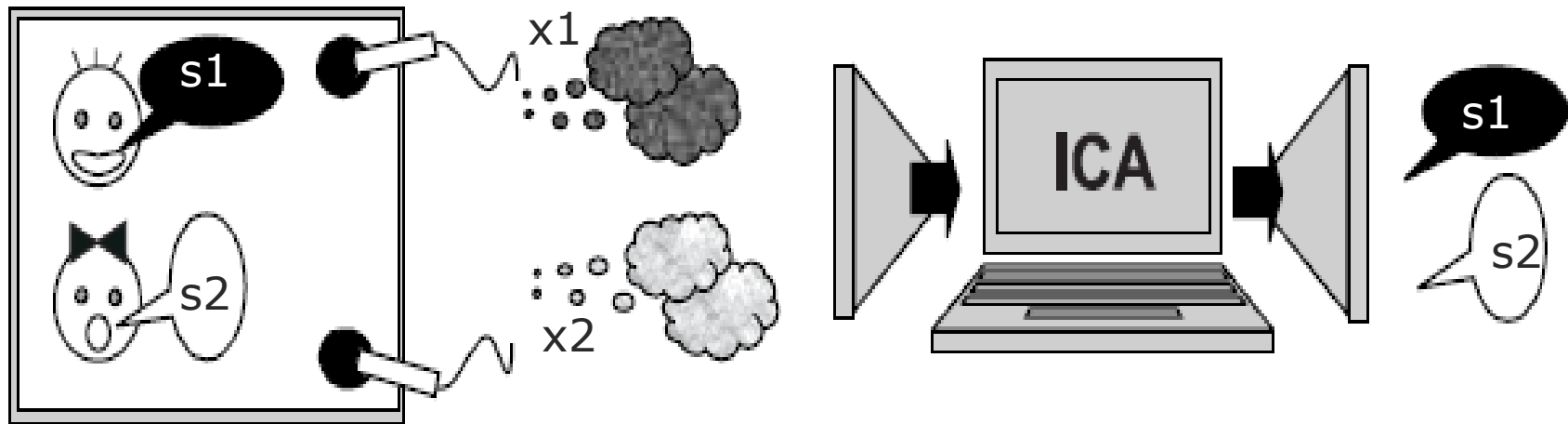


# X. ANALÝZA NEZÁVISLÝCH KOMPONENT





# ANALÝZA NEZÁVISLÝCH KOMPONENT PRINCIP METODY

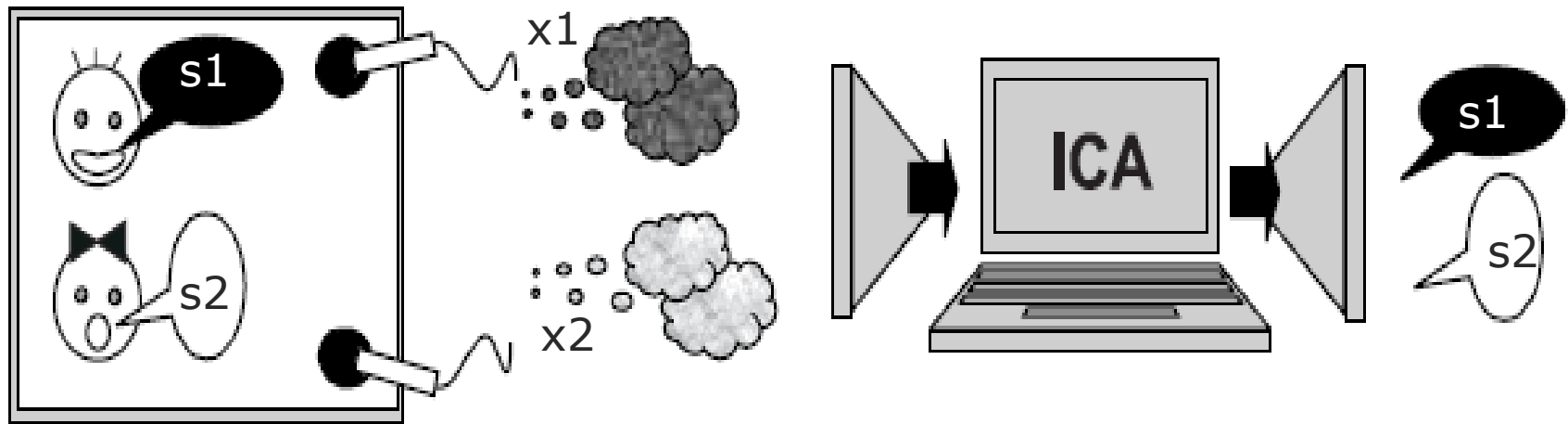


$$x_1(t) = a_{11} \cdot s_1(t) + a_{12} \cdot s_2(t)$$

$$x_2(t) = a_{21} \cdot s_1(t) + a_{22} \cdot s_2(t)$$

Úloha spočívá v nalezení originálních neznámých signálů z jednotlivých zdrojů  $s_1(t)$  a  $s_2(t)$  máme-li k dispozici pouze zaznamenané signály  $x_1(t)$  a  $x_2(t)$ .

# ANALÝZA NEZÁVISLÝCH KOMPONENT PRINCIP METODY



ICA umožňuje určit koeficienty  $a_{ij}$  za předpokladu, že známé signály jsou dány lineárních kombinací zdrojových a za předpokladu statistické nezávislosti zdrojů v každém čase  $t$ .

# ANALÝZA NEZÁVISLÝCH KOMPONENT

## MODEL DAT

- ☑ necht'  $\mathbf{x} = T(x_1, x_2, \dots, x_m)$  je  $m$ -rozměrný náhodný vektor (s nulovou střední hodnotou  $E(\mathbf{x})=0$ ).

$$x_i = a_{i1}^{\text{orig}} \cdot s_1^{\text{orig}} + a_{i2}^{\text{orig}} \cdot s_2^{\text{orig}} + \dots + a_{im}^{\text{orig}} \cdot s_m^{\text{orig}} \\ i = 1, 2, \dots, m$$

nebo

$$\mathbf{x} = \mathbf{A}^{\text{orig}} \cdot \mathbf{s}^{\text{orig}}$$

$\mathbf{s}^{\text{orig}}$  je vektor originálních skrytých nezávislých komponent a  $s_1^{\text{orig}}$  jsou nezávislé komponenty (předpoklad vzájemně statisticky nezávislosti);

$\mathbf{A}^{\text{orig}}$  je transformační matice

# ANALÝZA NEZÁVISLÝCH KOMPONENT

## MODEL DAT

- ☑ definice

$$\mathbf{s} = \mathbf{W} \cdot \mathbf{x},$$

- ☑ cíl:

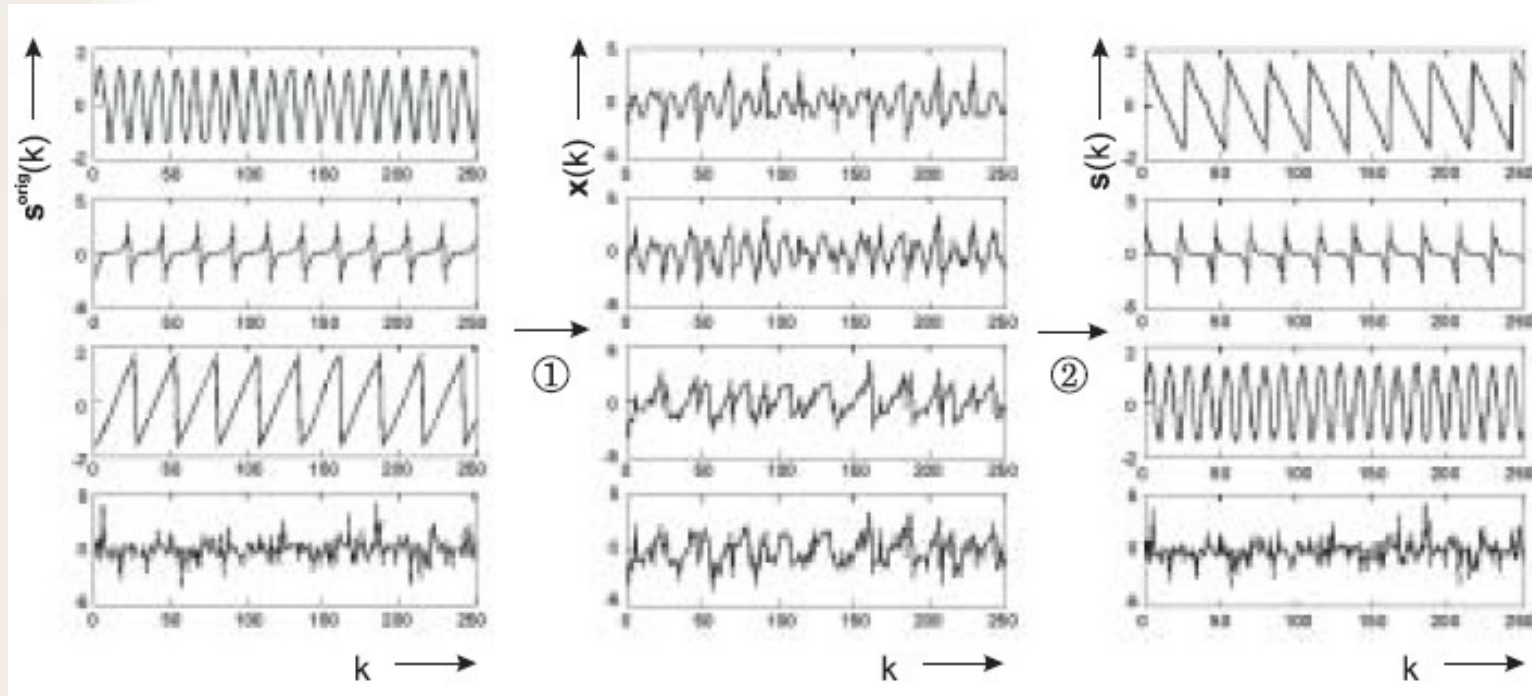
nalézt lineární transformaci (koeficienty transformační matice  $\mathbf{W}$  tak, aby vypočítané nezávislé komponenty  $s_i$  byly vzájemně statisticky nezávislé [ $\mathbf{W} = \mathbf{A}^{-1}$ ]

$$[p(s_1, s_2, \dots, s_m) = p_1(s_1) \cdot p_2(s_2) \dots p_m(s_m)]$$

# ANALÝZA NEZÁVISLÝCH KOMPONENT OMEZENÍ

- ☑ pouze jedna originální nezávislá komponenta může mít normální rozložení pravděpodobnosti (pokud má více zdrojů normální rozložení není ICA schopna tyto zdroje ze vstupních dat extrahovat);
- ☑ pro dané  $m$ -rozměrné obrazové vektory je ICA schopna najít pouze  $m$  nezávislých komponent;
- ☑ nelze obecně určit polaritu nezávislých komponent;
- ☑ nelze určit pořadí nezávislých komponent (?!)

# ANALÝZA NEZÁVISLÝCH KOMPONENT OMEZENÍ



# ODHAD NEZÁVISLÝCH KOMPONENT

- ☑ optimalizace pomocí zvolené optimalizační (účelové, kritériální, objektové) funkce



- nalézt kritériální funkci
- vybrat optimalizační algoritmus

ad a) možnost ovlivnit statistické vlastnosti metody;

ad b) spojitá optimalizační úloha s „rozumnou“ kritériální funkcí – gradientní metoda, Newtonova metoda – ovlivňujeme rychlost výpočtu (konvergenci), nároky na paměť,...

# ODHAD NEZÁVISLÝCH KOMPONENT

## ZÁKLADNÍ ÚVAHA

- ☑ nechť existuje  $m$  nezávislých náhodných veličin s určitými pravděpodobnostními rozděleními (jejich součet za dosti obecných podmínek konverguje s rostoucím počtem sčítanců k normálnímu rozdělení – **centrální limitní věta**);
- ☑ o vektoru  $\mathbf{x}$  (který máme k dispozici) předpokládáme, že vznikl součtem nezávislých komponent  $\mathbf{s}^{\text{orig}}$



jednotlivé náhodné veličiny  $x_i$  mají pravděpodobnostní rozdělení, které je „bližší“ normálnímu než rozdělení jednotlivých komponent  $s_i^{\text{orig}}$



# ODHAD NEZÁVISLÝCH KOMPONENT ZÁKLADNÍ ÚVAHA

- ☑ odhad nezávislých komponent si probíhá tak, že hledáme takové řádkové vektory  $\mathbf{w}_i$  transformační matice  $\mathbf{W}$ , aby pravděpodobnostní rozdělení součinu  $\mathbf{w}_i \cdot \mathbf{x}$  bylo „co nejvíce **nenormální**“



tj. nalézt takovou transformační matici  $\mathbf{W}$ , aby proměnné  $\mathbf{w}_i \cdot \mathbf{x}$  měly pravděpodobnostní rozdělení, které se co nejvíce liší od normálního



potřeba nalézt míru náhodné veličiny, která by mohla být použita pro kvantifikaci míry (podobnost, vzdálenost) nenormality

# ODHAD NEZÁVISLÝCH KOMPONENT POUŽÍVANÉ MÍRY NENORMALITY

- ☑ koeficient špičatosti
- ☑ negativní normalizovaná entropie;
- ☑ aproximace negativní normalizované entropie;

# ODHAD NEZÁVISLÝCH KOMPONENT KOEFIČIENT ŠPIČATOSTI

$$\text{kurt}(s) = E\{s^4\} - 3(E\{s^2\})^2$$

Gaussovo rozložení má koeficient špičatosti roven nule, zatímco pro jiná rozložení (ne pro všechna) je koeficient nenulový.

Při hledání nezávislých komponent hledáme extrém, resp. kvadrát koeficientu špičatosti veličiny  $\mathbf{s} = \mathbf{w}_i \cdot \mathbf{x}$

# ODHAD NEZÁVISLÝCH KOMPONENT KOEFIČIENT ŠPIČATOSTI

## **výhody:**

- ☑ rychlost a relativně jednoduchá implementace;

## **nevýhody:**

- ☑ malá robustnost vůči odlehlým hodnotám (pokud v průběhu měření získáme několik hodnot, které se liší od skutečných, výrazně se změní KŠ a tím i nezávislé komponenty nebudou odhadnut korektně);
- ☑ existence náhodných veličin s nulovým KŠ, ale nenormálním rozdělením;

# ODHAD NEZÁVISLÝCH KOMPONENT NEGATIVNÍ NORMALIZOVANÁ ENTROPIE

(NNE, negentropy)

**Informační entropie** - množství informace  
náhodné veličiny

☑ pro diskrétní náhodnou veličinu  $s$  je

$$H(s) = -\sum_i P(s=a_i) \cdot \log_2 P(s=a_i),$$

kde  $P(s=a_i)$  je pravděpodobnost, že náhodná veličina  $S$  je rovna hodnotě  $a_i$ .

☑ pro spojitou proměnnou platí

$$H(s) = -\int_{-\infty}^{\infty} p(s) \log_2 p(s) ds$$

# ODHAD NEZÁVISLÝCH KOMPONENT NEGATIVNÍ NORMALIZOVANÁ ENTROPIE

- ☑ entropie je tím větší, čím jsou hodnoty náhodné veličiny méně predikovatelné;
- ☑ pro normální rozdělení má entropie největší hodnotu ve srovnání v dalšími rozděleními

## NNE

$$J(s) = H(s_{\text{gauss}}) - H(s),$$

kde  $s_{\text{gauss}}$  je náhodná veličiny s normálním rozdělením

# ODHAD NEZÁVISLÝCH KOMPONENT NEGATIVNÍ NORMALIZOVANÁ ENTROPIE

## **výhody:**

- ✓ přesné vyjádření nenormality;
- ✓ dobrá robustnost vůči odlehlým hodnotám;

## **nevýhody:**

- ✓ časově náročný výpočet  $\Rightarrow$  snaha o vhodnou aproximaci NNE aby byly zachovány její výhody a současně byl výpočet nenáročný

# ODHAD NEZÁVISLÝCH KOMPONENT APROXIMACE NEGATIVNÍ NORMALIZOVANÉ ENTROPIE

- ✓ použití momentů vyšších řádů

$$J(\epsilon) \approx \frac{1}{2} E\{s^3\}^2 + \frac{1}{8} k_{u4}(s)^2$$

kde  $s$  je náhodná veličina s nulovou střední hodnotou a jednotkovým rozptylem

## **nevýhoda:**

- ✓ opět menší robustnost vůči odlehlým hodnotám



# ODHAD NEZÁVISLÝCH KOMPONENT

## APROXIMACE NEGATIVNÍ NORMALIZOVANÉ ENTROPIE

- ☑ Použití tzv. p-nekvadratických funkcí

$$J(s) \approx \sum_{i=1}^p [G_i(s)]^2 - [G_i(s_{\text{gauss}})]^2$$

kde  $k_i > 0$  je konstanta,  $G_i$  jsou šikovně navržené nelineární funkce a  $s_{\text{gauss}}$  je normální náhodná proměnná, která spolu s  $s$  má nulovou střední hodnotu a jednotkový rozptyl.

Je-li použita pouze jedna funkce  $G$ , pak je

$$J(s) \approx [E\{G(s)\}]^2 - [E\{G(s_{\text{gauss}})\}]^2$$

# ODHAD NEZÁVISLÝCH KOMPONENT APROXIMACE NEGATIVNÍ NORMALIZOVANÉ ENTROPIE

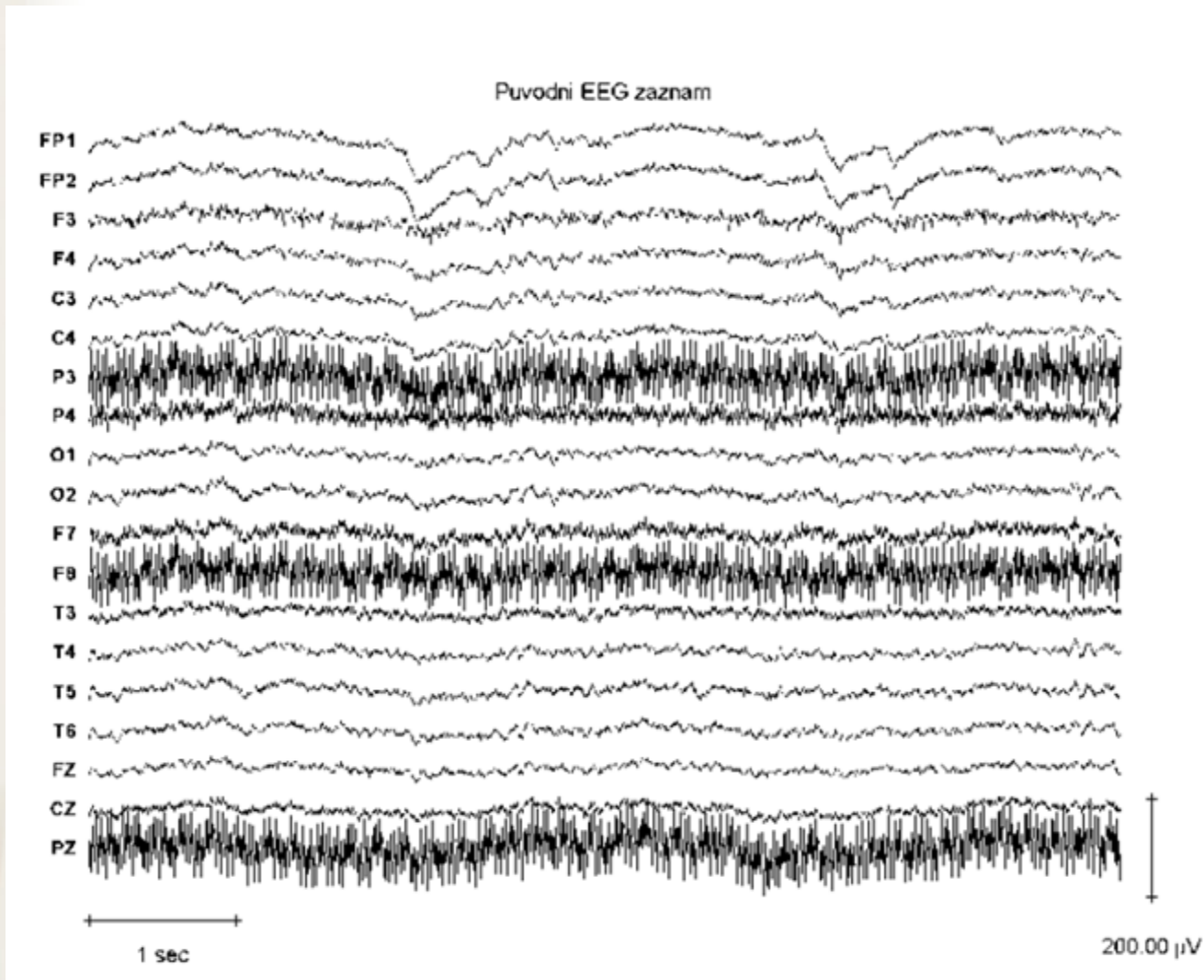
☑ doporučujeme:

$$G_1(\hat{s}) \approx \log(c_1 \hat{s})$$

kde  $a_1 \in \langle 1, 2 \rangle$  nebo

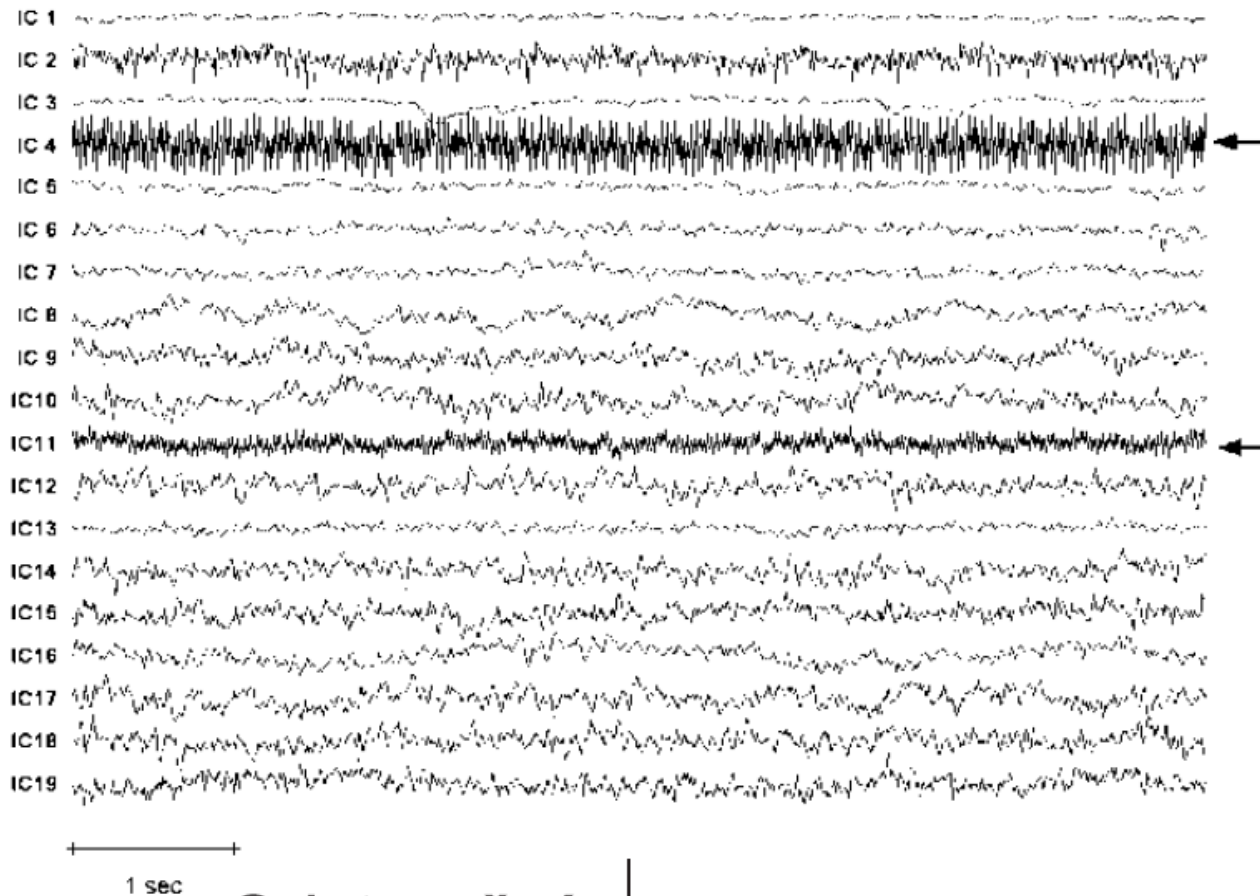
$$G_2(\hat{s}) \approx xp(\hat{s})$$

# ODHAD NEZÁVISLÝCH KOMPONENT PŘÍKLAD POUŽITÍ



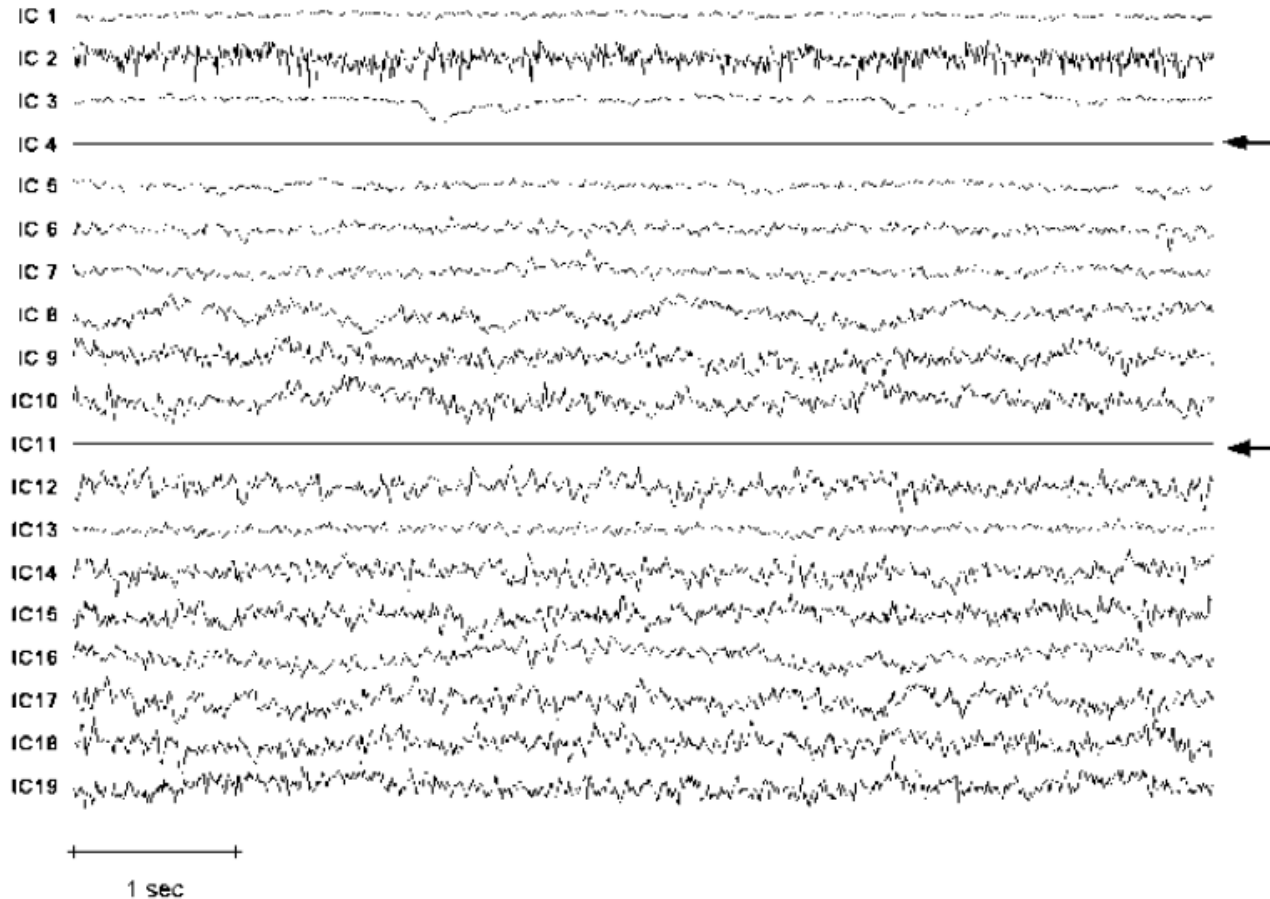
# ODHAD NEZÁVISLÝCH KOMPONENT PŘÍKLAD POUŽITÍ

Nezávisle komponenty (ICs)

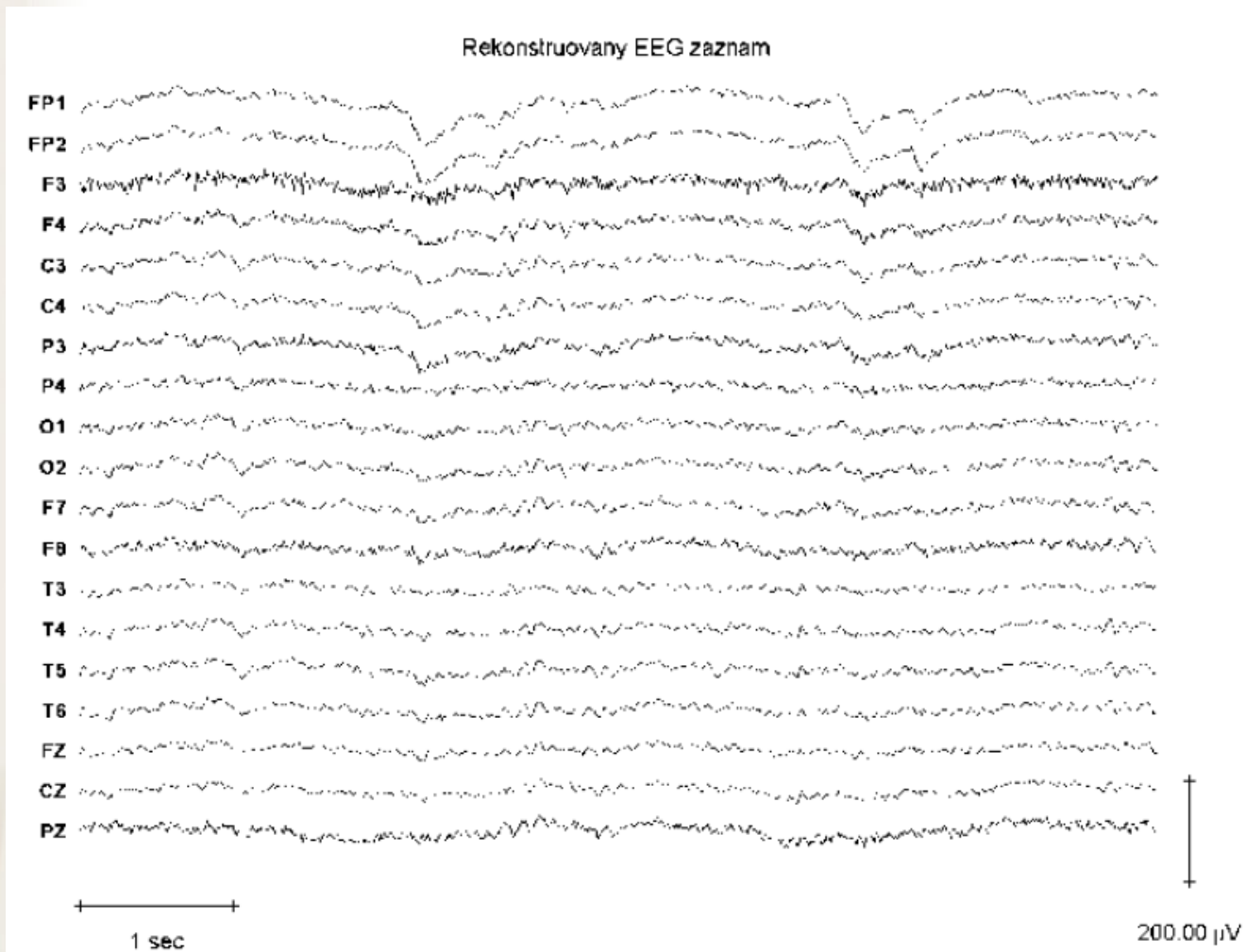


# ODHAD NEZÁVISLÝCH KOMPONENT PŘÍKLAD POUŽITÍ

Nezávislé komponenty (IC4 a IC11 byly odstraněny)



# ODHAD NEZÁVISLÝCH KOMPONENT PŘÍKLAD POUŽITÍ



# ODHAD NEZÁVISLÝCH KOMPONENT PŘÍKLAD POUŽITÍ

