Předmět B2M31DSP/Př. 11

PS

Přednáška11: Slepá separace a dekonvoluce

Obsah

Úvod

Slepá separace a dekonvoluce používají metody datově závislé techniky schopné se učit

Typy učení

- Učení s učitelem $\leftrightarrow$ existují trénovací signály, pro které známe požadovanou odezvu systému
- Učení bez učitele = slepá adaptace $\leftrightarrow$ neexistuje trénovací signál (požadovaná odezva) X existuje soubor pravidel umožňující nastavit specifický vztah mezi vstupem a výstupem

Slepá separace zdrojů (signálů)

Slepá separace zdrojů (signálů) - problém označovaný jako „cocktail party effect“

Předpoklady

- lineární kombinace (= skalární mixáž[1]) $x_i(t)$,
- nezávislých signálů $s_i(t)$

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t)$$

[1] Obecnější model je konvoluční mixáž (přítomno zpoždění mezi zdrojovými signály)

Slepá separace zdrojů (signálů)

- Matice a vektory vztahu

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t)$$

mají pro případ stejného počtu zdrojů generujících $s_i(t)$ a senzorů snímajících $x_i(t)$ tvar:

$\mathbf{s}(t) = [s_1(t), ..., s_m(t)]^T$

$\mathbf{x}(t) = [x_1(t), ..., x_m(t)]^T$

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,m} \end{bmatrix}$$

je neznámá mixážní matice s reálnými prvky

Slepá separace zdrojů (signálů)

Slepá separace zdrojů (signálů)

Cíl a řešení

- cíl $=$ pro daný vektor $\mathbf{x}(t)$ určit vektor $\mathbf{s}(t)$ v režimu bez učitele
- řešení existuje - jednoznačné až na
  - změnu měřítka prvků (signálů) vektoru $\mathbf{s}(t)$
  - změnu pořadí prvků vektoru $\mathbf{s}(t)$

Slepá separace zdrojů (signálů)

Slepá separace zdrojů (signálů)

Řešení

- je-li $\mathbf{A}$ nesingulární, pak existuje matice „demixážní“ matice $\mathbf{W}$
- a pro vektor separovaných signálů platí[2]

$$\mathbf{y} = \mathbf{W}\mathbf{x} = \mathbf{W}\mathbf{A}\mathbf{s} = \mathbf{D}\mathbf{P}\mathbf{s}$$

$\mathbf{D}$ je diagonální matice, $\mathbf{P}$ je permutační matice[3]

Slepá separace je často využívána pro separaci signálů z pole elektrod snímající EEG nebo pro odstranění artefaktů z EEG; byla použita i pro separaci akustických signálů

[2] Označení času t je vypuštěno

[3] Vznikne záměnou sloupců z jednotkové matice

Slepá separace zdrojů (signálů)

Slepá separace zdrojů (signálů)

Jeden z možných postupů je znám jako analýza nezávislých komponent (Independent Component Analysis - ICA) = rozšíření pojmu PCA

- PCA využívá statistiky druhých řádů a pojem nekorelovanost, hledá směry největšího rozptylu dat
- ICA[4] využívá statistiky vyšších řádů a pojem nezávislost, hledá směry největších hodnot statistik vyšších řádů v datech

[4] Jeden z prvních algoritmů využívající statistiky vyšších řádů se nazývá FastICA

Slepá separace zdrojů (signálů)

Slepá separace zdrojů (signálů)

Nekorelovanost a nezávislost

- Nutná a postačující podmínka nezávislosti náhodných veličin $\xi_1, \xi_2, \ldots$ vícerozměrné náhodné veličiny $(\xi_1, \xi_2, \ldots)$ je aby její sdružená distribuční funkce byla rovna součinu marginálních distribučních funkcí, tedy $F(x_1, x_2, \ldots) = F(x_1)F(x_2)\ldots$
- Pokud neznáme hustotu pravděpodobnosti, nahrazujeme jí souborem statistik – momentů nebo kumulantů až do zvoleného řádu
- Potom nekorelovanosti odpovídá nalezení maximální hodnoty vzájemné korelace a její nulování pomocí PCA
- Při použití nezávislosti náhodných veličin (a procesů) hledáme maxima statistik vyšších řádů

  Pozn.: Při použití statistik nelze úplné nezávislosti dosáhnout, proto se snažíme alespoň o co nejvyšší míru nezávislosti

Algoritmy slepé separace bez učitele

A. Algoritmy využívající pojmy z teorie informace

- entropie zdroje s výstupem $\mathbf{x}$

$$h(X) = -E[log(p(\mathbf{x})] = -\int_{-\infty}^{\infty} p(\mathbf{x})log(p(\mathbf{x}))d\mathbf{x}$$

  $p(\mathbf{x})$ je hustota rozdělení
  Cíl: hledá se maximum entropie

- vzájemná informace $I(X;Y)$

$$I(X;Y) = h(X) - h(X|Y)$$

$$I(X;Y) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(\mathbf{x},\mathbf{y})log(p(\frac{p(\mathbf{x}|\mathbf{y})}{p(\mathbf{x})}))d\mathbf{x}d\mathbf{y}$$

  Cíl: hledá se minimální vzáj. informace

Algoritmy slepé separace bez učitele

B. Algoritmy využívající statistiky vyšších řádů[5] - kumulanty, polyspektra

Příklady statistik:
Př.1 $c_2(\tau) = E[x(t)x(t+\tau)]$, $x \in \mathbb{R}$ (korelace - statistika druhého řádu)
Př.2 $c_3(\tau) = E[x(t)x(t+\tau_1)x(t+\tau_2)]$, $x \in \mathbb{R}$ (kumulant 3. řádu)
Př.3 $C_3(\omega_1, \omega_2) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_1=-\infty}^{\infty} c_k(\tau_1, \tau_2) e^{-j(\omega_1\tau_1 + \omega_2\tau_2)}$ (polyspektrum)

C. Algoritmy vyžívající statistik druhých řádů a časové struktury signálu[6]

[5]Např. FastICA hledá maximum kumulantu 4. řádu - špičatosti
[6]Např. SOBI (second-order blind identification)

Stručný popis algoritmu FastICA

FastICA - iterační algoritmus, který je řešením variační úlohy: hledání max. špičatosti za podmínky jednotkového modulu vektoru $|\mathbf{w}| = 1$

Dílčí kroky algoritmu separace signálů

- dána směs $\mathbf{x} = \mathbf{As} + \mathbf{u}$
- centrování $\mathbf{x}_c = \mathbf{x} - \mu_{\mathbf{x}}$
- bělení $\tilde{\mathbf{x}} = \mathbf{D}^{-1/2}\mathbf{V}\mathbf{x}_c$, matice $\mathbf{D}$ a $\mathbf{V}$ jsou získány pomocí PCA
- nalezení maxima špičatosti (kurtosis) $E[\tilde{\mathbf{x}}_c^4]$ iteračním algoritmem; pro jeden vektor $\mathbf{w}$ (směr maxima špičatosti) má tvar

$$\mathbf{w}_{i+1} = E\left[\tilde{\mathbf{x}} g(\mathbf{w}_i^T \tilde{\mathbf{x}})\right] - E\left[\tilde{\mathbf{x}} g'(\mathbf{w}_i^T \tilde{\mathbf{x}})\right] \mathbf{w}_i$$

$$\mathbf{w}_{i+1} = \frac{\mathbf{w}_{i+1}}{||\mathbf{w}_{i+1}||}$$

- vlastní separace $\hat{\mathbf{s}} = \mathbf{W}\tilde{\mathbf{x}}_c$, kde $\mathbf{W}$ je ortogonální matice vytvořená z vektorů $\mathbf{w}$

Používané nelinearity v algoritmu FastIca

Příklady nelineárních funkcí:
nelineární funkce $g(.)$ a její derivace $g^{'}(.)$

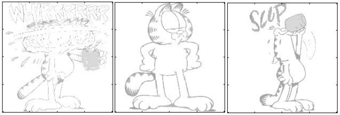
- $g_1(u) = \tanh(ku)$, $1 \le k \le 2$
- $g_2(u) = ue^{-u^2/2}$
- $g_3(u) = u^3$

- $g_1^{'}(u) = k(1 - \tanh^2(ku))$, $1 \le k \le 2$
- $g_2^{'}(u) = (1 - u^2)e^{-u^2/2}$
- $g_3^{'}(u) = 3u^2$

Slepá separace - ilustrace

zdrojové obrázky - nejsou k dispozici .

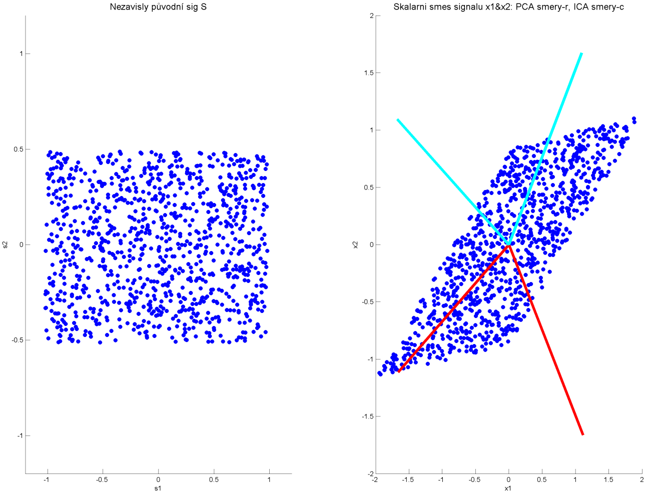
mix obrázků - je k dispozici .



separované obrázky - výstup metody slepé separace - jiné pořadí než zdrojové obrázky .

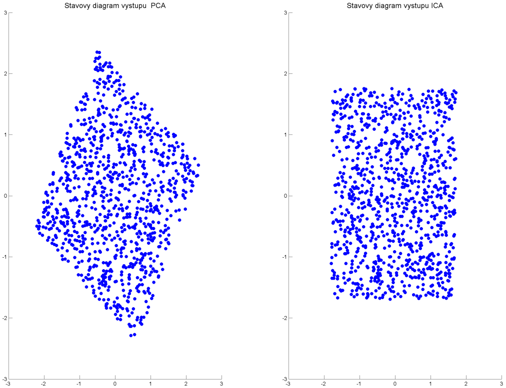
zvýrazněné obrázky .

Slepá separace - ilustrace



Stavový diagram – vlevo: nezávislých signálů - šum s uniformním rozdělením, vpravo: skalární směsi signálů neortogonální maticí; hlavní (PCA) a nezávislé (ICA) směry se liší

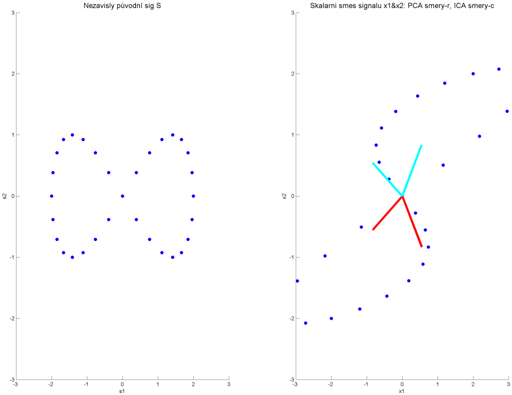
Slepá separace - ilustrace



Stavový diagram – vlevo: výstup PCA - separace není hotová, vpravo: výstup ICA - separace proběhla

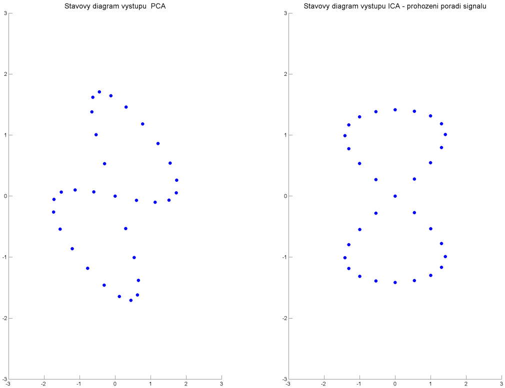
Proč ICA separuje a PCA nikoliv? Obě transformace realizují rotaci souřadné soustavy: PCA do hlavních směrů pomocí maxima rozptylu a ICA do nezávislých směrů pomocí špičatosti. Směry se ovšem liší - viz předchozí obr.

Slepá separace - ilustrace



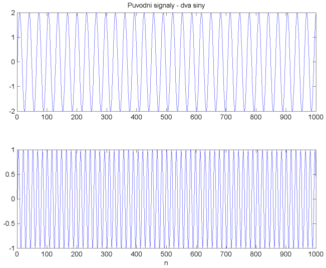
Stavový diagram – vlevo: nezávislých signálů - dva siny s rozdílnými frekvencemi, vpravo: skalární směsi signálů neortogonální maticí; hlavní (PCA) a nezávislé (ICA) směry se opět liší

Slepá separace - ilustrace

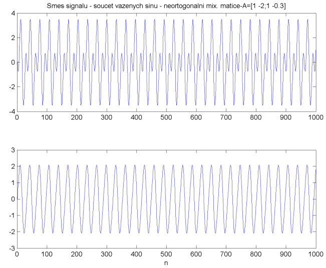


Stavový diagram – vlevo: výstup PCA - separace není hotová, vpravo: výstup ICA - separace proběhla, ale došlo k záměně pořadí signálů - „osmička“ nyní neleží

Slepá separace - ilustrace

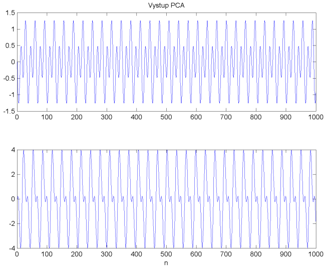


zdrojové signály - nejsou k dispozici .

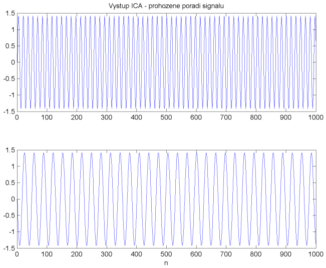


mix signálů - je k dispozici .

Slepá separace - ilustrace



výstup PCA - signály neseparovány .



výstup ICA - signály separovány - záměna pořadí .

Slepá dekonvoluce

Slepá dekonvoluce[7] - popis

- známe **pouze výstup**[8] $x(t)$ LTI soustavy se vstupem $s(t)$ a impulsovou odezvou $h(t)$, na výstupu systému může být přítomen aditivní šum
- pokud $H(f) = \mathcal{F}\{h(t)\}$ přísluší systému s minimální fází[9], pak nalezený inverzní systém $H^{-1}(f)$ je stabilní
- pokud $H(f)$ je systém s neminimální fází, pak je nutné řešit stabilizaci inverzního systému $H^{-1}(f)$

[7] Slepá dekonvoluce nebo nebo též slepá ekvalizace je využívána např. pro ekvalizaci linky pro přenos symbolů v různých typech modulací nebo ostranění neostrostí obrazů, apod.

[8] Viz kepstrální analýza

[9] Nuly přenosu spojitého systému jsou v levé polorovině nebo pro diskrétní systém nuly leží uvnitř jednotkové kružnice

Dekonvoluce a slepá dekonvoluce

A. Metody s učitelem lze pro slepou ekvalizaci přenosové linky použít, ale je nutné použít dodatečné informace

- pokud je linka LTI, pak lze použít Wienerovu filtraci[10] pro dekonvoluci signálu v šumu
- pokud je linka LTV, pak je nutné použít segmentaci signálu[11] nebo lépe adaptivní algoritmy[12]

---

[10] Je nutné znát vstupní signál nebo vzájemnou korelaci vstupu a požadovaného signálu a přenosovou funkci systému, který provádí konvoluci - pak se ovšem nejedná o slepou dekonvoluci ale pouze dekonvoluci

[11] a opět Wienerovu filtraci

[12] Typickým zástupcem je LMS (Least Mean Squares) algoritmus nastavující koeficienty FIR filtru - algoritmus v tomto případě vyžaduje trénovací fázi a pracuje pouze při nízké chybovosti přenosu

Slepá dekonvoluce

B. Metody bez učitele

i) využívají např. Bussgangův teorém[13]

$CE[x(t)x(t+\tau)] = E[f\,(x(t))\,x(t+\tau)]$,

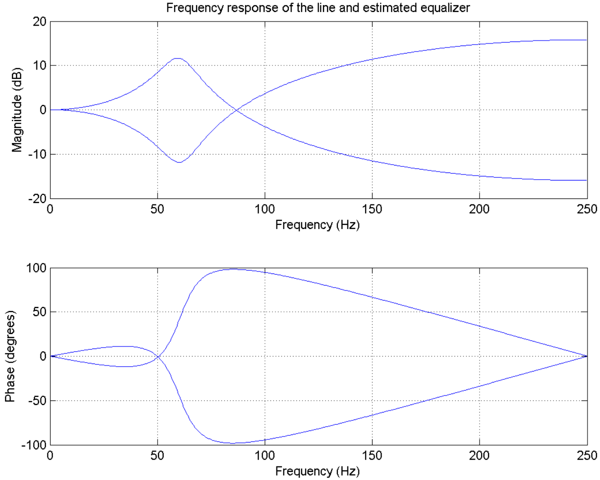
kde $f(t)$ je monotónní nelineární funkce, $x(t)$ je gaussovský stacionární signál s nulovou střední hodnotou, konstanta C je určena funkcí $f(t)$. Tyto metody[14] nastavující koeficienty inverzního filtru pomocí adaptivního algoritmu, jehož výstup je zpracován nelinearitou (např. znaménkovou funkcí nebo obecněji kvantizérem), nevyžadují trénovací fázi a jsou robustnější vůči chybám.

ii) Bez znalosti vstupního signálu a vzájemné korelace vstupu a požadovaného signálu lze využít i jiné přístupy, např. kepstální analýzu aplikovanou buď na signál nebo na parametry signálu, statistiky vyšších řádů - polyspektra, bayesovské metody nebo skryté Markovovy modely.

[13] Teorém říká, že autokorelace gaussovského signálu $x(t)$ se až na konstantu rovná vzájemné korelaci signálu $x(t)$ s jeho nelineárně transformovanou verzí $f(x(t))$

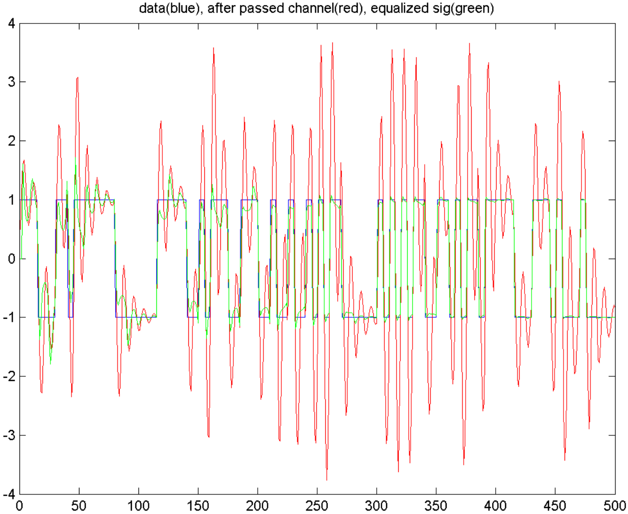
[14] Např. Satův algoritmus nebo "stop and go" algoritmus

Slepá dekonvoluce - ilustrace



Slepá dekonvoluce = inverzní úloha: Frekvenční charakteristika linky a ekvalizéru - modulová (nahoře), fázová (dole) - charakteristiky jsou vzájemně inverzní

Slepá dekonvoluce - ilustrace



Dvoustavový vstupní signál - modrý, po průchodu kanálem - červený, po ekvalizaci - zelený

Slepá separace a dekonvoluce - shrnutí

Slepá separace

- zdrojové signály tvořící směs musí být nezávislé - tyto signály neznáme a snažíme se je odhadnout
- existuje mnoho metod a přístupů - existují též metody pro separaci signálů, kdy je při mixáži přítomné zpoždění mezi zdrojovými signály
- metody využívající statistiky vyšších řádů (např. FastICA) navíc předpokládají, že zdrojové signály nemají gausovské rozdělení hustoty pravděpodobnosti
- metody slepé separace lze použít pro separaci signálů (např. akustických a biologických) i obrazů

Slepá dekonvoluce

- pouze jeden zdrojový signál, který neznáme a snažíme se jej odhadnout pomocí inverzní filtrace z výstupu přenosové linky nebo změřených dat
- pokud není použit pro ekvalizaci dat z přenosové linky FIR filtr, může nastat problém se stabilitou