

PŘEDMĚT B2M31DSP/PŘ. 12

PS

Přednáška 14: Stručný úvod do řídké reprezentace signálů

- 1 ÚVOD
- 2 ZÁKLADNÍ POJMY
- 3 METODA NEJMENŠÍCH ČTVERCŮ
- 4 ŘÍDKÉ ŘEŠENÍ NEDOURČENÉ SOUSTAVY ROVNIC
- 5 ILUSTRACE
- 6 KOMPRIMOVANÉ SNÍMÁNÍ

ÚVOD

Řídká reprezentace signálů (sparse signal processing) a její aplikace komprimované snímání (compressive/compressed sensing, compressive sampling, or sparse sampling) zahrnuje techniky pro efektivní získávání a rekonstrukci signálů vzorkovaných pod Nyquistovou frekvencí. Snížení počtu měření lze provést za předpokladu řídkosti signálu ve vhodné reprezentaci. Matematicky se jedná o nedourčenou soustavu lineárních rovnic se snahou nalézt její řešení s minimálním počtem nenulových koeficientů. Jedná se o poměrně nový obor rozvíjený pracovníky zabývajícími se teorií vlnkové transformace a matematiky¹. Tato oblast, podobně jako teorie vlnkové transformace a bank filtrů, využívá teorii bazí vektorových prostorů konečné a nekonečné dimenze. Setkáme se s pojmy jako je pseudoinverze, podurčené soustavy lineárních rovnic, algoritmy pro optimalizaci, apod. Jedná se o poměrně komplikovanou problematiku, proto se zde změříme pouze na povrchní vysvětlení principů.

Aplikace řídké reprezentace signálů lze nalézt v mnoha oblastech od biomedicíny, radiolokace, komunikačních technologií až po analýzu vibrací.

¹ Jmenujme alespoň některá jména známá z teorie vlnkové transformace, např. matematiky a statistiky Stephane Mallata a Davida Donoho a jeho bývalého doktoranda Emmanuela Candese. V ČR se touto problematikou zabývá doc. Mgr. Pavel Rajmíc, Ph.D. se svým kolektivem

ZÁKLADNÍ POJMY

- Řídké matice a vektory - většina prvků je nulových, např.
 $x = [0, 0, 1, 0, 2, 0, 0, -1]^T$
- Řídká reprezentace signálu představují řešení nedourčené soustavy lineárních rovnic, které má většinu proměnných nulových
- Používané normy² l_0 : $p=0$, l_1 : $p=1$, l_2 : $p=2$

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p},$$

popř.

$$\|x\|_2^2 = \sum_{i=1}^n |x_i|^2$$

²[https://en.wikipedia.org/wiki/Norm_\(mathematics\)#p-norm](https://en.wikipedia.org/wiki/Norm_(mathematics)#p-norm)

METODA NEJMENŠÍCH ČTVERCŮ - LS

Soustava lineárních rovnic

$$\mathbf{Ax} = \mathbf{y},$$

rozměr matice \mathbf{A} je $[m,n]$, vektor \mathbf{y} je měření, \mathbf{x} je hledané řešení

METODA NEJMENŠÍCH ČTVERCŮ

Typy soustav lineárních rovnic

- pře určená soustava rovnic $m > n$, hledáme $\underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{Ax} - \mathbf{y}\|_2^2$

$$\mathbf{x} = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{y}$$

- nedourčená³ soustava rovnic $m < n$, hledáme $\underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_2^2$ tak, aby
 $\mathbf{Ax} = \mathbf{y}$

$$\mathbf{x} = \mathbf{A}^H (\mathbf{A} \mathbf{A}^H)^{-1} \mathbf{y}$$

- jednoznačně určená soustava rovnic $m = n$ nevede na LS

³https://cs.wikipedia.org/wiki/Nedour%C4%8Den%C3%A1_soustava_rovnic

METODA NEJMENŠÍCH ČTVERCŮ

V případě šumu je řešení nedourčené soustavy

$$\mathbf{x} = (\mathbf{A}^H \mathbf{A} + \lambda \mathbf{I})^{-1} \mathbf{A}^H \mathbf{y}$$

Tato metoda je známá jako Tikhonova regularizace⁴, kde parametr λ je potřeba určit v závislosti na typu úlohy

Metoda nejmenších čtverců založená na l_2 normě neposkytuje řídké řešení, proto se postupně prosadily metody používající normu l_0 a l_1 , které řídké řešení poskytují.

⁴https://en.wikipedia.org/wiki/Tikhonov_regularization

MOTIVACE PRO ŘÍDKÉ ŘEŠENÍ

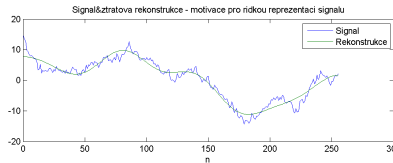
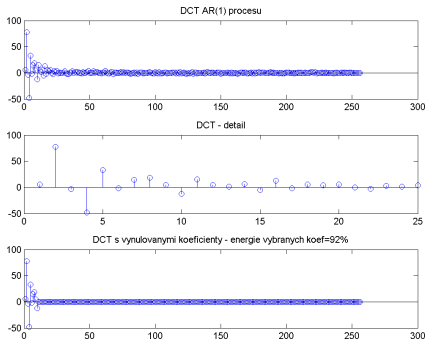
V minulých lekcích jsme se seznámili s metodami ztrátové komprese, kdy se pro rekonstrukci signálu používá pouze několik největších transformovaných hodnot (např. několik spektrálních čar u DFT).

Víme, že PCA, DFT, DCT i vlnková transformace signál nahradí vektorem hodnot, kde pouze několik z nich má velkou hodnotu. Pokud malé hodnoty anulujeme a poté provedeme rekonstrukci signálu, bude rozdíl mezi původním a komprimovaným signálem malý. Toho využívají metody ztrátové komprese.

MOTIVACE PRO ŘÍDKÉ ŘEŠENÍ

Ilustrace ztrátové komprese AR(1) signálu pomocí DCT

1. vlevo shora dolů: DCT koeficienty signálu, detail, DCT s vynulovanými složkami
 2. vpravo: původní signál a rekonstruovaný signál pomocí DCT koeficientů ze spodního obrázku
- vlevo - je patrné, že charakter signálu zůstal zachován, potlačeny jsou rychlé změny s malou amplitudou



MOTIVACE PRO ŘÍDKÉ ŘEŠENÍ

Zatímco DCT koeficienty nepředstavují řádký vektor, tak DCT s vynulovanými koeficienty už lze pokládat za řádký vektor

Nicméně popsáný způsob nepředstavuje použitelné řešení – toho se dosáhne pouze tak, že soustavu

$$\mathbf{Ax} = \mathbf{y}$$

řešíme za podmínky, aby norma l_0 řešení soustavy \mathbf{x} , tedy $\|\mathbf{x}\|_0$, byla minimální, přičemž platí, že $\|\mathbf{x}\|_0$ je počet nenulových prvků vektoru \mathbf{x}

ŘÍDKÉ ŘEŠENÍ NEDOURČENÉ SOUSTAVY ROVNIC

Jak jsme uvedli, na rozdíl od metody nejmenších čtverců (norma l_2), používáme pro nalezení řídkého řešení normu l_0 – hledáme tedy řídké řešení ve tvaru⁵

$$\underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_0, \text{ tak, aby } \mathbf{Ax} = \mathbf{y}$$

Norma l_0 není konvexní norma a její výpočetní nároky jsou neúměrně vysoké, proto se l_0 nahrazuje (aproximuje) normou⁶ l_1

$$\underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_1, \text{ tak, aby } \mathbf{Ax} = \mathbf{y}$$

Tento typ úlohy bývá v literatuře označován jako Basic Pursuit (BP)

⁵Matice \mathbf{A} se nazývá slovník, \mathbf{y} je měření (signál) a \mathbf{x} je řídké řešení soustavy

⁶Platí $\|\mathbf{x}\|_1 = |x_1| + |x_2| + \dots + |x_m|$, a tedy všechny prvky přispívají k výsledné normě (velikosti) vektoru stejnou vahou, a proto bude mít výsledný hledaný vektor \mathbf{x} pouze málo nenulových prvků – toto nás opravňuje k náhradě normy l_0 normou l_1

ŘÍDKÉ ŘEŠENÍ NEDOURČENÉ SOUSTAVY ROVNIC

Při přítomnosti šumu (chyb měření) se upouští od podmínky rovnosti a připouští se malá chyba

$$\underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_1, \text{ tak, aby } \|\mathbf{Ax} - \mathbf{y}\|_2 \leq \delta$$

Tento typ úlohy⁷ bývá v literatuře označován jako Least Absolute Shrinkage and Selection Operator (LASSO)

⁷Lze nalézt i jiné, i když ekvivaletní, formulace této úlohy.

ŘÍDKÉ ŘEŠENÍ NEDOURČENÉ SOUSTAVY ROVNIC

Obě úlohy vedou na iterativní výpočet, neboť nelze získat soustavu rovnic pro jednorázové řešení

Pro BP lze použít lineární programování, kdy hledáme

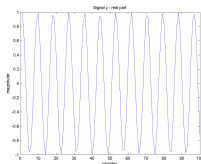
$$\min \sum_{i=1}^n t_i, \text{ tak, aby}$$

$$x_i < t_i \ \& \ x_i > -t_i, \ \& \ \mathbf{Ax} = \mathbf{y}, \ x_i, t_i \in \mathbb{R}$$

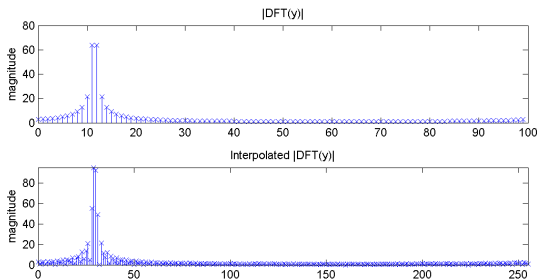
Pro řešení LASSO⁸ úlohy existují iterativní optimalizační algoritmy využívající jiné techniky, např. relaxační algoritmy (např. Iterative Reweighted Least Squares) nebo hledání „nejvýznamnějších“ sloupců matice \mathbf{A} (např. Matching Pursuit), a další, kombinující metodu největšího spádu s prahováním

⁸I pro řešení BP problému

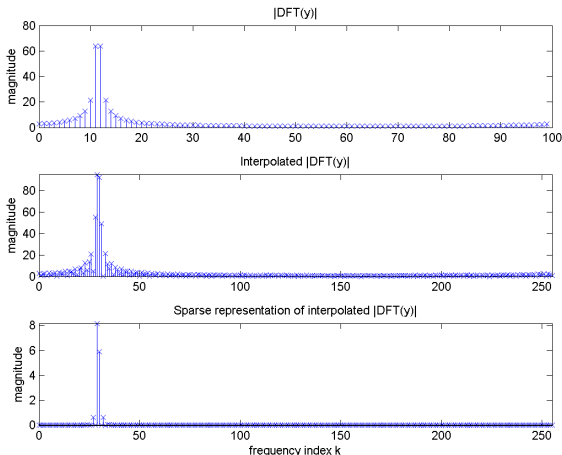
SIGNÁL A JEHO REPREZENTACE POMOCÍ DFT



Mnoho nenulových vzorků v čase - teoreticky jedna hodnota ve spektru = řídká reprezentace
ale v důsledku proskakování ovšem získáme více hodnot - tedy nikoliv řídkou reprezentaci

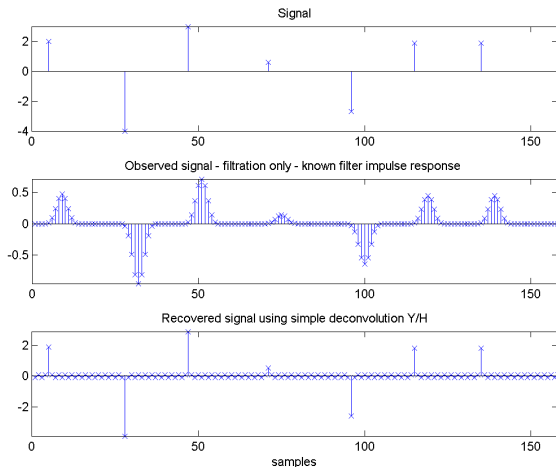


SIGNÁL A JEHO REPREZENTACE - l_1 A l_2 NORMA



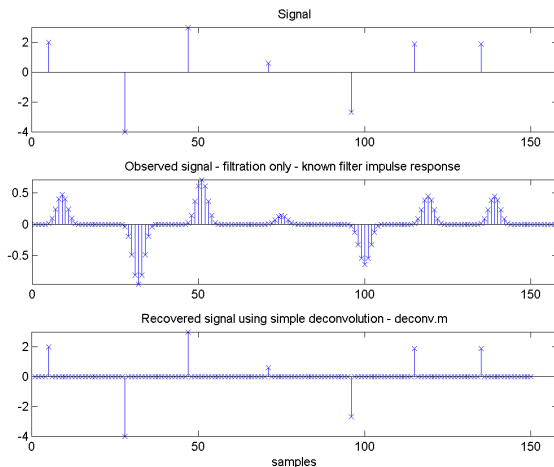
1. nahoře-DFT, uprostřed - interpolace ve spektru - odpovídá metodě nejmenších čtverců, tedy normě l_2 (nenulových čar je ještě více než u neinterpolovaného spektra)
2. dole - řídké spektrum získané pomocí normy l_1 obsahuje nejmenší počet čar - došlo k redukci prosakování, což je ve spektrální analýze velmi žádoucí

DEKONVOLUCE PROSTOU INVERZÍ BEZ ŠUMU



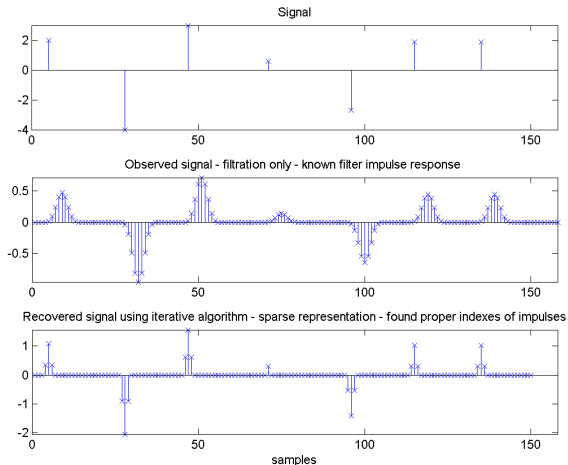
Obrázky: (1) řidký signál $s[n]$ je (2) filtrován dolní propustí = konvoluce s imp. odezvou DP $h[n]$, (3) dekonvoluce signálu $y[n]$ ve spektrální oblasti prostou inverzí $S = \frac{Y}{H}$

DEKONVOLUCE PROSTOU INVERZÍ BEZ ŠUMU



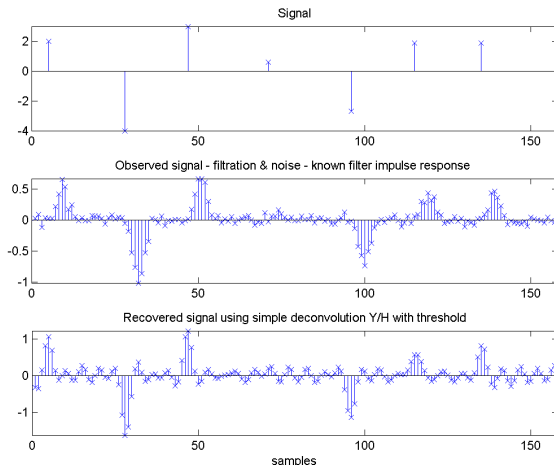
Rekurentní výpočet dekonvoluce bez prahování pomocí funkce `deconv.m`

DEKONVOLUCE POMOCÍ l_1 -NORMY - ÚLOHA BP



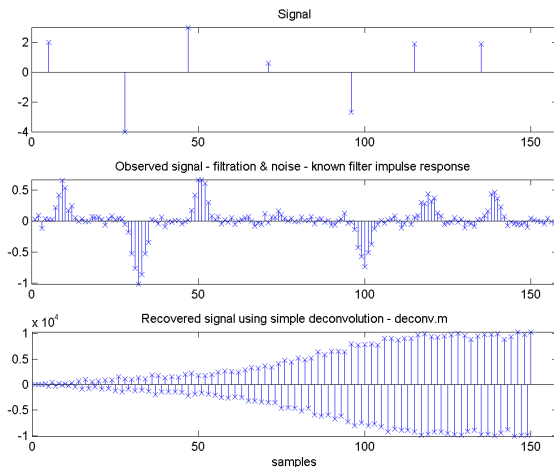
Iterativní algoritmus pro realizaci dekonvoluce s použitím normy l_1 - BP úloha

DEKONVOLUCE PROSTOU INVERZÍ S PRAHOVÁNÍM PRO KONVOLUCI SE ŠUMEM



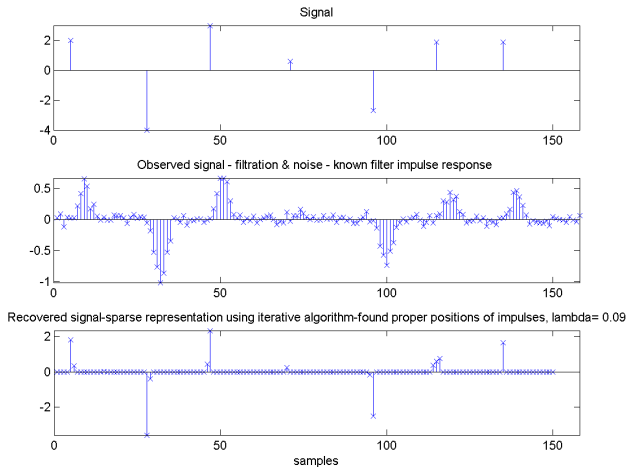
(1) řídký signál $s[n]$ je (2) filtrován dolní propustí = konvoluce s imp. odezvou DP $h[n]$ a následně je přičten šum, (3) dekonvoluce signálu $y[n]$ ve spektrální oblasti s prahováním

DEKONVOLUCE PROSTOU INVERZÍ - REKURENTNÍ ALGORITMUS BEZ PRAHOVÁNÍ



Rekurentní výpočet dekonvoluce - deconv.m :-)

DEKONVOLUCE POMOCÍ l_1 -NORMY



Iterativní algoritmus pro realizaci dekonvoluce s použitím normy l_1 - LASSO problém. Je zřejmé, že postup s použitím l_1 normy je odolný vůči šumu a vede na řídké řešení

KOMPRIMOVANÉ SNÍMÁNÍ

Jedna z nejvíce rozšířených aplikací řídké reprezentace signálů je komprimované snímání, kdy za splnění jistých podmínek, snímáme pouze tolik vzorků, kolik je potřeba, i když to zdánlivě odporuje vzorkovacímu teorému⁹. Vtip je v tom, že místo jednoho vzorku¹⁰ používáme postup, kdy za vzorek signálu považujeme lineární kombinaci více vzorků signálu. K tomu je ovšem třeba použít takové postupy, kdy matice **A** (slovník) je tvořena součinem tří matic, z nichž jedna představuje matici vhodné ortogonální transformace¹¹ (např. DCT), druhá je permutační matice vybírající náhodně řádky z matice¹² náhodných čísel s daným rozdělením (např. s Gaussovým rozdělením).

⁹Kromě zde uvedeného příkladu 1-D signálu, existuje možnost získat vzorky scény jednopixelovou kamerou a následně provést rekonstrukci scény. Kvalita obrazu je sice nižší, ale výhody jednopixelového systému výrazně převažují

¹⁰Získaného součinem Diracova impulsu se signálem spojeným v čase

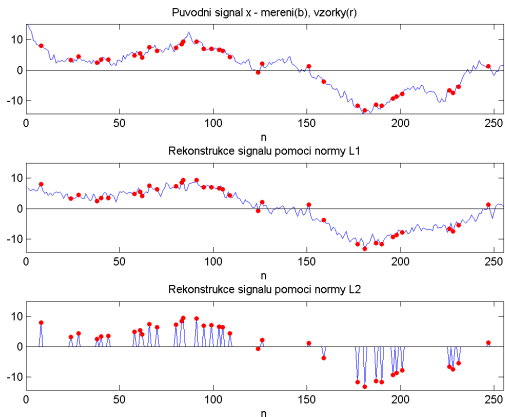
¹¹Nicméně existují postupy, kdy lze pomocí speciálních algoritmů získat slovník **A** přizpůsobený dané množině signálů, což zaručuje poměrně vysokou přesnost rekonstrukce signálu

¹²Tuto matici lze někdy vynechat.

KOMPRIMOVANÉ SNÍMÁNÍ

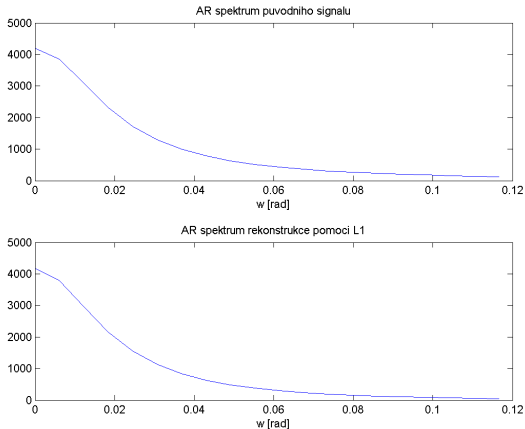
Příklad rekonstrukce signálu získaného komprimovaným snímáním pro slovník **A** tvořený dvěma maticemi

signál modře-vzorky červeně: 1. obr. nahoře: originální AR(1) signál & vzorky=měření získané komprimovaným snímáním, 2. obr. uprostřed: rekonstrukce pomocí l_1 -normy ze vzorků, 3. obr. dole: rekonstrukce pomocí l_2 -normy - toto řešení selhává, neboť neposkytuje řídkou reprezentaci signálu



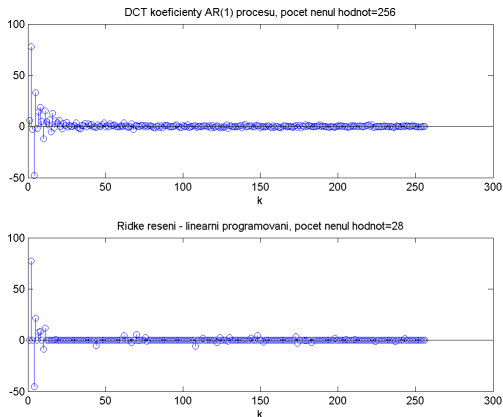
KOMPRIMOVANÉ SNÍMÁNÍ

Porovnání PSD originálního a rekonstruovaného signálu



KOMPRIMOVANÉ SNÍMÁNÍ

Porovnání DCT koeficientů získaných pomocí DCT transformace AR1 signálu a prvků řídkého vektoru x získaného z AR1 signálu pomocí BP úlohy (použito lineární programování)=řídká reprezentace měření získaná řešením nedourčené soustavy rovnic pomocí normy l_1 : místo 256 nenulových DCT koeficientů BP dává pouze 28 nenulových koeficientů, což umožňuje mnohem úspornější reprezentaci signálu



KOMPRIMOVANÉ SNÍMÁNÍ

Příklad rekonstrukce signálu získaného komprimovaným snímáním pro slovník **A** tvořený třemi maticemi

signál modře-vzorky červeně: 1. obr. nahoře: originální AR(1) signál & vzorky=měření získané komprimovaným snímáním, 2. obr. uprostřed: rekonstrukce pomocí l_1 -normy ze vzorků, 3. obr. dole: rekonstrukce pomocí l_2 -normy - toto řešení selhává, neboť neposkytuje řídkou reprezentaci signálu

