

Projekt hrvatske zaklade za znanost 9.01/232:
"Analiza nelinearnih komponenata s primjenama u kemometriji i patologiji"
Institut Ruđer Bošković, 17. prosinca 2012.



Analiza nelinearnih komponenata s primjenama u kemometriji i patologiji

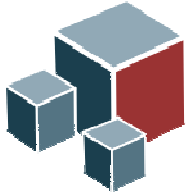
Ivica Kopriva

Institut Ruđer Bošković
17. prosinca 2012.

e-mail: ikopriva@irb.hr ikopriva@gmail.com

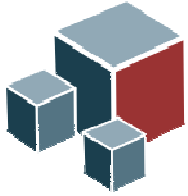
Web: <http://www.lair.irb.hr/ikopriva/projects/hrzz.html>.

"Mišljenja, nalazi i zaključci ili preporuke navedene u ovom materijalu označavaju mišljenje autora i ne reflektiraju nužno stajališta Hrvatske zaklade za znanost."



Sažetak prezentacije

- Informacije o projektu
- Ciljevi projekta
- Linearan bezmemorijski problem slijepog razdvajanja signala (SRS) i analiza rijetkih komponenata (ARK)
- Nelinearan bezmemorijski problem SRS
- Demonstracije i primjene



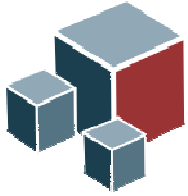
Informacije o projektu

Naziv projekta: Analiza nelinearnih komponenata s primjenama u kemometriji i patologiji.

Voditelj projekta: Dr.sc. Ivica Kopriva, ZLAIR, IRB

Trajanje projekta: 1. 12. 2012. godine do 30. 11. 2015. godine (3 godine).

Iznos financiranja: 997,599,70 KN.



Istraživači i konzultanti na projektu

Dr.sc. Marijana Popović-Hadžija, ZMM, IRB

Marko Filipović, dipl. inž, znanstveni novak, ZLAIR, IRB

Ante Jukić, dipl. inž, znanstveni novak, ZLAIR, IRB

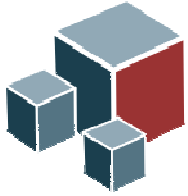
Lidija Brkljačić, dipl. inž, znanstvena novakinja, ZOKiB, IRB

Dr.sc. Ivanka Jerić (konzultant), ZOKiB, IRB

Dr.sc. Mirko Hadžija (konzultant), ZMM, IRB

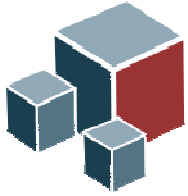
Dr.sc. Gorana Aralica (konzultant), KBC Dubrava

Prof. dr.sc. Šimun Križanac (konzultant), KBC Dubrava



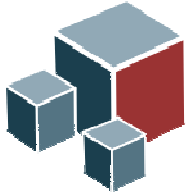
Ciljevi projekta

- Razvoj i komparativna analiza metoda za nelinearno slijepo razdvajanje rijetkih signala (NSRS), odnosno analizu nelinearnih rijetkih komponenata (ANRK). Izdvajanje komponenata iz nelinearnih smjesa će biti utemeljeno na eksplicitnim i implicitnim preslikavanjima u Hilbertov prostor s reprodukcijским jezgrama (engl. RKHS) čime je omogućeno rješenje problema NSRS bez strukturnog ograničenja na vrstu nelinearnosti. To predloženi koncept čini praktično važnim.



Ciljevi projekta

- Primjena NSRS/ANRK u kemometriji na izdvajanje analita (čistih komponentata) iz nelinearnih modela spektara mase kemijskih reakcija. Pri tome je broj analita (znatno) veći od broja smjesa (pod-određeno SRS). Kemijske reakcije će predstavljati modele biološki važnih reakcija: formiranja i hidrolize estera te formiranje peptidne veze. Izdvajanje analita iz smjese spektara je važno u dijagnostici oboljenja (identifikaciji biomarkera), praćenju toksičnosti ksenobiotika, praćenju odgovora živih organizama na promjene u prehrani i okolišu,...



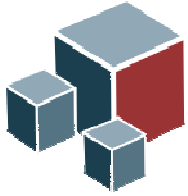
Ciljevi projekta

- Spektrometrija mase je izabrana zbog velike važnosti za modernu metabolomiku i proteomiku. Štoviše, spektrometrija mase visoke rezolucije s velikom vjerojatnosti ispunjava zahtjev međusobne rijetkosti analita (analiti se ne preklapaju u domeni *masa/naboj*):

I. Kopriva, I. Jerić, *Anal. Chem.* 82, 1911-1920, 2010.

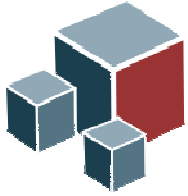
I. Kopriva, I. Jerić, *J. Mass. Spec.* 44, 1378-1388, 2009.

Time je omogućeno izdvajanje većeg broja analita od broja mješavina.



Ciljevi projekta

- Primjena NSRS/NARK u dekompoziciji (segmentaciji) višespektralne slike histopatološkog uzorka u anatomske smislene komponente, a bez uporabe kontrastnog agensa. Formulacija problema segmentacije slike u RKHS ima za učinak poboljšanje kontrasta između objekata (tkiva) prisutnih u slici. Moguća eliminacija uporabe kontrastnog agensa je praktično važna za histopatologiju kroz skraćivanje histološke analize i primjenu tijekom operativnog procesa (intraoperativna analiza tkiva).



Ciljevi projekta

- Koncept segmentacije RGB slike histopatološkog uzorka živčanih vlakana (*nervus ischiadicus*) i bijele pulpe u tkivu slezene bez kontrastnog reagensa je preliminarno demonstriran u radu:

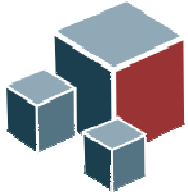
I. Kopriva, M. Hadžija, M. Popović-Hadžija, M. Korolija, A. Cichocki, *American Journal of Pathology*, vol. 179, No. 2, 547-553, 2011,

na segmentaciji RGB fluorescentne slike tumora kože:

I. Kopriva, A. Peršin, *Medical Image Analysis*, vol. 13, 507-518, 2009.

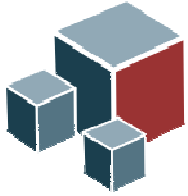
te na segmentaciji MR slike mozga:

I. Kopriva, A. Cichocki. *Lecture Notes Computer Science* 6365, pp. 490-497, 2010.



Ciljevi projekta

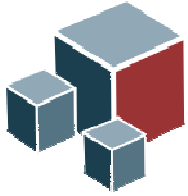
- U citiranim radovima NARK je temeljena na eksplicitnim preslikavanjima drugog i trećeg reda (dobiju se faktorizacijom polinomske jezgre drugog/trećeg reda). Na projektu se planira primjena naprednije verzije NARK za segmentaciju slike tkiva jetre bez kontrastnog agensa, a za vizualizaciju patoloških promjena uzrokovanih poremećenim metabolizmom lipida i ugljikohidrata (šećerna bolest tipa I i II).



Ciljevi projekta

- Očekuje se da razvijene metode NARK daju dobre rezultate i u izdvajanju površinskih raspodjela biopolimera (posebice proteina na površini stanica i tkiva) iz slikovnih spektara mase bioloških uzoraka, a bez uporabe specifičnih reagensa. To je važno za analizu ekspresije proteina u tkivu sisavaca te posljedično za patološke studije i analizu biomarkera. Ovo očekivanje temelji se na dobrim rezultatima u izdvajanju površinske i prostorne raspodjele mekanih (nisko kontrastnih) tkiva (jetra) iz višefazne CT slike abdomena i raspodjele bazičnih tkiva (siva i bijela masa, cerebrospinalna tekućina) iz višespektralne MR slike mozga:

I. Kopriva, X. Chen, A. Jukić, J. Yao, "Nonlinear Nonnegative Sparse Component Analysis for Unsupervised Segmentation of Multi-phase Computed Tomography and Multi-spectral Magnetic Resonance Images," *IEEE Transactions on Image Processing* – pod recenzijom.



Linearan bezmemorijski problem SRS i ARK

Problem:

$$\mathbf{X}=\mathbf{A}\mathbf{S} \quad \mathbf{X} \in \mathbb{R}_{0+}^{N \times T}, \mathbf{S} \in \mathbb{R}_{0+}^{M \times T}, \mathbf{A} \in \mathbb{R}_{0+}^{N \times M} \quad (1)$$

Cilj: procijeniti \mathbf{A} i \mathbf{S} koristeći samo podatke \mathbf{X} , pri čemu je broj izvornih signala M nepoznat i može biti manji, jednak ili veći od broja smjesa N . T označava broj realizacija (npr. piksela u RGB slici ili m/z koordinata u spektrima mase).

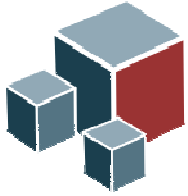
SRS problem (1) ima beskonačno rješenja, tj. $\mathbf{Y}=\mathbf{B}\mathbf{S}$ gdje je \mathbf{B} proizvoljna invertibilna kvadratna matrica $M \times M$. To je zato što je $\mathbf{X}=\mathbf{A}\mathbf{S}=\mathbf{A}\mathbf{B}^{-1}\mathbf{B}\mathbf{S}$. Smisljena rješenja su karakterizirana sa $\mathbf{B}=\mathbf{P}\mathbf{\Lambda}$ gdje su \mathbf{P} permutacijska, a $\mathbf{\Lambda}$ diagonalna matrica, tj.: $\mathbf{Y} \cong \mathbf{P}\mathbf{\Lambda}\mathbf{S}$.

Za $\mathbf{B}=\mathbf{P}\mathbf{\Lambda}$ potrebna su ograničenja na \mathbf{S} i/ili \mathbf{A} .

A. Hyvarinen, J. Karhunen, E. Oja, "Independent Component Analysis," John Wiley, 2001.

A. Cichocki, S. Amari, "Adaptive Blind Signal and Image Processing," John Wiley, 2002.

P. Comon, C. Jutten, editors, "Handbook of Blind Source Separation," Elsevier, 2010.



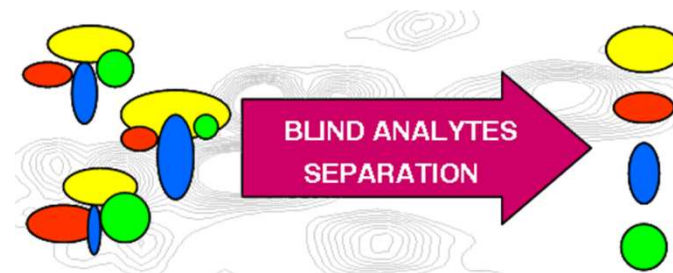
Linearan bezmemorijski problem SRS i ARK

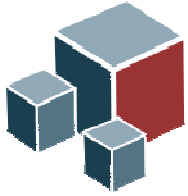
Interpretacija modela:

$$\mathbf{X} = \mathbf{A}\mathbf{S} \quad \mathbf{X} \in \mathbb{R}_{0+}^{N \times T}, \mathbf{S} \in \mathbb{R}_{0+}^{M \times T}, \mathbf{A} \in \mathbb{R}_{0+}^{N \times M} \quad (1)$$

U kemometriji (spektroskopiji NMR ili spektrometriji mase) redci od \mathbf{X} su spektri uzoraka (bioloških, itd.), stupci od \mathbf{A} su koncentracijski profili analita koje se nalaze u spektrima \mathbf{X} , a redci od \mathbf{S} su spektri analita koje se nalaze u spektrima \mathbf{X} .

Problem SRS sastoji se u izdvajanju spektara analita (i njihovih koncentracija?) koristeći samo spektre mješavina \mathbf{X} :





Linearan bezmemorijski problem SRS i ARK

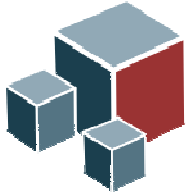
Izdvajanje analita je važno u studijama identifikacije biomarkera u različitim vrstama bioloških uzoraka. Te komponente ukazuju na perturbacije u biološkim sustavima uslijed bolesti ili tretmana lijekovima.

To počiva na pretpostavci da se stanje bolesti (ili neko drugo stanje koje se razlikuje od referentnog neporemećenog stanja sustava) reflektira u promjenama i koncentracijama analita (metabolita) prisutnih u biološkom uzorku:

Madsen, R., Lundstedt, T., Trygga, J. *Anal. Chim. Acta* 659, 23-33, 2010.

Međutim, stotine komponenata mogu biti prisutne u spektru biološkog uzorka. Na primjer, u GC-MS studiji uzorka lista biljke *Arabidopsis thaliana* identificirano je čak 497 kemijskih komponenata:

Jonsson, P.; Johansson, A. I.; Gullberg, J.; Trygg, J.; Jiye, A.; Grung, B.; Marklund, S.; Sjöström, M.; Antti, H.; Moritz, T. *Anal. Chem.*, 77, 5635-5642. 2005.



Linearan bezmemorijski problem SRS i ARK

Realno je očekivati, posebice kod spektrometara niske rezolucije, da će se spektri analita prisutnih u kompleksnim (biološkim) uzorcima preklapati.

Probabilistička analiza pokazuje da će približno 20% komponentata prisutnih u kromatogramu ko-eluirati, što onda implicira preklapanje spektara mase:

Rosenthal, D. *Anal. Chem.* 54, 63-66, 1982.

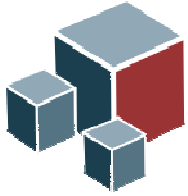
Davies, J. M.; Giddings, J. C. *Anal. Chem.* 55, 418-424, 1983.

Dakle, problem SRS iz spektara mase bioloških uzoraka karakteriziran je nenegativnim preklapajućim (koreliranim) izvornim signalima (analitima), pri čemu je njihov broj (znatno?) veći od broja spektara/uzoraka. Radi se o nenegativnom pododređenom problemu SRS.

U općem slučaju biokemijski procesi nisu linearni:

Yuh, C.H., Bolouri, H., Davidson, E.H. *Science* 279, 1896-1902, 1998.

Kako izabrati ograničenja (regularizaciju)?



Linearan bezmemorijski problem SRS i ARK

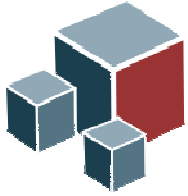
Interpretacija modela:

$$\mathbf{X}=\mathbf{A}\mathbf{S} \quad \mathbf{X} \in \mathbb{R}_{0+}^{N \times T}, \mathbf{S} \in \mathbb{R}_{0+}^{M \times T}, \mathbf{A} \in \mathbb{R}_{0+}^{N \times M} \quad (1)$$

U slikovnoj spektroskopiji (primjer: višespektralna/RGB slika) redci od \mathbf{X} su vektorizirane slike na pojedinim valnim duljinama (npr. crvenoj, zelenoj i plavoj boji), stupci od \mathbf{A} su spektralni profili objekata (tkiva, organa) koji se nalaze u slici \mathbf{X} , a redci od \mathbf{S} su površinske (ili prostorne) raspodjele intenziteta objekata (tkiva, organa) koji se nalaze u slici \mathbf{X} .

Model (1) je uz ekvivalentnu interpretaciju primjenjiv na druge vrste ko-registrirane višekanalne slike: hiperspektralna slika, višefazni CT, višespektralni MR, funkcijska MR slika, slikovna spektrometrija mase, DTI, OCT, multimodalna slika dobivena fuzijom: PET/CT, ...

Problem SRS sastoji se u nenadziranoj segmentaciji (dekompoziciji) slike \mathbf{X} u anatomske (morfološke) smislene komponente: raspodjele intenziteta objekata prisutnih u slici \mathbf{X} .



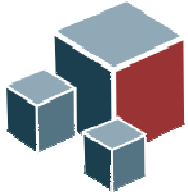
Linearan bezmemorijski problem SRS i ARK

Dakle, problem nenadzirane segmentacije višespektralne slike je problem nenegativnog SRS.

U slikovnim sustavima visoke prostorne rezolucije (RGB slika histopatoloških uzoraka) u svakom slikovnom elementu uglavnom je prisutan samo jedan objekt. To znači da su komponente (raspodjele intenziteta objekata) međusobno rijetke, tj. da se ne preklapaju.

U slikovnim sustavima poput funkcijske ili višespektralne MR učinak parcijalnog volumena implicira prisutnost više objekata u istom slikovnom elementu:

R.K.S. Kwan, A.C. Evans, and G.B. Pike, *IEEE Trans. Med. Imag.*, vol. 18, no. 11, pp. 1085-1097, 1999.



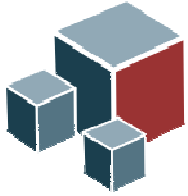
Linearan bezmemorijski problem SRS i ARK

Opravdano je postaviti pitanje da li je linearni model višekanalne slike adekvatan? Kod višespektralne MR slike, prostorna neuniformnost intenziteta (uzrokovana nehomogenošću magnetskog polja) daje nelinearan (multiplikativan) odnos između snimljenog i stvarnog intenziteta:

B. Belaroussi, J. Milles, S. Carme, Y. M. Zhu, H. Benoit-Cattin, *Med. Image Anal.*, vol. 10, no.2, pp. 234-246, 2006.

Teorija prijenosa topline zračenjem (engl. radiative transfer theory) opisuje prijenos energije pri interakciji fotona sa objektima na sceni. To rezultira fizikalno utemeljenim nelinearnim modelom koji implicira rješavanje ekstremno složenog i slabo-uvjetovanog inverznog problema s *a priornim* znanjima o parametrima scene koje je vrlo teško (nemoguće) dobiti. Parametarski modeli, npr. *kernel*-metode, omogućavaju rad sa nelinearnim modelima uz razuman stupanj računske složenosti:

H. Kwon, N. M. Nasrabadi, *IEEE Trans. Geosc. and Remote Sens.*, vol. 43, pp. 388-397, 2005.



Linearan bezmemorijski problem SRS i ARK

Kako izabrati ograničenja (regularizaciju) za rješenje linearnog nenegativnog pododređenog problema SRS?

$$\mathbf{X}=\mathbf{A}\mathbf{S} \quad \mathbf{X} \in \mathbb{R}_{0+}^{N \times T}, \mathbf{S} \in \mathbb{R}_{0+}^{M \times T}, \mathbf{A} \in \mathbb{R}_{0+}^{N \times M} \quad (1)$$

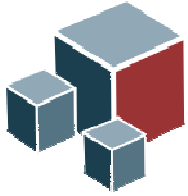
Metode analize nezavisnih komponentata (ANK) podrazumijevaju da su izvorni signali: (i) ne-Gaussovi; (ii) statistički nezavisni; (iii) da je njihov broj manji ili jednak broju smjesa:

P. Common, "Independent Component Analysis – a new concept ?", Signal Processing, 36(3):287-314, 1994.

A. Hyvarinen, J. Karhunen, E. Oja, "Independent Component Analysis," John Wiley, 2001.

Ova ograničenja nisu primjenjiva na problem SRS u analizi spektara mase bioloških uzoraka čak i ako je linearan model (1) prihvatljiv (?)

Takav komentar se, u načelu, može primijeniti i na analizu RGB slike histopatološkog uzorka.



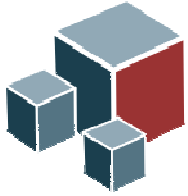
Linearan bezmemorijski problem SRS i ARK

Metode faktorizacije nenegativnih matrica (FNM) prepostavljaju: (i) nenegativnost spektralnih/koncentracijskih profila i komponentata; (ii) omogućavaju eksplicitno ili implicitno ograničenje rijetkosti izvornih signala; (iii) da je broj komponentata manji, jednak ili veći od broja smjesa:

A. Cichocki, R. Zdunek, A. H. Phan, S. Amari, Nonnegative Matrix and Tensor Factorizations-Applications to Exploratory Multi-way Data Analysis and Blind Source Separation, John Wiley, 2009.

Ova ograničenja su primjenjiva na problem SRS u analizi spektara mase bioloških uzoraka, te analizu RGB slike u patologiji.

FNM metode procjenjuju matrice \mathbf{A} i \mathbf{S} iz matrice podataka \mathbf{X} istovremeno. U slučaju nelinearnih modela, FNM metode od interesa postaju direktno primjenjive nakon transformacije modela u RKHS.



Linearan bezmemorijski problem SRS i ARK

Od posebnog interesa za projekt su FNM metode za izdvajanje djelomično koreliranih komponentata:

N. Gillis, F. Glineur, *Pattern Recognition*, vol. 43, pp. 1676-1687, 2010.

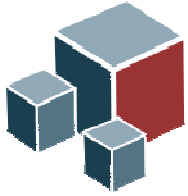
C. F. Caifa, A. Cichocki, *Neural Computation*, vol. 21, pp. 3487-3518, 2009.

R. Peharz, F. Pernkopf, *Neurocomputing*, vol. 80, pp. 38-46, 2012.

Y. Qian, S. Jia, J. Zhou, A. Robles-Kelly, *IEEE Trans. Geo. Rem. Sens.*, Vol. 49, No. 11, pp. 4282-4297, 2011.

T. H. Chan, C. Y. Chi, Y. M. Huang, W. K. Ma, *IEEE Trans. Sig. Proc.*, vol. 57, No. 11, pp. 4418-4432, 2009.

A. Pascual-Montano, J. M. Carazo, K. Kochi, D. Lehman, R. D. Pascual-Marqui, *IEEE Trans. Pat. Anal. Mach. Intel.*, Vol. 28, No. 3, pp. 403-415, 2006.



Linearan bezmemorijski problem SRS i ARK

Alternativa FNM metodama su algoritmi koji pododređeni problem nenegativnog SRS rješavaju u dvije faze: (i) procjena matrice \mathbf{A} metodama grupiranja podataka matrice \mathbf{X} (engl. data clustering); (ii) rješavanje linearnog pododređenog sustava jednadžbi $\mathbf{X}=\mathbf{AS}$, pri čemu je \mathbf{A} poznato iz (i).

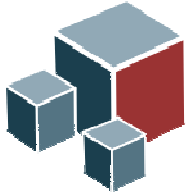
Od posebnog interesa za projekt su metode grupiranja podataka za procjenu matrice \mathbf{A} koje ne zahtijevaju postojanje realizacija u kojima je (dominantno) prisutna samo jedna komponenta:

Z. Yang, Y. Xiang, S. Xei, S. Ding, Y. Rong, *IEEE Trans. Neural Net. Learn. Syst.*, vol. 23, No. 10, pp. 1601-1610, 2012.

N. Gillis, S. A. Vavasis, preprint, 2012; <https://sites.google.com/site/nicolasgillis/>.

A. Ambikapathi, T. H. Chan, W. K. Ma, C. Y. Chi, *IEEE Trans. Geo. Rem. Sen.*, vol. 49, No. 11, pp. 4194-4209, 2011.

G. H. Ritter, G. Urcid, *Information Sciences*, vol. 181, pp. 1787-1803, 2011.



Linearan bezmemorijski problem SRS i ARK

Od posebnog interesa za projekt su algoritmi za rješavanje pododređenog sustava linearnih jednadžbi koji zahtijevaju manji broj mjerenja odnosno dozvoljavaju manju razinu rijetkosti rješenja (veći stupanj preklapanja):

M. Hayder, K. Mahata, *IEEE Trans. Sig. Proc.*, vol. 58, No. 4, pp. 2194-2205, 2010.

C. F. Caifa, A. Cichocki, *Neural Computation*, vol. 21, pp. 3487-3518, 2009.

Z. Hu, X. Chang, F. Xu, X. Zhang, *IEEE Trans. Neural Net. Learning Syst.*, vol. 23, No. 7, pp. 1013-1027, 2012.

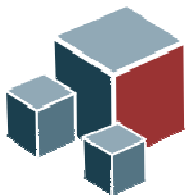
R. Chartrand, W. Yin, *IEEE ICASSP*, pp. 3869-3872, 2008.

D. Wipf, S. Nagarajan, *IEEE Journal Sel. Top. Sig. Proc.*, Vol. 4, No. 2, 317-329, 2010.

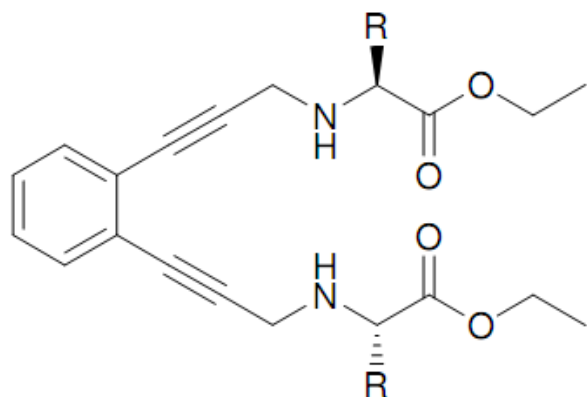
J. Vila, P. Schniter, "Expectation-maximization Gaussian-mixture approximate message passing", submitted, 2011.

Z. Zhang, B. D. Rao, *IEEE J. Selected Topics Sig. Proc.*, 5 (2011), 5: 912-926.

G. Marjanović, V. Solo, *IEEE Trans. Sig. Proc.*, vol. 60, No. 11, pp. 5714-5724, 2012.



Primjena ARK u izdvajanju pet analita iz dvije smjese



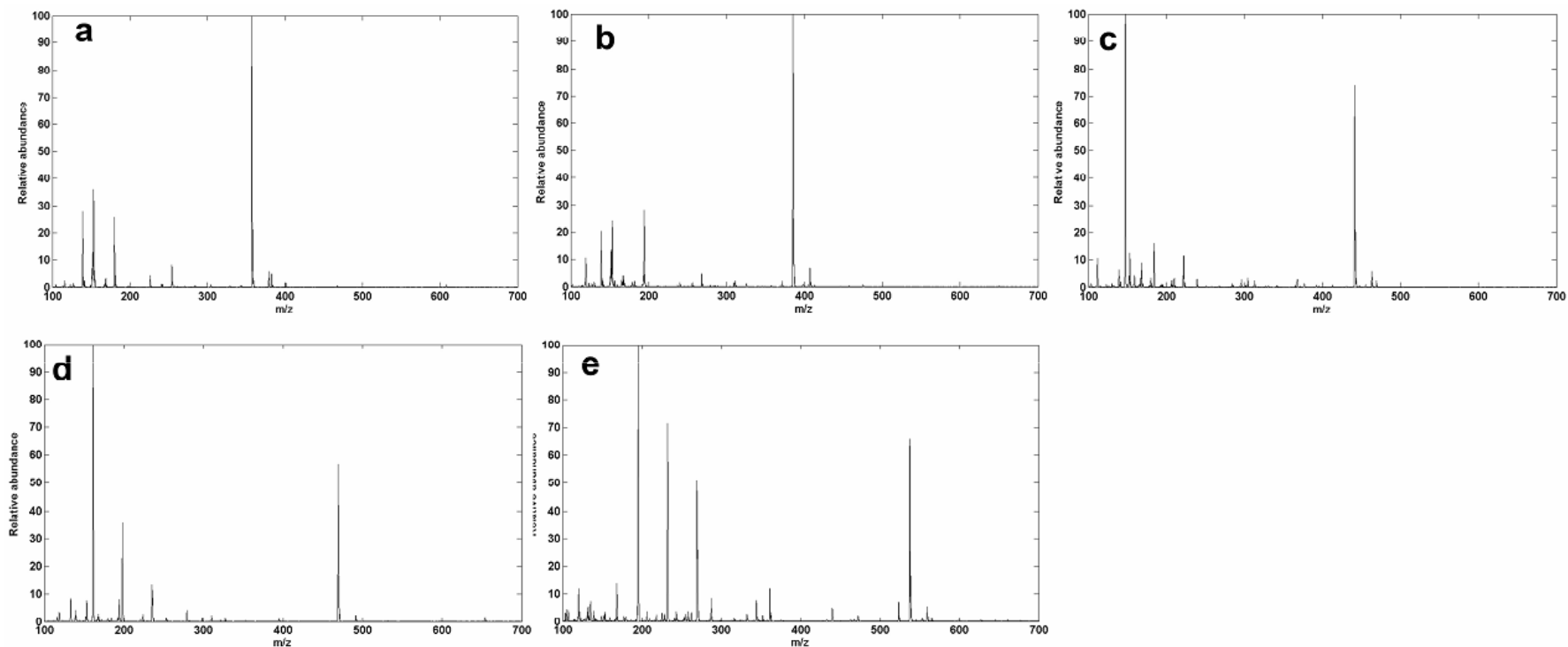
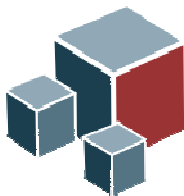
- 5 R=H
- 6 R=CH₃
- 7 R=CH(CH₃)₂
- 8 R=CH₂CH(CH₃)₃
- 9 R=CH₂C₆H₅

Kemijska struktura pet analita.

Figure S-1.

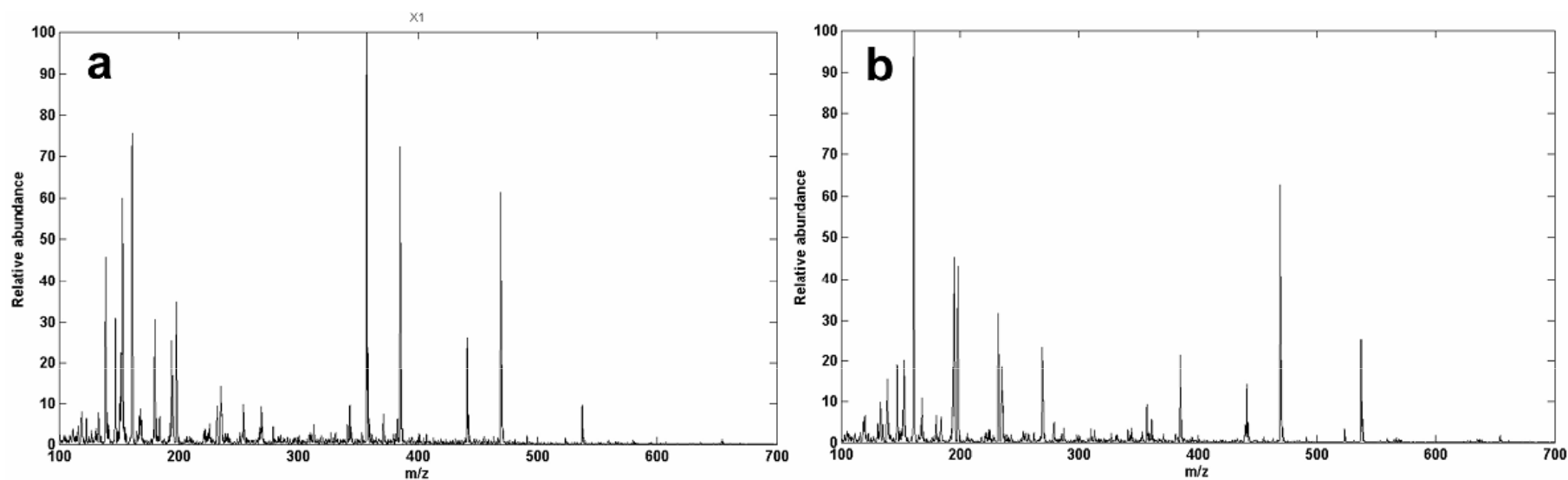
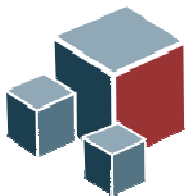
I. Kopriva, I. Jerić. Blind separation of analytes in nuclear magnetic resonance spectroscopy and mass spectrometry: sparseness-based robust multicomponent analysis, *Analytical Chemistry*, vol. 82, pp. 1911-1920, 2010.

I. Kopriva, I. Jerić, Method of and system for blind extraction of more pure components than mixtures in 1D and 2D NMR spectroscopy and mass spectrometry combining sparse component analysis and single component points, *US Patent 8,165,373*, 24. 4. 2012.

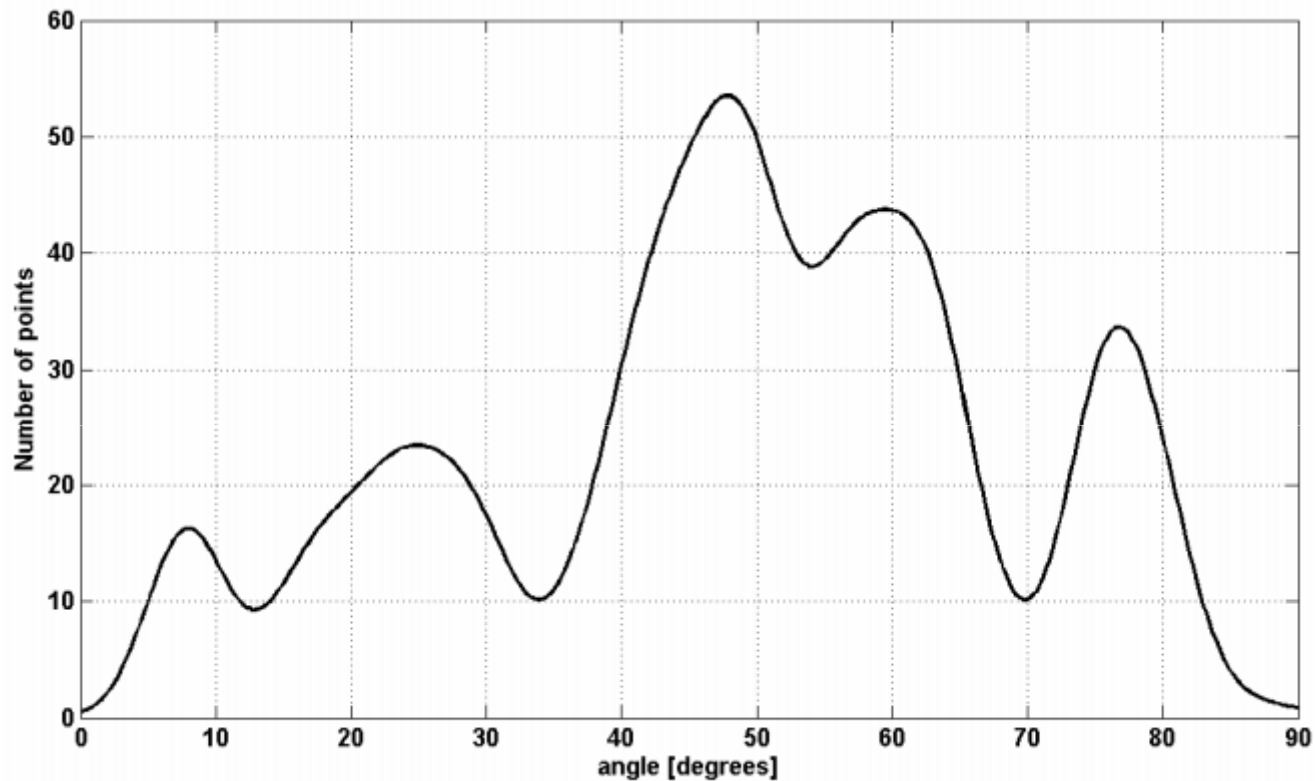
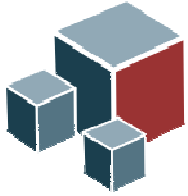


Spektri mase pet analita.

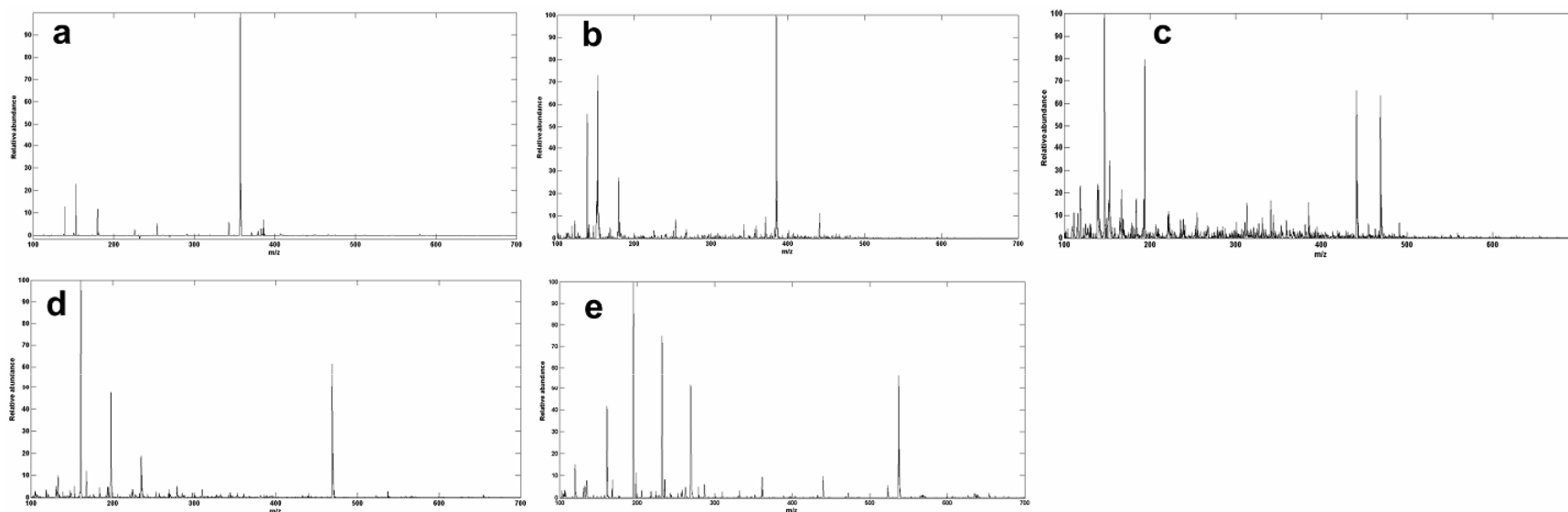
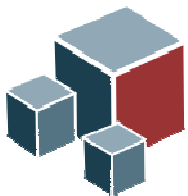
Stupanj preklopanja među analitima je vrlo mali (maksimalna korelacija ≈ 0.1).



Spektri mase dvije smjese.



Funkcija grupiranja podataka u domeni kuta vektora koncentracije. Pet vrhova ukazuje na prisutnost pet komponentata u spektrima mješavina.



Maseni spektri izdvojenih analita.

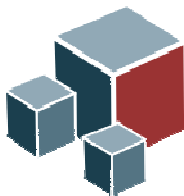


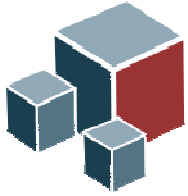
Table S-1. Normalized correlation coefficients for (a) pure analytes **5-9**; (b) analytes **5-9** estimated on 290 SAPs detected by using analytical representation (3) and *clusterdata* algorithm.*

entry		An_5	An_6	An_7	An_8	An_9
a	An_5	1	0.1268	0.0456	0.0266	0.0075
	An_6	0.1268	1	0.0321	0.0332	0.0379
	An_7	0.0456	0.0321	1	0.0134	0.0030
	An_8	0.0265	0.0332	0.0134	1	0.0029
	An_9	0.0075	0.0379	0.0030	0.0029	1
b	\hat{An}_5	0.9038	0.0305	0.0044	0.0002	0.0120
	\hat{An}_6	0.3162	0.8294	0.1198	0.0325	0.0043
	\hat{An}_7	0.0959	0.2334	0.7275	0.2009	0.0038
	\hat{An}_8	0.0043	0.0038	0.0124	0.9736	0.0293
	\hat{An}_9	0.0121	0.0161	0.0073	0.2097	0.9437

* An_5 - An_9 pure analytes **5-9**; \hat{An}_5 - \hat{An}_9 estimated analytes **5-9**.

Normalizirane korelacije poznatih i izdvojenih analita.

Što kada su analiti značajno korelirani i/ili smjese (reakcije) nisu linearne?



Nelinearan bezmemorijski problem SRS i NARK

Problem:

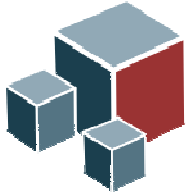
$$\mathbf{X} = F(\mathbf{S}) \quad \mathbf{X} \in \mathbb{R}_{0+}^{N \times T}, \mathbf{S} \in \mathbb{R}_{0+}^{M \times T} \quad (3)$$

gdje je $F: \mathbb{R}^M \rightarrow \mathbb{R}^N$ viševarijabilno preslikavanje oblika $\{\mathbf{x}(t) = F(\mathbf{s}(t))\}_{t=1}^T$ tako da

je $F(\mathbf{s}(t)) = [f_1(\mathbf{s}(t)) \dots f_N(\mathbf{s}(t))]^T$, odnosno $\{f_n: \mathbb{R}^M \rightarrow \mathbb{R}\}_{n=1}^N$.

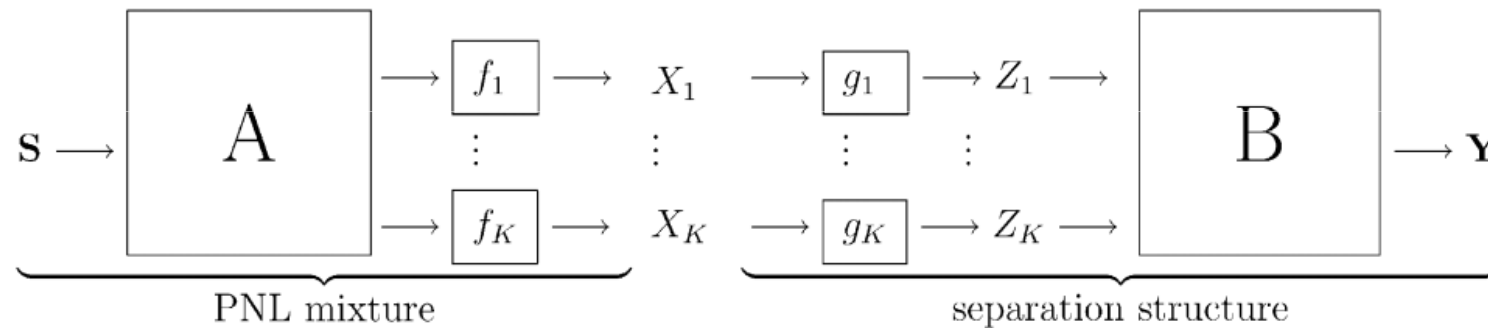
Cilj: procijeniti \mathbf{S} koristeći samo podatke \mathbf{X} , pri čemu je broj izvornih signala M nepoznat i može biti veći od broja smjesa N . Nelinearno preslikavanje F također nije poznato.

Nelinearan problem SRS (3) je puno teži od linearnog problema (1). Ograničenja na \mathbf{S} oblika statističke nezavisnosti, nenegativnosti i/ili rijetkosti nisu dovoljna da se procjeni \mathbf{S} .



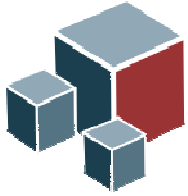
Nelinearan bezmemorijski problem SRS i NARK

Velik dio metoda NSRS podrazumijeva posebne vrste nelinearnosti F , što onda ograničava njihovu praktičnu primjenjivost. Tipičan slučaj je post-nelinearni (PNL) model:



Nelinearne funkcije $\{f_n : \mathbb{R} \rightarrow \mathbb{R}\}_{n=1}^N$ u PNL modelu djeluju nakon linearnog modela.

A. Taleb, and C. Jutten, "Source Separation in Post-Nonlinear Mixtures," *IEEE Trans. Sig. Proc.*, vol. 47, pp. 2807-2820, 1999.



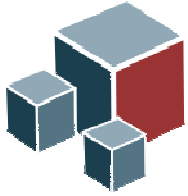
Nelinearan bezmemorijski problem SRS i NARK

Nelinearne funkcije $\{g_n \approx f_n^{-1} : \mathbb{R} \rightarrow \mathbb{R}\}_{n=1}^N$ lineariziraju model, pri čemu se statistička nezavisnost koristi kao ograničenje pri učenju $\{g_n \approx f_n^{-1} : \mathbb{R} \rightarrow \mathbb{R}\}_{n=1}^N$ i

B. Pri tome se višeslojni perceptron koristi za modeliranje inverznih nelinearnih funkcija.

Računska složenost metode temeljenih na PNL modelu je velika, te su demonstrirane na problemima NSRS niskog reda (2×2 do 4×4). U najvećem broju slučajeva zahtijevaju da broj komponenta bude manji ili jednak broju smjesa.

Metode NSRS od interesa za ovaj projekt temelje se na nelinearnoj transformaciji problema NSRS u RKHS i primjeni metoda za linearni problem SRS u induciranom RKHS:



Nelinearan bezmemorijski problem SRS i NARK

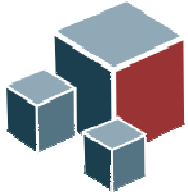
Predloženi koncept temelji se na nelinearnim preslikavanjima: $\{\mathbf{x}(t) \mapsto \phi(\mathbf{x}(t))\}_{t=1}^T$, gdje je $\phi: \mathbb{R}^N \rightarrow \mathbb{R}^D$ u visoko (moguće beskonačno) dimenzijski funkcijski prostor, tj. $D \gg N$, tako da je:

$$\phi(\mathbf{X}) \approx \mathbf{C}\mathbf{S}, \quad \phi(\mathbf{X}) \in \mathbb{R}_{0+}^{D \times T}, \quad \mathbf{S} \in \mathbb{R}_{0+}^{M \times T}, \quad \mathbf{C} \in \mathbb{R}_{0+}^{D \times M} \quad (4)$$

Veliki D je važan jer je vjerojatnost postojanja hiperravnina koje razdvajaju klase (izvorne signale), hiperravnine su redci matrice razdvajanja u modelu: $\mathbf{Y} = \mathbf{W}\phi(\mathbf{X}) \approx \mathbf{S}$, direktno proporcionalna D (Cover-ov teorem):

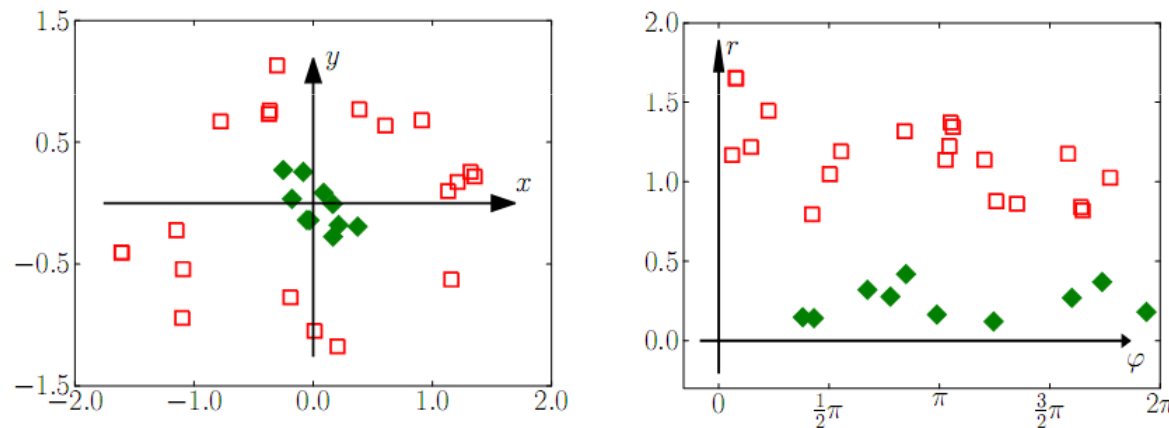
Cover, T. M., *IEEE Trans. Electronic Computers* EC-14 (3), 326-334, 1965.

Problem: (i) za jako veliki (moguće beskonačan) D faktorizacija implicirana sa (4) postaje računski neizvediva; (ii) koja preslikavanja $\phi(\mathbf{X})$ daju linearan model u transformiranom prostoru za široko područje nelinearnih funkcija $\{f_n^{-1}\}_{n=1}^N$?



Nelinearan bezmemorijski problem SRS i NARK

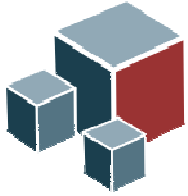
Primjer. Lijevo: empirijska raspodjela 2D uzoraka iz dvije klase: crvena i zelena. Ne postoji linearni klasifikator (pravac) koji može razdvojiti ove klase. Desno: nelinearna transformacija iz Kartezijevih u polarne koordinate daje linearno separabilan problem u preslikanom prostoru.



Komentar. Ova transformacija je optimalna za sferično separabilan polazni problem. U općem slučaju priroda površine koja razdvaja klase nije poznata!

Problem NSRS je načelno rješiv ako nelinearno preslikavanje inducira dovoljno bogat RKHS:

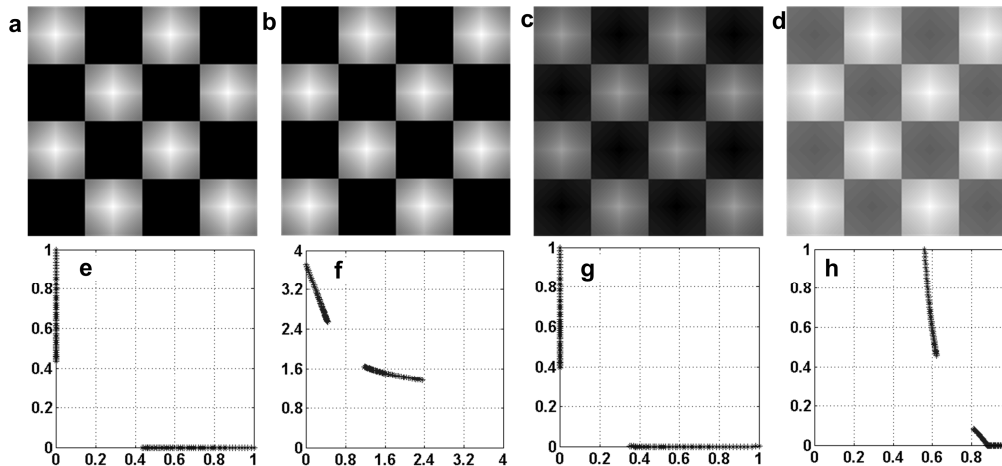
34/43



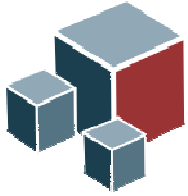
Nelinearan bezmemorijski problem SRS i NARK

Primjer. Problem NSRS: $x_1 = e^{s_1} - e^{s_2 - 1}$
 $x_2 = e^{-s_1} + e^{s_2}$

gdje su signali (slike) s_1 i s_2 međusobno rijetki (1-rijedak problem SRS).
Nelinearno preslikavanje u RKHS rađeno je sa Gaussovom jezgrom ($\sigma^2=16$)
sa $D=10$.



a) i b) izvorni signali s_1 i s_2 ; c) i d) nelinearne smjese x_1 i x_2 . Scatter dijagrami: e) izvornih signala; f) nelinearnih smjesa; g) signala izdvojenih NARK metodom; h) signala izdvojenih linearnom ARK metodom.



Nelinearan bezmemorijski problem SRS i NARK

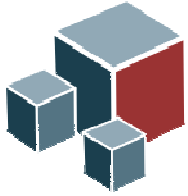
Primjer. Segmentacija RGB slike živca i tumora kože koristeći eksplicitna preslikavanja drugog i trećeg reda.

Eksplicitna preslikavanja $\Phi(\mathbf{x})$ dobiju se faktorizacijom izabrane jezgrene funkcije, a koristeći svojstvo reproduktivnosti: $k(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle_{H_k}$.

Za $\mathbf{x} \in \mathbb{R}^3$ faktorizacija polinomske jezgre $k(\mathbf{x}_i, \mathbf{x}_j) = (1 + \langle \mathbf{x}_i, \mathbf{x}_j \rangle)^d$ sa $d=3$ daje eksplicitno preslikavanje trećeg reda:

$$\phi(\mathbf{x}) = \begin{bmatrix} 1 & \sqrt{3}x_1 & \sqrt{3}x_2 & \sqrt{3}x_3 & \sqrt{3}x_1^2 & \sqrt{3}x_2^2 & \sqrt{3}x_3^2 & x_1^3 & x_2^3 & x_3^3 & \sqrt{3}x_1^2x_2 & \sqrt{3}x_1^2x_3 & \dots \\ \dots & \sqrt{3}x_1x_2^2 & \sqrt{3}x_1x_3^2 & \sqrt{3}x_2^2x_3 & \sqrt{3}x_2x_3^2 & \sqrt{6}x_1x_2x_3 & \sqrt{6}x_1x_2 & \sqrt{6}x_1x_3 & \sqrt{6}x_2x_3 & \dots & \dots & \dots \end{bmatrix}^T$$

3D vektor (piksel u RGB slici) se preslikava u 20-D vektor. Preslikavanje drugog reda daje 10-D vektor u RKHS.

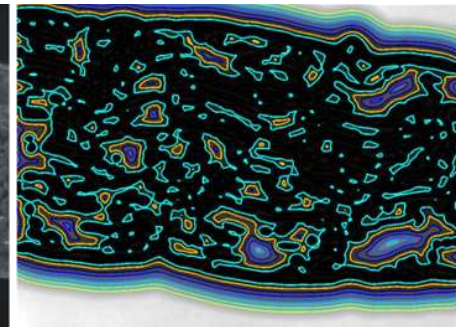
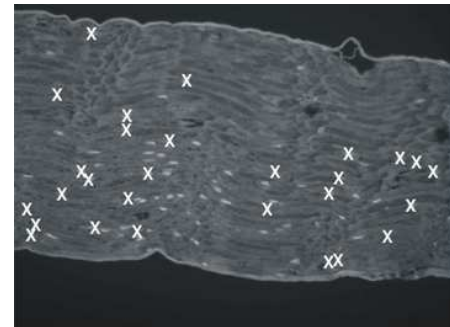


Nelinearan bezmemorijski problem SRS i NARK

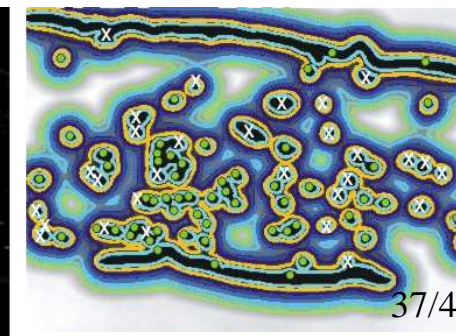
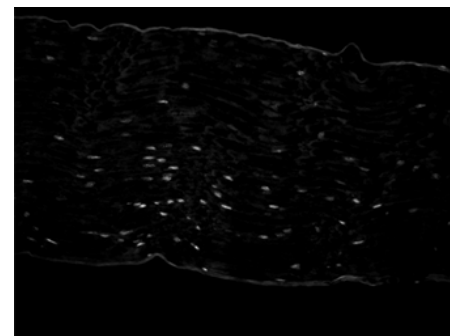
Rijetkosti ograničena FNM metoda u kombinaciji sa eksplicitnim preslikavanjem trećeg reda primijenjena na segmentaciju slike živca (*nervus ischiadicus*) bez kontrastnog reagensa:

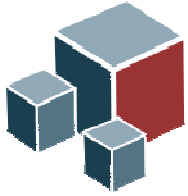
I. Kopriva, M. Hadžija, M. Popović-Hadžija, M. Korolija, A. Cichocki . *American Journal of Pathology*, vol. 179, No. 2, pp. 547-553, 2011.

Prvi red: siva slika (510 nm) bez kontrastnog reagensa i pripadajuća mapa isokontura.



Drugi red: odgovarajuća komponenta nakon NARK i pripadajuća mapa isokontura.



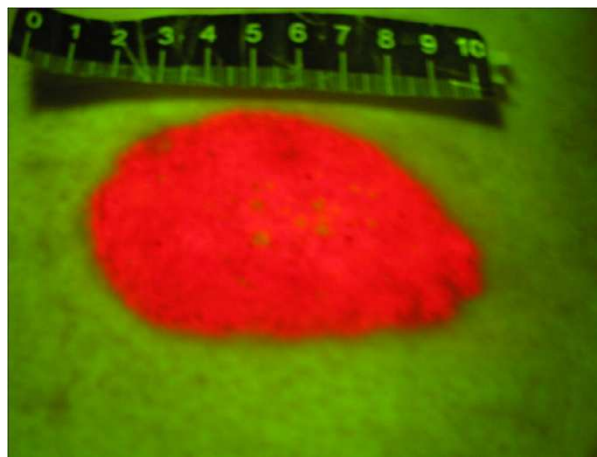


Nelinearan bezmemorijski problem SRS i NARK

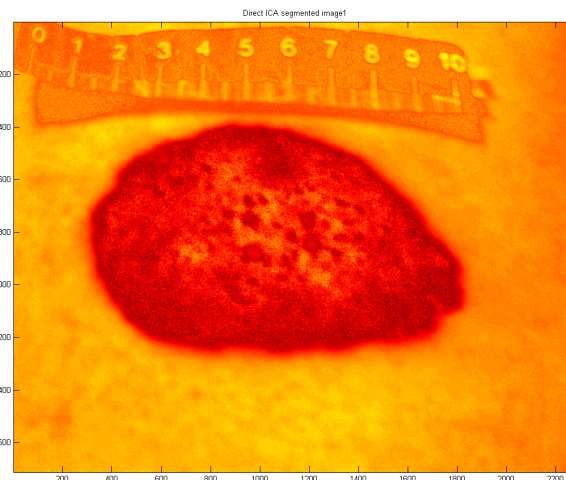
Segmentacija fluorescentne slike tumora kože koristeći analizu zavisnih komponentata u kombinaciji sa eksplicitnim preslikavanjem drugog reda demonstrirana je:

I. Kopriva and A. Peršin. *Medical Image Analysis* 13, 507-518, 2009.

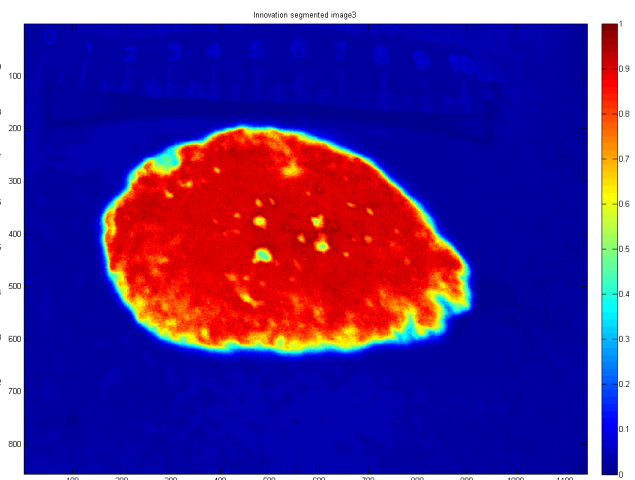
I. Kopriva, Method for real time tumour visualisation and demarcation by means of photodynamic diagnosis, *US Patent* 8,224,427, 17. 7. 2012.



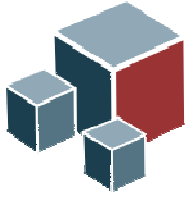
RGB fluorescentna slika visokog intenziteta



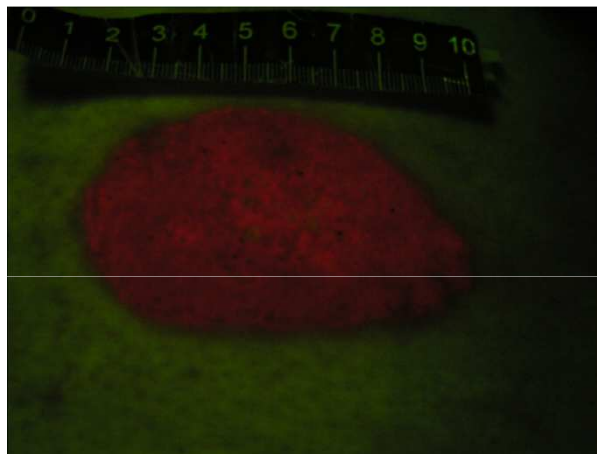
Linearni ICA algoritam



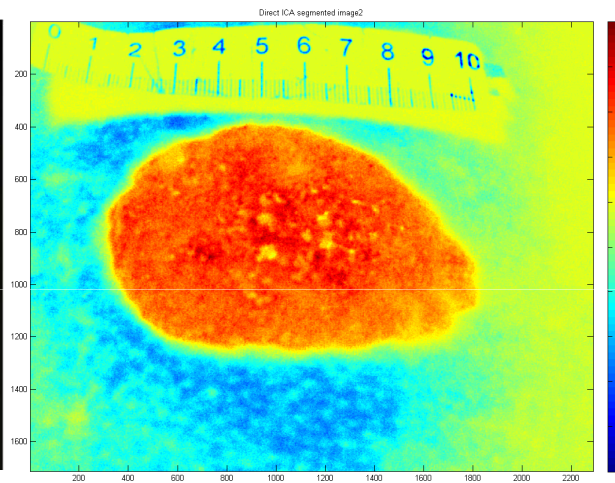
Patentirani NDCA algoritam



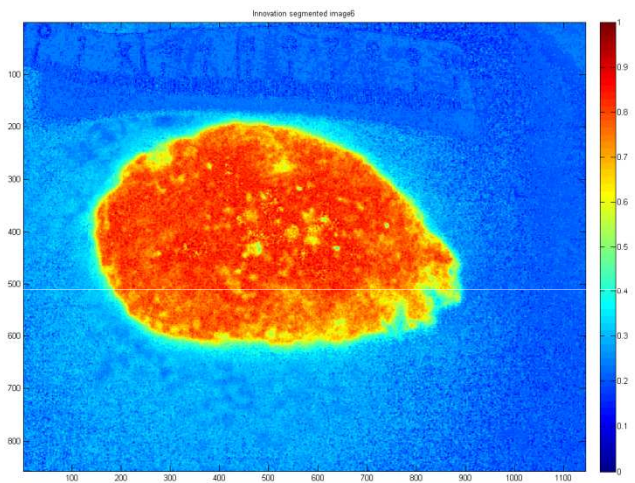
Nelinearan bezmemorijski problem SRS i NARK



RGB fluorescentna slika niskog intenziteta

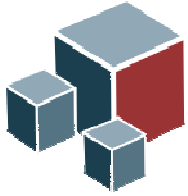


Linearni ICA algoritam



Patentirani NDCA algoritam

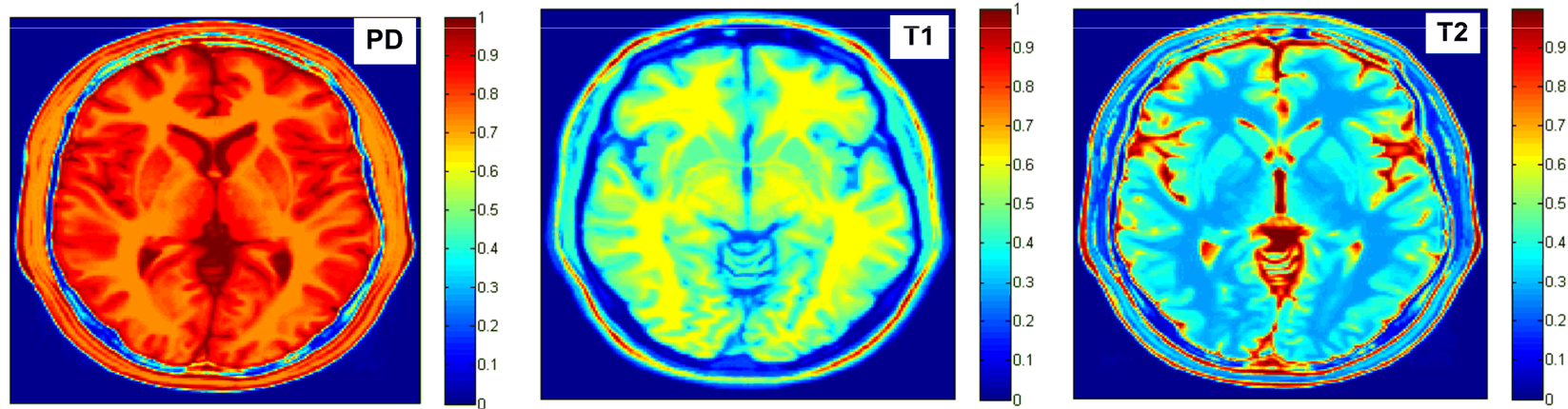
Poboljšanje kontrasta uporabom nelinearnog slijepog razdvajanja signala omogućava točniju demarkaciju tumora pri niskom intenzitetu fluorescencije.



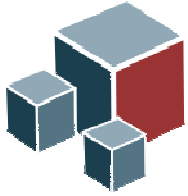
Nelinearan bezmemorijski problem SRS i NARK

Segmentacija MR slike mozga faktorizaciju nenegativnih tenzora (FNT) kombinaciji sa eksplicitnim preslikavanjem drugog reda demonstrirana je:

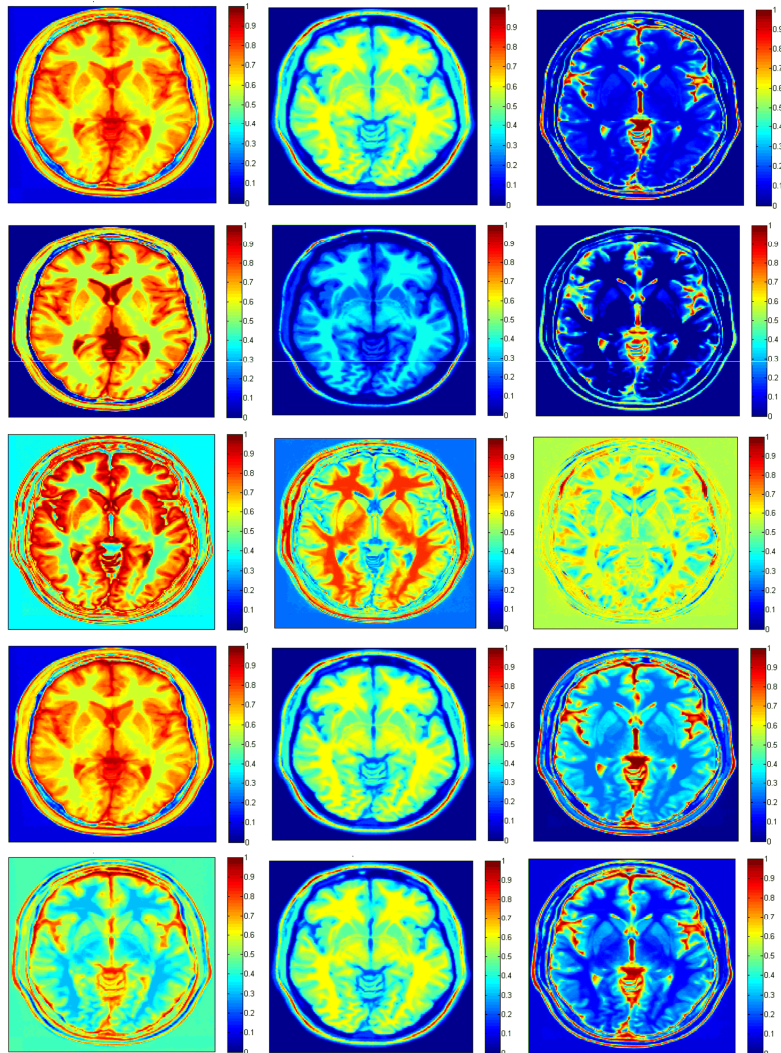
I. Kopriva, A. Cichocki. *Lecture Notes Computer Science* 6365, pp. 490-497, 2010.



Realistično simulirana slika mozga. S lijeva na desno: PD (siva masa-SM), T1 (bijela masa-BM), T2 (cerebrospinalna tekućina - CST).
www.bic.mni.mcgill.ca/brainweb/.

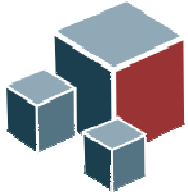


Nelinearan bezmemorijski problem SRS i NARK



SM, BM i CST supstance segmentirane od vrha prema dnu: nelinearni α -FNT ($\alpha=2$); nelinearni β -FNT ($\beta=1$); nelinearni EFICA; linearni α -FNT; linearni β -FNT.

Izbor nelinearnog preslikavanja je zavisian od problema. Kroz primjene na eksperimentalne podatke u spektrometriji mase i patologiji biti će provedena parametrizacija.

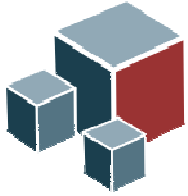


Nelinearan bezmemorijski problem SRS i NARK

Primjena NARK na segmentaciju slike histopatološkog uzorka (ZMM, IRB i KBC Dubrava).

Planira se izrada stakalaca uzoraka tkiva jetre NOD miševa sa svrhom vizualizacije poremećaja uzrokovanih narušenim metabolizmom lipida i ugljikohidrata u dijabetesu tipa I i II. U svakoj od tri skupine izradit će se 40 stakalaca (20 sa kontrastnim reagensom i 20 bez reagensa) – ukupno 120 stakalaca.

Planira se izrada 100 stakalaca (50 sa kontrastnim reagensom i 50 bez reagensa) humanog uzorka tkiva jetre u suradnji sa KBC Dubrava sa svrhom vizualizacije promjene na jetri uzrokovane metabolizmom lipida (*degeneratio lipomatosis*).



Nelinearan bezmemorijski problem SRS i NARK

Primjena NARK na dekompoziciju spektara mase nelinearnih kemijskih reakcija (ZOKB, IRB).

Planiraju se provesti reakcije esterifikacije aminokiselina i peptida, hidroliza estera i reakcije sinteze kratkih peptida.

U prvoj fazi planiraju se jednostavnije reakcije s manjim brojem komponenata u smjesi (reaktanti i produkti).

U drugoj fazi planiraju se reakcije s većim brojem komponenata, čime će se povećati kompleksnost sustava.