

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 1356

**Reprezentacija spline funkcija korištenjem
sažimajućeg očitavanja**

Željka Zjalić

Zagreb, lipanj 2016.

Zahvala:

Zahvaljujem svom mentoru, prof. dr. sc. Davoru Petrinoviću na strpljenju, pomoći i vodstvu prilikom pisanja ovog diplomskog rada.

Hvala svim kolegama i prijateljima na savjetima i olakšavanju studentskih dana.

Posebno hvala mojoj obitelji i M. koji su uvijek bili uz mene tijekom studija te na pruženom razumijevanju i podršci bez koje bi sve ovo bilo puno teže ostvariti.

Zagreb, 11. ožujka 2016.

DIPLOMSKI ZADATAK br. 1356

Pristupnik: **Željka Zjalić (0036462824)**
Studij: Elektrotehnika i informacijska tehnologija
Profil: Elektroničko i računalno inženjerstvo

Zadatak: **Reprezentacija spline funkcija korištenjem sažimajućeg očitavanja**

Opis zadatka:

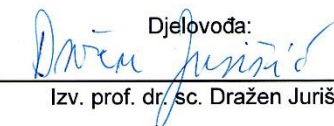
U okviru diplomskog rada potrebno je istražiti mogućnost rekonstrukcije uniformne i neuniformne spline funkcije koristeći tehniku sažimajućeg očitavanja (engl. compressed sensing) uz pronalaženje odgovarajuće osjetilne matrice i matrice transformacije u rijetku domenu s potrebnim uvjetima koji omogućuju potpunu rekonstrukciju. Odrediti potreban broj uzoraka dovoljnih za rekonstrukciju uz različite stupnjeve spline funkcije i različit broj njenih segmenata pri korištenju različitih algoritama za rekonstrukciju. Analizirati i primjenu rekonstrukcije za zašumljenu spline funkciju uz pronalaženje potrebnih uvjeta. Odstupanja rekonstruirane i originalne funkcije vrednovati pogodnim mjerama sličnosti. Sve analize, simulacije i vrednovanje ostvarenih rezultata potrebno je provesti korištenjem Matlaba.

Zadatak uručen pristupniku: 18. ožujka 2016.

Rok za predaju rada: 1. srpnja 2016.

Mentor:


Prof. dr. sc. Davor Petrinović

Djelovođa:


Izv. prof. dr. sc. Dražen Jurišić

Predsjednik odbora za
diplomski rad profila:



Prof. dr. sc. Mladen Vučić

Sadržaj

1.	Uvod	1
2.	Spline funkcije	2
2.1	Povijest i upotreba spline funkcija	2
2.2	Usporedba spline interpolacije i drugih interpolacijskih tehnika	5
2.2.1	Općenito o interpolaciji	5
2.2.2	Polinomijalna interpolacija.....	6
2.2.3	Spline interpolacija	11
2.3	Izvod i matrični zapis kubne spline funkcije	12
3.	Sažimajuće očitavanje	17
3.1	Općenito o tehnici sažimajućeg očitavanja	17
3.2	Problem uzorkovanja	18
3.3	Inkoherencija i uzorkovanje rijetkih signala	19
3.3.1	Rijetkost.....	19
3.3.2	Inkoherentno uzorkovanje	21
3.3	Poduzorkovanje i rekonstrukcija rijetkog signala	22
3.4	Sažeto očitavanje u šumovitom okruženju	25
3.4.1	Rekonstrukcija signala iz malog skupa podataka.....	27
3.4.2	Rekonstrukcija signala iz šumovitih podataka.....	28
3.5	Sažeto uzorkovanje.....	29
3.6	Primjene sažimajućeg očitavanja	30
4.	Primjena sažimajućeg očitavanja.....	34
4.1	Primjeri rijetkih baza	34
4.2	Primjeri osjetilnih matrica	35
4.1.1	Slučajne osjetilne matrica	35
4.1.2	Determinističke osjetilne matrice	37
4.1.2.1	Chirp osjetilna matrica	37
4.1.2.2	Osjetilna matrica sa statističkim restriktivnim izometrijskim svojstvom	38
4.3	Rekonstrukcijski algoritmi za ℓ_1 – minimizaciju	39
4.4	Mjera točnosti rekonstrukcije	40
5.	Pronalazak spline rijetke baze	42
5.1	Algoritam za generiranje spline rijetke baze.....	42
5.2	Tridijagonalni matrični zapis i invertiranje.....	43
5.3	Evaluacija polinoma.....	45

6. Provedba sažimajućeg očitavanja na spline funkcijama.....	48
7. Zaključak	51
Literatura.....	52

1. Uvod

Spline funkcije su polinomijalne funkcije po intervalima koje se koriste za opis i manipulaciju krivuljama te se često upotrebljavaju u numeričkoj matematici. Jedna od njihovih bitnijih primjena je interpolacija diskretnog skupa točaka, gdje se iz manjeg broja uzoraka može dobiti kontinuirana krivulja. Ideja interpolacije spline funkcijama je da se segment interpolacije razbije na niz manjih podintervala na kojima se zadana krivulja interpolira polinomima relativno niskog stupnja. Upravo zbog svog svojstva da iz ograničenog broja uzoraka, tzv. čvorova (eng. knots) spline funkcije generiraju kontinuirane krivulje, prikladne su za korištenje tehnike sažimajućeg očitavanja (eng. compressive sensing). Sažimajuće očitavanje podrazumijeva novu paradigmu uzorkovanja koja se protivi uobičajenom pristupu u akviziciji podataka. Teorija sažimajućeg očitavanja tvrdi da se određeni signali i slike mogu rekonstruirati iz mnogo manje mjerenja ili uzoraka od onog što tradicionalne metode koriste. Tehnika sažimajućeg očitavanja je zapravo jako efikasan i jednostavan protokol akvizicije koji na, o signalu neovisan način, uzorkuje signal niskom frekvencijom i kasnije koristi moć računanja za rekonstrukciju iz nečeg što se čini kao nepotpun skup podataka. Da bi to bilo moguće, jedan od bitnih podrazumijevajućih principa je rijetkost (eng. sparsity) signala, u smislu da signal ima sažetu reprezentaciju kad je izražen u prikladnoj bazi, što je svojstvo koje upravo spline funkcije sadržavaju. U okviru ovog rada proučena su svojstva i različiti zapisi spline funkcija te mogućnosti njihovog korištenja u tehnici sažimajućeg očitavanja u različitim uvjetima. Analize i simulacije su provedene koristeći program Matlab te je na kraju donesen zaključak o mogućim budućim primjenama.

2. Spline funkcije

2.1 Povijest i upotreba spline funkcija

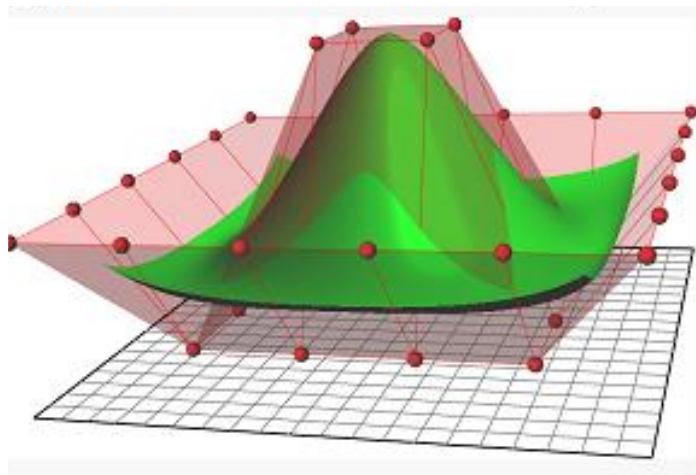
Iako mnogi ljudi koji nisu tehničke struke možda nisu upoznati s riječi spline, svijet je ispunjen krivuljama koje su stvorene korištenjem splineova. Splineovi su način modeliranja krivulja koristeći skup točaka koje mogu biti grupirane koristeći matematičke funkcije. Međutim, sam koncept spline funkcija prethodio je računalima te čak i matematičkim izračunima. Ime spline funkcija dolazi od engleske riječi *spline* istog korijena kao engleska riječ *splinter* koja znači dijelić ili komadić. Spline je označavao naprave koje su se upotrebljavale u tehničkom crtanju za konstrukciju glatkih krivulja, uglavnom u strojarstvu i brodogradnji. Preciznije, riječ je o dužoj elastičnoj napravi koja se može saviti tako da prođe kroz određen broj fiksiranih čvorova. Originalno su splineovi bili napravljeni od drva jer je drveni spline bio dovoljno fleksibilan da se lagano savija kad se čvorovi pomaknu ili se njihova težina promijeni. To je omogućavalo široku manipulaciju krivuljama i rezultiralo najgladom mogućom krivuljom između dvije točke.



Slika 1 Primjena splineova u brodogradnji [4] [5]

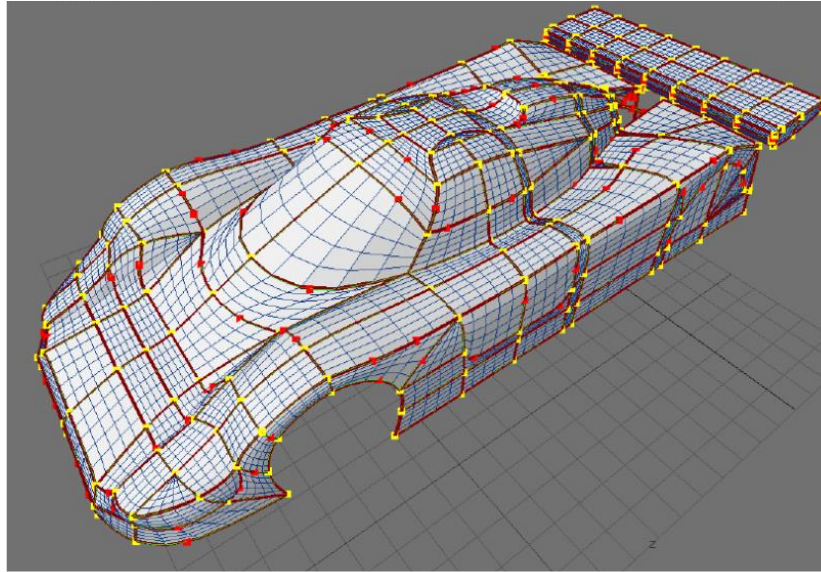
Brodovi su stoljećima građeni koristeći ovu metodu, međutim tek nakon Drugog svjetskog rata je matematičar I.O. Schoenberg razvio konciznu matematičku formulu poznatu kao spline funkcija ili krivulja. Spline funkcija je u svojoj biti skup dviju ili više polinomijalnih krivulja koje su glatko spojene zajedno.

Korištenje matematičke reprezentacije površine zrakoplova umjesto fizikalnog modela uštedilo je tisuće mjerenja u procesu dizajniranja i konstrukcije. U zračnoj industriji, osim njihove upotrebe u modeliranju, spline funkcije su neprocjenjive za crtanje trajektorija leta.



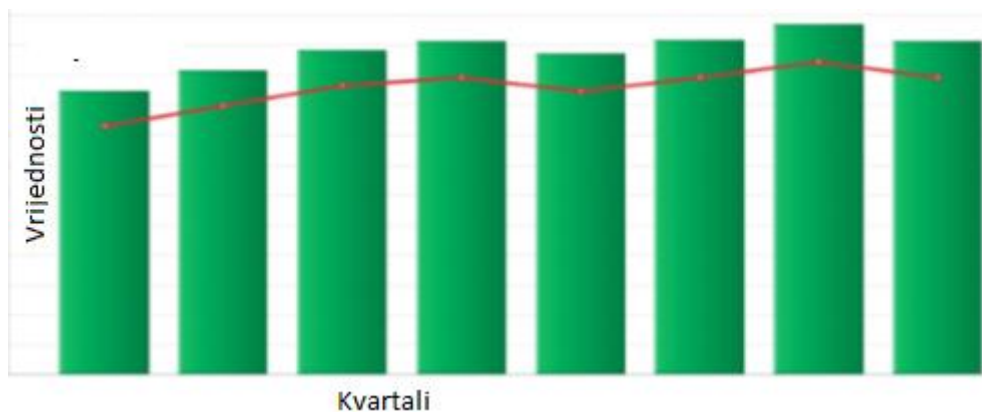
Slika 2 Primjena splineova u zrakoplovstvu [2]

U 1950-ima su se splineovi počeli koristiti u automobilskom dizajnu te kasnije u CAD-u (eng. Computer-aided Design) i drugim računalnim sustavima za dizajniranje. U digitalnim sustavima, čvorovi korišteni u brodogradnji su transformirani u kontrolne točke koje se mogu trenutačno mijenjati i micati da bi stvorile bilo koju vrstu krivulje. U automobilske industriji, upotreba spline funkcija rezultirala je automobilima s glatkim organskim linijama koje su prenijele smisao aerodinamičke učinkovitosti.



Slika 3 Primjena splineova u automobilskoj industriji [4]

U 1960-ima i 1970-ima upotreba CAD-a je bila ograničena na velike korporacije koje su si mogle priuštiti računsku moć potrebnu za izračun spline krivulja. Međutim, kako je računska moć ubrzo postala sve pristupačnija i programi sve dostupniji, današnji rezultati su uzeti zdravo za gotovo. Spline funkcije se također često koriste u poslovnoj inteligenciji. Statistički programi kao R imaju raznolikost spline opcija koje se primjenjuju na podacima. Primjerice, zaglađujući spline paket može transformirati zašumljene podatke u graf sa značenjem. Danas su splineovi nezaobilazni zbog njihove jednostavnosti i točnosti koje im dopuštaju da modeliraju visoko kompleksne podatke da bi bili razumljivi korisnicima [2].



Slika 4 Primjena splineova u poslovnoj inteligenciji [6]

2.2 Usporedba spline interpolacije i drugih interpolacijskih tehnika

2.2.1 Općenito o interpolaciji

U realnim primjenama vrlo često su nam na raspolaganju podaci koji su dobiveni eksperimentalno te su nam poznate samo vrijednosti u nekim zadanim točkama. Želja nam je iz djelomičnog poznavanja funkcije f čiji su dio te vrijednosti, na temelju raspoloživih podataka dobiti što je više moguće informacija o samoj funkciji. Ideja interpolacije je naći funkciju Φ koja prolazi kroz zadane točke, nadajući se pritom da se funkcije f i Φ u proizvoljnoj točki u intervalu zadanih vrijednosti neće mnogo razlikovati. Za takvu funkciju Φ kažemo da interpolira, odnosno povezuje zadane podatke. Često se ta funkcija naziva interpoland ili interpolirajuća funkcija.

Preciznije, interpolacija je zahtjev da se vrijednosti funkcija f i Φ podudaraju na nekom konačnom skupu argumenata (ili kraće točaka). Te točke obično nazivamo čvorovima interpolacije. Ovom zahtjevu se može, ali i ne mora dodati zahtjev da se u čvorovima, osim funkcijskih vrijednosti, poklapaju i vrijednosti nekih derivacija.

Drugim riječima, u najjednostavnijem obliku interpolacije, kad tražimo samo podudaranje funkcijskih vrijednosti, od podataka o funkciji f koristi se samo informacija o njenoj vrijednosti na skupu od $(n + 1)$ točaka, tj. podaci oblika (x_k, f_k) , gdje je

$$f_k = f_{(x_k)}, \text{ za } k = 0, \dots, n \quad (1)$$

Parametri (a_0, \dots, a_n) (kojih mora biti točno onoliko koliko i podataka) određuju se iz uvjeta:

$$\Phi(x_k; a_0, a_1, \dots, a_n) = f_k, \quad k = 0, \dots, n \quad (2)$$

što je, općenito, nelinearni sustav jednadžbi.

Ako je aproksimacijska funkcija Φ linearna, onda za parametre a_k dobivamo sustav linearnih jednadžbi koji ima točno $(n + 1)$ jednadžbi i $(n + 1)$ nepoznanica. Matrica tog sustava je kvadratna, što bitno olakšava analizu egzistencije i jedinstvenosti rješenja za parametre interpolacije [3].

Ako je $[a, b]$ segment na kojem koristimo interpolaciju, u praksi su točke obično numerirane tako da vrijedi

$$a \leq x_0 < x_1 < \dots < x_n \leq b. \quad (3)$$

Važno pitanje povezano s problemom interpolacije jest odabir skupa funkcija među kojima tražimo interpolirajuću funkciju. Ovisno o prirodi zadanih podataka za interpolacijske funkcije mogu se koristiti polinomi, eksponencijalna funkcija, trigonometrijske funkcije ili racionalne funkcije. Naime, ako originalna funkcija f ima svojstvo periodičnosti, onda i interpolirajuća funkcija također treba biti periodična. Periodične funkcije gotovo se uvijek interpoliraju pomoću trigonometrijskih funkcija. Funkcije koje posjeduju singularitete interpoliraju se uglavnom pomoću racionalnih funkcija koje imaju sličnu strukturu singulariteta. Polinomijalna interpolacija se zbog svojih nedostataka uglavnom upotrebljava kao pomoćno sredstvo u konstrukciji nekih numeričkih postupaka, npr. numeričkoj integraciji. S druge strane, nedostatci polinomijalne interpolacije se mogu djelomično otkloniti ako interpoliramo po dijelovima promatranog segmenta na način da prvo podijelimo promatrani segment na podintervale, a zatim na svakom od njih radimo interpolaciju polinomom odgovarajućeg stupnja, pazeći pritom na glatkoću dobivene funkcije u zadanim vrijednostima. Tako dobivamo spline interpolaciju. Primijetimo da postupkom interpolacije dobivamo funkciju Φ koja aproksimira nepoznatu funkciju f . U nastavku ćemo pobliže promotriti polinomijalnu interpolaciju.

2.2.2 Polinomijalna interpolacija

U prethodnom potpoglavlju spomenuto je da polinomijalna interpolacija ima nedostatke zbog kojih uglavnom nije korisna za primjene. Ti nedostatci će biti detaljnije prikazani u nastavku.

Označimo sa \mathbb{P}_n skup svih realnih polinoma p čiji je stupnj jednak ili manji $n \in \mathbb{N}$. Svaki element $p \in \mathbb{P}_n$ može se prikazati u obliku:

$$p(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \quad (4)$$

gdje su $a_i \in \mathbb{R}, i = 0, \dots, n$ koeficijenti polinoma p .

Prvo pitanje je ima li interpolacijska zadaća rješenje i je li ono jedinstveno.

Za polinomnu interpolaciju vrijedi sljedeći teorem, čiji dokaz koristi činjenicu da linearni sustav s regularnom matricom ima jedinstveno rješenje. U iskazu teorema koristi se oznaka \mathbb{N}_0 za skup cijelih nenegativnih brojeva.

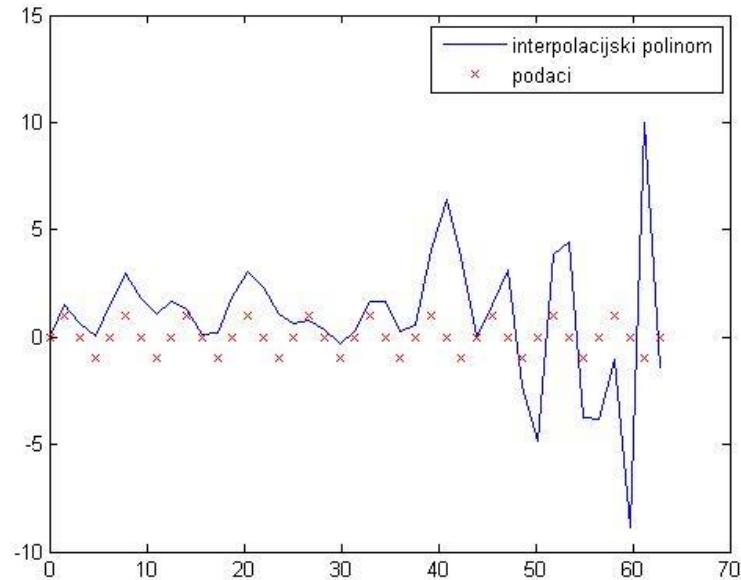
Teorem 1: Neka je $n \in \mathbb{N}_0$. Za zadane točke $(x_k, f_k), k = 0, \dots, n$, gdje je $x_i \neq x_j$, za $i \neq j$ postoji jedinstveni (interpolacijski) polinom stupnja najviše n $\Phi(x) := p_n(x) = a_0 + a_1x + \dots + a_nx^n$ za koji vrijedi $p_n(x_k) = f_k, k = 0, \dots, n$.

Ovaj teorem u potpunosti rješava prvo ključno pitanje egzistencije i jedinstvenosti rješenja problema polinomne interpolacije u njegovom najjednostavnijem obliku — kad su zadane funkcijske vrijednosti u međusobno različitim točkama [1].

Interpolacijska zadaća zapisana u matričnom zapisu izgleda:

$$\begin{bmatrix} 1 & \dots & x_0^n \\ \vdots & \ddots & \vdots \\ 1 & \dots & x_n^n \end{bmatrix} \begin{bmatrix} a_0 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} f_0 \\ \vdots \\ f_n \end{bmatrix} \quad (5)$$

Matrica sustava, tzv. Vandermondeova matrica ovisi o broju zadanih polaznih točaka. Porastom broja interpolacijskih čvorova stupanj interpolacijskog polinoma i red matrice se povećavaju. Rješavanje interpolacijske zadaće matričnim postupkom nije odgovarajuće jer su koeficijenti za veći broj čvorova neprecizno izračunati, kao što je prikazano na slici 5.



Slika 5 Interpolacija funkcije $\sin(x)$ na $[0, 20\pi]$

```
n = 40;
x = linspace(0,20*pi,n+1)';
y=sin(x);
V=zeros(n+1);
for i = 1:n+1
    V(:,i) = x.^(i-1);
end
a = V\y;
f = horner (a,x);
plot(x,f, 'b-', x,y, 'rx');
```

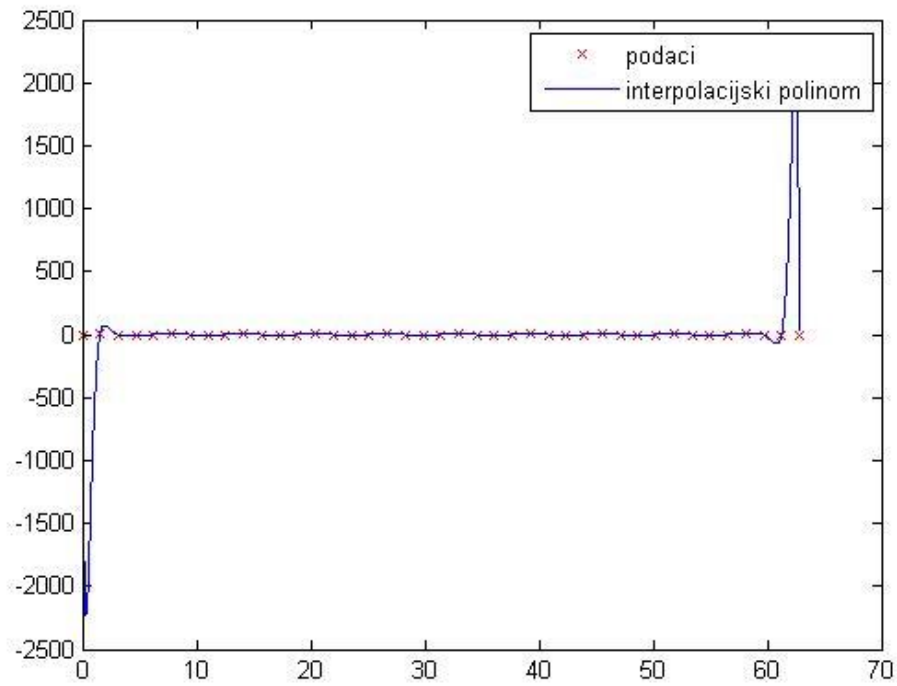
```
function y = horner( a,x )
k = length(a)-1;
n = length(x);
y(1:n,1) = a(k+1);
for j = k:-1:1
    y = a(j) + y.*x;
end

end
```

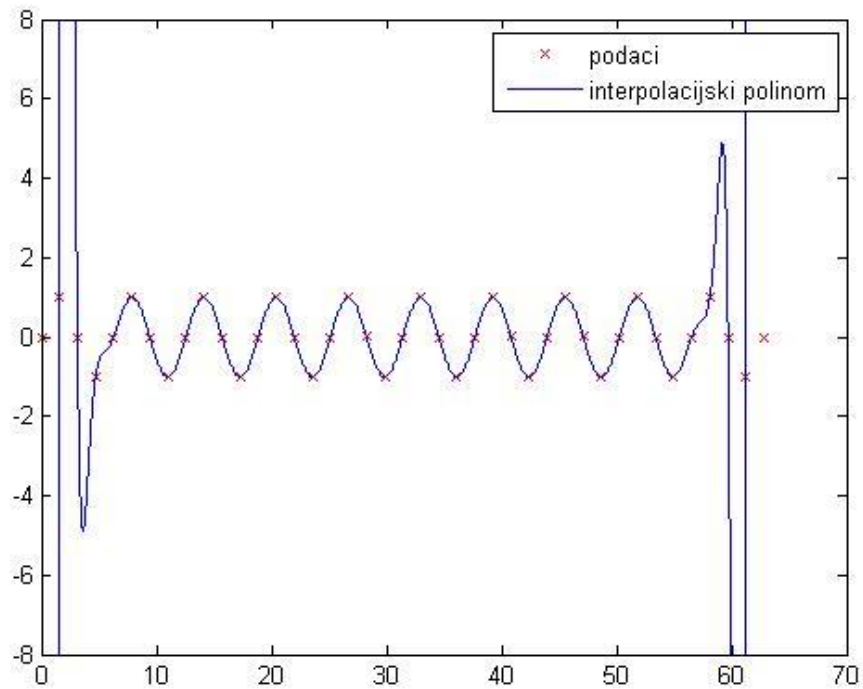
Kod 1 Izračun interpolacijske zadaće korištenjem matričnog zapisa

Da bismo izbjegli problem nepreciznosti koeficijenata, interpolacijsku zadaću možemo rješavati bez korištenja matričnog zapisa, pomoću Lagrangeova oblika polinoma:

$$p(x) = \sum_{j=0}^n \beta_j \varphi_j(x), \quad \varphi_j(x) = \prod_{i=0, i \neq j}^n \frac{x-x_i}{x_j-x_i}, j = 0, 1, \dots, n \quad (6)$$



Slika 6a Lagrangeov interpolacijski polinom za $\sin(x)$



Slika 6b Lagrangeov interpolacijski polinom za $\sin(x)$ povećane rezolucije

```

function v = polyinterp( x,y,u )
n = length(x);
v = zeros(size(u));
for k = 1:n
    w = ones(size(u));
    for j = [1:k-1 k+1:n]
        w = (u-x(j))./(x(k)-
x(j)).*w;
    end
    v = v+w*y(k);
end

end

```

```

n = 40;
x =
linspace(0,20*pi,n+1)
';
y=sin(x);
xu = [0:0.01:20*pi];
v =
polyinterp(x,y,xu);
plot(x,y,'rx',xu,v,'b
-');

```

Kod 2 Izračun interpolacijske zadaće korištenjem Lagrangeovog zapisa

Međutim, kao što je vidljivo na slici 6, iako interpolacijski polinom prolazi kroz zadane točke (slika 6a), u blizini rubova segmenta se pojavljuju oscilacije (slika 6b) te je u tim područjima aproksimacija izuzetno loša.

Na temelju ovog primjera vidimo da se prijelazom na Lagrangeovu bazu može riješiti problem preciznog računanja koeficijenata interpolacijskog polinoma, ali ostaje pitanje koliko dobro interpolacijski polinom visokog stupnja aproksimira traženu funkciju. Naime, polinomi višeg stupnja koji imaju međusobno različite nultočke su oscilatorne funkcije i teže u beskonačnost kada x raste. Loša aproksimacija interpolacijskim polinomom na rubnim segmentima upravo je posljedica oscilatornih svojstava polinoma.

U praksi se obično koriste interpolacijski polinomi niskih stupnjeva, najčešće do petog stupnja jer kod nekih funkcija za određeni izbor točaka interpolacije, povećavanje stupnja interpolacijskog polinoma može dovesti do povećanja grešaka.

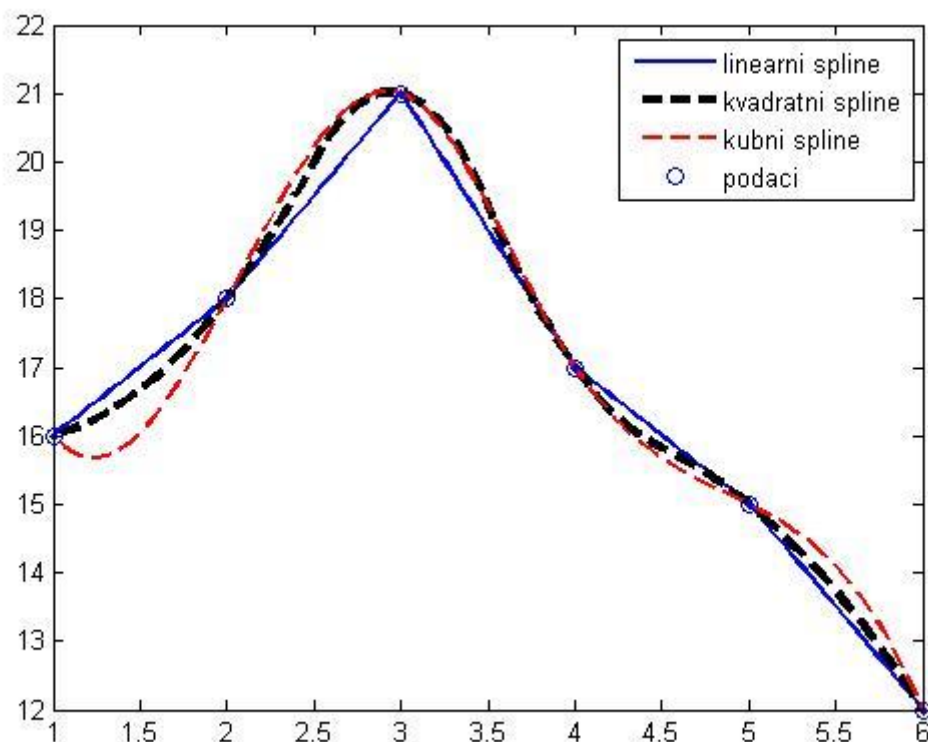
Nadalje, na rubu područja interpolacije polinomi višeg stupnja pokazuju jake oscilacije. Zbog toga se umjesto visokog stupnja interpolacijskog polinoma u praksi koristi po dijelovima polinomna interpolacija.

2.2.3 Spline interpolacija

Ideja spline interpolacije ili polinomijalne interpolacije po dijelovima počiva na tome da segment interpolacije razbijemo na niz manjih podintervala na kojima ćemo zadanu funkciju interpolirati polinomima relativno niskog stupnja. Na rubovima podintervala dobivena interpolirajuća funkcija neće biti proizvoljno glatka, no ipak se može postići zadovoljavajući stupanj diferencijabilnosti. U većini slučajeva važnije je izbjeći moguće oscilacije nego težiti za postizanjem velike glatkoće interpolirajuće funkcije. Po dijelovima glatke funkcije koje interpoliraju zadane točke nazivaju se splineovi. Slijedi matematička definicija splinea.

Definicija 1. Neka je na intervalu $[a, b]$ zadana mreža $\Delta = \{a \leq x_0 < x_1 < \dots < x_n \leq b\}$. Spline reda $l \in \mathbb{N}$ na zadanoj mreži Δ je neprekinuta funkcija $s(x)$ koja se na svakom podintervalu $[x_i, x_{i+1}]$ podudara s polinomom stupnja najviše l , odnosno vrijedi: $s|_{[x_i, x_{i+1}]} \in \mathcal{P}_l, i = 0, \dots, n-1$.

Kao što vidimo na slici 7, po dijelovima linearni interpoland je neprekidna funkcija čija prva derivacija nije neprekidna. Kvadratni spline je manje zanimljiv za primjene jer tek krivulje čija je druga derivacija neprekidna smatramo vizualno glatkima i kvadratni spline može imati skokove prve derivacije u zadanim točkama. Vidljivo je i da je kubni splajn glađi od kvadratnog te se zato najčešće upotrebljava.



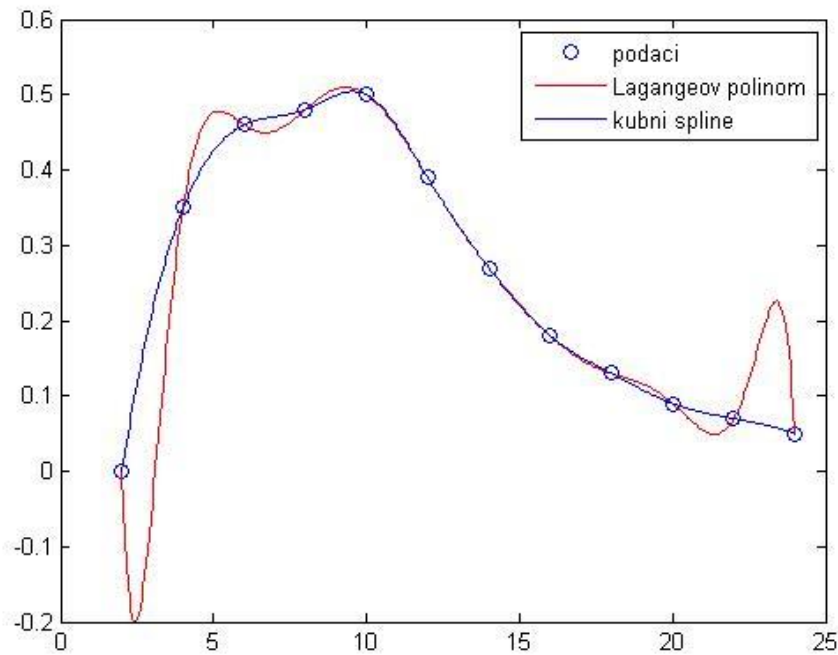
Slika 7 Primjeri spline interpolacija različitog reda

2.3 Izvod i matrični zapis kubne spline funkcije

Ovisno u udaljenosti x koordinata zadanih točaka, spline funkcija može biti uniformna ili neuniformna. Ako su x koordinate jednako razmaknute, govorimo u uniformnoj spline funkciji, u suprotnom je riječ o neuniformnoj spline funkciji.

Ponekad je aproksimacija linearnim splineom zadovoljavajuća, no ako želimo postići veću glatkoću interpolanda, potrebno je konstruirati spline višega reda.

Na slici 8 vidimo da se korištenjem po dijelovima polinomijalne interpolacije uklonio problem oscilacija na rubovima intervala.



Slika 8 Usporedba Lagrangeovog interpolacijskog polinoma i kubnog splinea

```
x = [2:2:24];
y = [0 0.35 0.46 0.48 0.5 0.39 0.27 0.18 0.13 0.09 0.07
0.05];
xu = [2:0.05:24];
p = polyinterp(x,y,xu);
s = interp1(x,y,xu,'spline');
plot(x,y,'o');hold on;
plot(xu,p,'r-');hold on;
plot(xu,s);
```

Kod 3 Usporedba polinomijalne i spline interpolacije

Kubni spline za podatke $\{(x_i, y_i), i = 0, \dots, n\}$ je neprekinuta funkcija koja je na svakom podintervalu $[x_i, x_{i+1}]$, za $i = 0, 1, \dots, n-1$, polinom najviše trećeg stupnja, u čvorovima interpolacije zadovoljava uvjete interpolacije te ima neprekidnu drugu derivaciju. U idućim zapisima kubni spline označavamo sa s .

Polinom trećeg stupnja ima četiri koeficijenta, tako da je ukupan broj parametara ili stupnjeva slobode $4n$.

Kubni spline mora udovoljavati sljedećim uvjetima:

a)

neprekidnost funkcije $s(x_i - 0) = s(x_i + 0)$ za $i = 1, 2, \dots, n - 1$ (7)

b)

uvjeti interpolacije $s(x_i) = y_i$ za $i = 0, 1, \dots, n$ (8)

c)

neprekidnost prve derivacije $s'(x_i - 0) = s'(x_i + 0)$ za $i = 1, 2, \dots, n - 1$ (9)

d)

neprekidnost druge derivacije $s''(x_i - 0) = s''(x_i + 0)$ za $i = 1, 2, \dots, n - 1$ (10)

e) $s'(x_0) = f'(x_0), s'(x_n) = f'(x_n)$ - potpuni (11)

različiti rubni uvjeti $s'(x_0) = s'(x_n), s''(x_0) = s''(x_n)$ - periodični (12)

$s''(x_0) = s''(x_n) = 0$ - prirodni (13)

Iz jednadžbi (7) do (10) vidimo da imamo ukupno $3(n - 1) + (n + 1) = 4n - 2$ uvjeta. Dakle, problem interpolacije ima $4n$ nepoznatih parametara koje trebamo odrediti koristeći $4n - 2$ poznatih uvjeta. Da bismo jedinstveno odredili sve nepoznate parametre, potrebno je postaviti dodatna dva uvjeta. Dodatni uvjeti se obično odnose na ponašanje spline funkcija na rubovima zadanog intervala. Ako je poznata derivacija funkcije f na rubovima, postavljaju se uvjeti oblika (11) i na taj način se dobije potpuni kubni spline.

Druga mogućnost je postavljanje prperiodičnih rubnih uvjeta na prvu i drugu derivaciju (12) i time se dobije periodični kubni spline. Radi jednostavnosti računa, uglavnom se koristi uvjet (13) i dobiveni spline se naziva prirodni kubni spline. Upravo prirodni kubni spline koristit će se u nastavku rada.

Kombinacijom danih uvjeta za prirodni kubni spline dobiju se sljedeći rezultati. Detaljni izvod moguće je pronaći u [1] i [10].

Koristeći notaciju iz jednadžbe (4) $\gamma = a_2$ slijedi [9]:

$$A\gamma = v$$

$$h_i = x_{i+1} - x_i, \quad \gamma_0 = \gamma_n = 0,$$

$$v_i = \frac{6}{h_i}(y_{i+1} - y_i) - \frac{6}{h_{i-1}}(y_i - y_{i-1})$$

$$A = \begin{bmatrix} 2(h_1 + h_2) & h_2 & 0 & \dots & 0 \\ h_2 & 2(h_2 + h_3) & h_3 & \ddots & \vdots \\ 0 & h_3 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & h_{n-2} \\ 0 & \dots & 0 & h_{n-2} & 2(h_{n-2} + h_{n-1}) \end{bmatrix}$$

$$\gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_{n-2} \\ \gamma_{n-1} \end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n-2} \\ v_{n-1} \end{bmatrix} \quad (14)$$

Preostali koeficijenti se dobiju sljedećim jednadžbama:

$$a_{0i} = y_i$$

$$a_{1i} = \frac{y_{i+1} - y_i}{h_i} - \gamma_i h_i - a_{3i} h_i^2$$

$$\begin{aligned}
 a_{2i} &= \gamma_i \\
 a_{3i} &= \frac{\gamma_{i+1} - \gamma_i}{3h_i}
 \end{aligned}
 \tag{15}$$

Definicija 2. Za matricu $A \in \mathbb{R}^{n \times n}$ kažemo da je strogo dijagonalno dominantna ako vrijedi

$$\sum_{j=1, j \neq k}^n |a_{kj}| < |a_{kk}|, \forall k = 1, 2, \dots, n$$

Propozicija 1. Neka je matrica $A \in \mathbb{R}^{n \times n}$ strogo dijagonalno dominantna matrica. Tada je A regularna matrica.

Matrica A iz jednadžbe (14) je tridijagonalna i strogo dijagonalno dominantna jer je u svakom njenom retku element na dijagonali veći od zbroja vrijednosti ostalih elemenata u promatranom retku, kao što je zapisano u definiciji 2. Zbog toga vrijedi da je matrica A regularna (propozicija 1) te sustav iz jednadžbe (14) ima jedinstveno rješenje. Isto vrijedi i za splineove s drugim rubnim uvjetima [1].

U Matlabu se spline funkcije mogu se dobiti pozivanjem raznih naredbi, neke od njih su interp1, spline i spapi.

3. Sažimajuće očitavanje

3.1 Općenito o tehnici sažimajućeg očitavanja

Konvencionalni pristupi uzorkovanju signala ili slika slijede Shannonov proslavljeni teorem: frekvencija uzorkovanja mora biti najmanje dvostruko veća od maksimalne frekvencije prisutne u signalu, tzv. Nyquistove frekvencije. Ovaj princip obuhvaća gotovo sve protokole akvizicije signala korištene u potrošačkoj audio i vizualnoj elektronici, uređajima za snimanje medicinskih slika, radio prijemnicima, itd. Za neke signale koji nisu prirodno pojasno ograničeni, primjerice slike, frekvencija uzorkovanja nije diktirana Shannonovim teoremom, nego željenom vremenskom ili prostornom rezolucijom. Međutim, uobičajeno je da se u takvim sustavima koristi antialiasing niskopropusni filter za pojasno ograničavanje signala prije uzorkovanja i tako Shannonov teorem ima implicitnu ulogu. U području pretvorbe podataka standardni analogno-digitalni pretvornik implementira uniformno kvantiziranu Shannonovu reprezentaciju: signal je uniformno uzorkovan na ili iznad Nyquistove frekvencije.

Sažimajuće očitavanje (eng. Compressed sensing), još poznat kao CS ili kompresivno uzorkovanje, je nova paradigma uzorkovanja koja se protivi uobičajenom pristupu u akviziciji podataka. CS teorija tvrdi da se određeni signali i slike mogu rekonstruirati iz mnogo manje mjerenja ili uzoraka od onog što tradicionalne metode koriste. Da bi to bilo moguće, CS se oslanja na dva principa: rijetkost (eng. sparsity) koja se odnosi na signal i nepovezanost (eng. incoherence) koja se odnosi na osjetilni način.

Rijetkost izražava ideju da „frekvencija informacije” kontinuiranog vremenskog signala može biti puno manja nego što je sugerirana frekvencijskim pojasom signala ili da vremenski diskretan signal ovisi o broju stupnjeva slobode, što je mnogo manje od njegove konačne dužine.

Preciznije, CS iskorištava činjenicu da su mnogi prirodni signali rijetki ili podložni kompresiji u smislu da imaju sažetu reprezentaciju kad su izraženi u prikladnoj bazi ψ .

Nepovezanost ili inkoherencija proširuje dualnost između vremena i frekvencije i izražava ideju da objekti koji imaju rijetku reprezentaciju u bazi ψ , moraju biti rašireni u domeni u kojoj su prikupljeni, baš kao što je Dirac u vremenskoj domeni raširen u frekvencijskoj domeni. Rečeno na drugi način, nepovezanost kaže da za razliku od promatranog signala, valni oblici korišteni za uzorkovanje imaju izrazito gustu reprezentaciju u bazi ψ .

Ključno zapažanje je da se mogu dizajnirati efikasni protokoli uzorkovanja koji hvataju korisni informacijski sadržaj ugrađen u rijetki signal i sažimaju ga u malu količinu podataka. Ti protokoli jednostavno zahtijevaju korelaciju signala s malim brojem fiksnih valnih oblika koji su inkoherentni s rijetkom bazom. Zanimljivo kod njih je da dopuštaju senzoru da jako efikasno uhvati informaciju u rijetkom signalu bez pokušaja razumijevanja tog signala. Nadalje, postoji način korištenja numeričke optimizacije kojim bi se rekonstruirao signal pune duljine iz male količine prikupljenih podataka. CS je zapravo jako efikasan i jednostavan protokol akvizicije koji na, o signalu neovisan način, uzorkuje signal niskom frekvencijom i kasnije koristi moć računanja za rekonstrukciju iz nečeg što se čini kao nepotpun skup podataka [7].

3.2 Problem uzorkovanja

Razmotrimo mehanizme uzorkovanja ili prikupljanja podataka u kojima je informacija o signalu $f(t)$ dobivena linearnim funkcijama koje spremaju vrijednosti.

$$y_k = \langle f | \varphi_k \rangle, \quad k = 1, \dots, m \quad (16)$$

To jest, jednostavno koreliramo objekt koji želimo prikupiti s valnim oblicima $\varphi_k(t)$. Ako su osjetilni valni oblici Dirac delta funkcije, tada je y vektor uzorkovanih vrijednosti od f u vremenskoj ili prostornoj domeni. Ako su osjetilni valni oblici indikator funkcije piksela, tada je y skup podataka slike tipično prikupljen senzorima digitalne kamere. Ako su valni oblici sinusne funkcije, tada je y vektor Fourierovih koeficijenata, što je osjetilna tehnika korištena kod slika magnetske rezonance.

Iako se CS teorija može prikazati na primjeru vremenski kontinuiranih signala, u ovom radu će biti pokazana na primjeru diskretnih signala $f \in \mathbb{R}^n$.

Razlog za to je što je konceptualno jednostavnije i dostupna CS diskretna teorija je mnogo razvijenija.

Za ovu teoriju zanimljive su poduzorkovane situacije u kojima je broj dostupnih mjerenja m mnogo manji od dimenzije n signala f . Takvi problemi su prilično uobičajeni zbog mnoštva razloga. Na primjer, broj senzora može biti ograničen ili mjerenja mogu biti jako skupa kao što je slučaj u nekim procesima slikanja neurona. Ili osjetilni proces može biti spor tako da se objekt može mjeriti samo par puta kao što je slučaj kod magnetske rezonance. Te okolnosti postavljaju važna pitanja. Je li točna rekonstrukcija moguća iz $m \ll n$ mjerenja? Je li moguće dizajnirati $m \ll n$ osjetilnih valnih oblika i tako prikupiti skoro sve informacije o f ? I kako se f može aproksimirati iz tih informacija?

Kad je problem tako postavljen, izgleda prilično izazovno jer je potrebno riješiti neodređen sustav linearnih jednadžbi. Ako A označava osjetilnu matricu $m \times n$ s vektorima $\varphi_1^*, \dots, \varphi_m^*$ kao redovima (gdje $*$ označava kompleksno transponiranje), onda je proces rekonstrukcije $f \in \mathbb{R}^n$ iz $y = Af \in \mathbb{R}^m$ nepotpuno postavljen u općenitom slučaju kad je $m < n$, što znači da postoji beskonačno mnogo kandidata signala \tilde{f} za koji vrijedi $A\tilde{f} = y$.

Međutim, može se zamisliti izlaz iz ovog problema ako se osloni na realistične modele f koji prirodno postoje. Shannonova teorija nam govori da ako $f(t)$ zaista ima mali frekvencijski pojas, onda će mali broj uniformnih uzoraka biti dovoljan za rekonstrukciju. U nastavku ćemo vidjeti da se rekonstrukcija može postići za puno širu klasu signala [7].

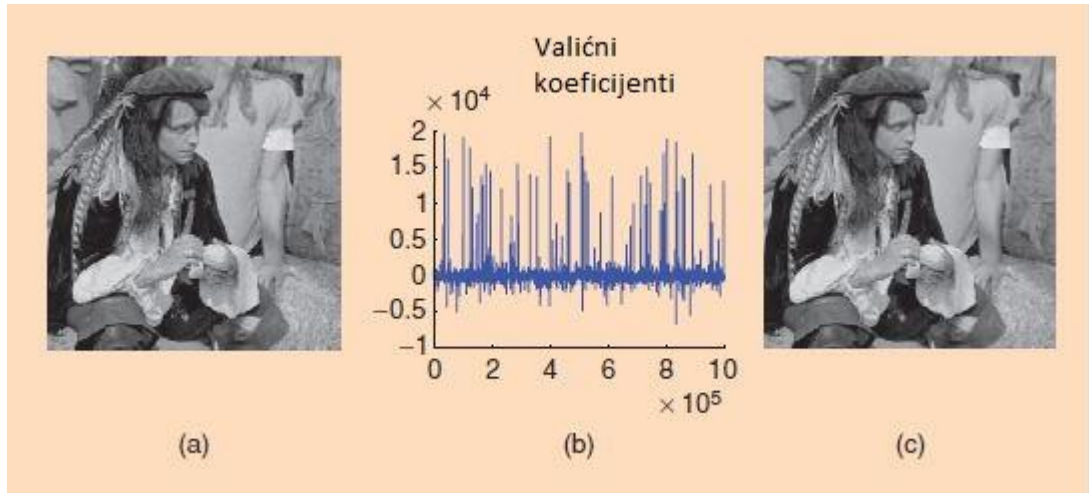
3.3 Inkoherencija i uzorkovanje rijetkih signala

U ovom odlomku će se obraditi dvije osnovne premise koje su temelj za funkcioniranje sažimajućeg očitavanja; rijetkost i inkoherencija.

3.3.1 Rijetkost

Mnogi prirodni signali imaju sažetu reprezentaciju kad su predstavljeni u prikladnoj bazi.

Primjer možemo vidjeti na slici 9., gdje je na 9(a) predstavljena slika i skoro svi njeni pikseli imaju vrijednosti različite od nule te na 9(b) valična transformacija te slike, tj. njeni valični koeficijenti koji su sažeti jer je većina koeficijenata mala, dok samo par većih sadrži većinu informacije.



Slika 9 Prikaz slike(a), njezinih valičnih koeficijenata(b) i rekonstrukcije(c) [7]

Matematički izraženo, imamo vektor $f \in \mathbb{R}^n$ kao što je n -pikselna slika na slici 9(a) kojeg proširujemo u ortonormiranoj bazi kao što je valična baza

$\psi = [\psi_1 \psi_2 \dots \psi_n]$ i slijedi:

$$f(t) = \sum_{i=1}^n x_i \psi_i(t), \quad (17)$$

gdje je x slijed koeficijenata od f , $x_i = \langle f | \psi_i \rangle$. Prikladno je izraziti f kao ψx , gdje je ψ $n \times n$ matrica s $\psi_1 \psi_2 \dots \psi_n$ kao stupcima. Implikacija rijetkosti je sada jasna, kad signal ima rijetku ekspanziju, mali koeficijenti se mogu odbaciti bez primjetnog perceptualnog gubitka.

Formalno, $f_S(t)$ je dobiven zadržavajući samo S najvećih vrijednosti od x_i u izrazu (17). Po definiciji, $f_S := \psi x_S$, gdje x_S predstavlja vektor koeficijenata x_i gdje su svi koeficijenti postavljeni u nulu osim S najvećih.

Ovaj vektor je rijedak u strogom smislu jer su svi osim par njegovih koeficijenata nula i nazivamo ga S -rijetkim jer ima S koeficijenata različitih od nule. Kako je ψ ortonormirana baza, vrijedi:

$$\|f - f_S\|_{l_2} = \|x - x_S\|_{l_2}$$

Ako je x rijedak ili se može komprimirati u smislu da sortirane amplitude od x_i padaju brzo, onda je x dobro aproksimiran sa x_S i greška $\|f - f_S\|_{l_2}$ je mala.

Jednostavno rečeno, možemo odbaciti veliku količinu koeficijenata s malo gubitka. Slika 9(c) pokazuje primjer gdje je perceptualni gubitak jedva primjetan između megapikselne slike i njezine aproksimacije dobivene odbacivanjem 97.5 % koeficijenata.

Taj princip je podloga mnogim modernim koderima kao što je JPEG-2000 jer je jednostavna metoda za podatkovnu kompresiju izračunati x iz f i onda kodirati vrijednosti i lokacije S najznačajnijih koeficijenata. Takav proces zahtijeva poznavanje svih n koeficijenata od x jer lokacije značajnih dijelova informacije možda nisu poznate unaprijed (ovisne su o signalu). U našem primjeru, koeficijenti su zgusnuti u rubovima slike.

U daljnjoj razradi, bit će pokazano kako rijetkost ima utjecaja na sam proces akvizicije podataka [7].

3.3.2 Inkoherentno uzorkovanje

Pretpostavimo da je dan par (ϕ, ψ) ortobaza u \mathbb{R}^n . Prva baza ϕ je korištena za uzorkovanje objekta f kao u (16) i druga je korištena za reprezentaciju f .

Restrikcija na par ortobaza nije važna i poslužit će samo za pojednostavljenje.

Koherencija između osjetilne baze ϕ i reprezentacijske baze ψ je definirana:

$$\mu(\phi, \psi) = \sqrt{n} \cdot \max_{1 \leq k, j \leq n} |\langle \phi_k | \psi_j \rangle| \quad (18)$$

Najjednostavnije rečeno, koherencija mjeri najveću korelaciju između bilo koja dva elementa od ϕ, ψ . Ako ϕ i ψ sadrže korelirane elemente, koherencija je velika.

Ako elementi nisu korelirani, onda je mala. Koliko mala ili velika, odgovor daje linearna algebra: $\mu(\phi, \psi) \in [1, \sqrt{n}]$.

Za sažimajuće očitavanje je važna mala koherencija parova baza.

Primjeri tih parova su ϕ kanonska baza $\phi_k(t) = \delta(t - k)$ i ψ Fourierova baza $\psi_j(t) = n^{-1/2} e^{i 2\pi j t/n}$. Kako je ϕ osjetilna baza, ovo je klasična shema uzorkovanja. Ovaj vremensko-frekvencijski par zadovoljava $\mu(\phi, \psi) = 1$ i tako osigurava maksimalnu inkoherenciju. Nadalje, Dirac delta funkcije i sinusne funkcije su maksimalno inkoherentne ne samo u jednoj dimenziji, već u bilo kojoj (dvije, tri, ...).

Slučajne matrice su u velikoj mjeri inkoherentne s bilo kojom fiksnom bazom ψ . Možemo slučajno izabrati ortobazu ϕ tako da ortonormiramo n vektora koji su nezavisno uzorkovani i uniformni na jediničnoj sferi. Tada je s visokom vjerojatnošću, koherencija između ϕ i ψ oko $\sqrt{2 \log n}$.

Općenito, slučajni valni oblici $\phi_k(t)$ s nezavisnim jednoliko distribuiranim koeficijentima, npr. gausovski ili binarni ± 1 koeficijenti će imati nisku koherenciju s bilo kojom fiksnom bazom ψ [7].

3.3 Poduzorkovanje i rekonstrukcija rijetkog signala

U idealnom slučaju, htjeli bismo izmjeriti svih n koeficijenata od f , ali možemo promatrati samo njihov podskup i prikupiti podatke

$$y_k = \langle f | \phi_k \rangle, \quad k \in M, \quad (19)$$

gdje je $M \subset \{1, \dots, n\}$ podskup kardinalnosti $m < n$. S tom informacijom, signal se rekonstruira koristeći minimizaciju ℓ_1 normom ($\|x\|_{\ell_1} := \sum_i |x_i|$). Predložena rekonstrukcija f^* je dana sa $f^* = \psi x^*$ gdje je x^* rješenje konveksnog optimizacijskog programa:

$$\min_{\tilde{x} \in \mathbb{R}^n} \|\tilde{x}\|_{\ell_1} \quad \text{podložno} \quad y_k = \langle \phi_k, \psi \tilde{x} \rangle \quad (20)$$

To znači da unutar svih rješenja $\tilde{f} = \psi \tilde{x}$ konzistentnih s podacima, izaberemo ono čiji slijed koeficijenata ima minimalnu ℓ_1 normu. Upotreba ℓ_1 norme datira otprije nekoliko desetljeća. Međutim, ta minimizacija nije jedini način za rekonstrukciju rijetkog signala, predložena su još neka rješenja kao što je pohlepni algoritam.

Prvi rezultat pokazuje da kad je f dovoljno rijedak, rekonstrukcija koristeći ℓ_1 minimizaciju je dokazano točna.

Teorem 2: Fiksirajmo $f \in \mathbb{R}^n$ i pretpostavimo da je slijed koeficijenata x od f u bazi ψ S -rijedak. Odaberimo n uniformnih slučajnih mjerenja u ϕ domeni.

Tada ako vrijedi:

$$m \geq C \cdot \mu(\phi, \psi)^2 \cdot S \cdot \log n \quad (21)$$

za neku pozitivnu konstantu C , rješenje jednadžbe (20) je točno s dokazanom vjerojatnošću.

Vjerojatnost uspjeha prelazi $1 - \delta$ ako je $m \geq C \cdot \mu(\phi, \psi)^2 \cdot S \cdot \log n / \delta$.

Oko ovog teorema bitne su tri stvari:

- 1) Uloga koherencije je u cijelosti transparentna. Što je manja koherencija, manje je uzoraka potrebno i zato se CS temelji na niskoj koherenciji.
- 2) Nema informacijskog gubitka mjerenjem bilo kojeg seta m koeficijenata, što može biti puno manje od zahtjeva veličine signala. Ako je $\mu(\phi, \psi)$ jednak ili blizu jedinice, tada je dovoljno $S \log n$ uzoraka umjesto n .
- 3) Signal f može biti točno rekonstruiran iz sažetog skupa podataka minimizacijom konveksne funkcije koja ne pretpostavlja znanje o broju nenul koordinata od x , njihove lokacije ili amplitude za koje pretpostavljamo da su *a-priori* potpuno nepoznate. Algoritam se samo provede i ako je signal dovoljno rijedak, dobije se točna rekonstrukcija.

Teorem sugerira konkretan protokol akvizicije: neadaptivno uzorkovanje u inkoherentnoj domeni i nakon akvizicijskog koraka provođenje linearnog programiranja. Slijeđenje tog protokola bi dovelo do akvizicije signala u komprimiranom obliku. Sve što je potrebno je dekođer za „dekompresiju” podataka, što je uloga ℓ_1 minimizacije.

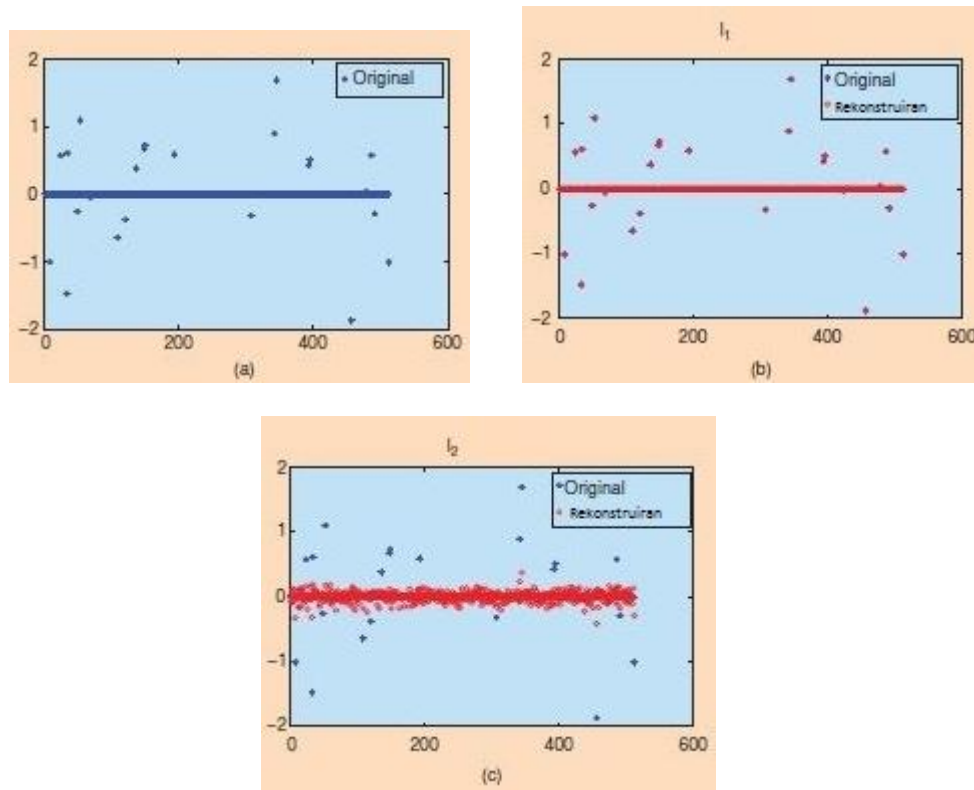
Ovaj teorem slučajnog inkoherentnog uzorkovanja proširuje raniji rezultat o uzorkovanju spektralno rijetkih signala koji je pokazao da je korištenje slučajnosti jako efikasan osjetilni mehanizam i podložan pravim dokazima.

Recimo da smo zainteresirani za uzorkovanje ultra-širokopojasnog signala koji je spektralno rijedak i oblika $f(t) = \sum_{j=0}^{n-1} x_j e^{i2\pi jt/n}$, $t = 0, \dots, n-1$, gdje je n jako velik, ali broj ne-nul komponenti x_j je jednak ili manji od S , koji je mali.

Ne znamo koje su frekvencije aktivne, ni kolike su amplitude aktivnog skupa. Kako aktivni skup nije nužno podskup slijednih cijelih brojeva, Nyquist-Shannonov teorem nije od pomoći.

U ovom specijalnom slučaju, teorem 2 tvrdi da se signal s proizvoljnim i nepoznatim frekvencijama veličine S može rekonstruirati iz $S \log n$ vremenskih uzoraka. Štoviše, ti uzorci ne moraju biti pažljivo izabrani, skoro svaki skup uzoraka će biti dovoljan.

Ilustrativni primjer je prikazan na slici 10(b), gdje je točno rekonstruiran realni signal iz 60 kompleksnih Fourierovih koeficijenata koristeći ℓ_1 minimizaciju. Na slici 10(c) je umjesto ℓ_1 minimizacije korištena ℓ_2 minimizacija i očito je da ne daje dobru aproksimaciju polaznog signala.



Slika 10 Usporedba L_1 i L_2 norme u primjeni sažimajućeg očitavanja [7]

Pogledajmo koju ulogu ima vjerojatnost u svemu ovome. Ključna točka je da je za dobivanje korisnih i moćnih rezultata, potrebno posegnuti za doprinosima vjerojatnosti jer se ne mogu koristiti svi mogući skupovi mjerenja veličine m . Objašnjenje slijedi.

Postoje posebni rijetki signali koji iščezavaju skoro svugdje u ϕ domeni. Drugim riječima, mogu se pronaći rijetki signali f i veliki podskupovi veličine gotovo n (primjerice $n - S$) za koje vrijedi $y_k = \langle f, \varphi_k \rangle = 0$ za sve $k \in M$. S jedne strane, koristeći takve podskupove, moglo bi se dogoditi da se dobije niz nula iz kojih nikakav algoritam ne bi mogao rekonstruirati točan signal. S druge strane, teorem jamči da je skupova za koje točna rekonstrukcija nije moguća zanemarivo malen (velika negativna potencija od n). Tako se mora tolerirati vjerojatnost neuspjeha koja je ekstremno mala. Za praktične primjene, vjerojatnost neuspjeha je nula, pod uvjetom da je broj uzoraka dovoljno velik.

Zanimljivo, studija posebnih rijetkih signala također pokazuje da je potrebno najmanje $\mu^2 S \log n$ uzoraka za točnu rekonstrukciju. S manje uzoraka, vjerojatnost da informacija bude izgubljena je prevelika i onda je rekonstrukcija bilo kojom metodom nemoguća. Primjerice, kad koherencija iznosi 1, ne treba nam više od $S \log n$ uzoraka, ali ne možemo raditi s manje od toga.

Možemo uzeti primjer slike 9(c) koja ima 25 000 ne-nul valićnih koeficijenata. Informacije su se prikupile koristeći 96 000 inkoherentnih mjerenja i rješavanjem jednadžbe (20). Rekonstrukcija ℓ_1 –minimizacijom je savršena, tj. vrijedi $f^* = f$. Ovaj primjer pokazuje da je dovoljan broj uzoraka oko $4 \times$ veći od razine rijetkosti. Mnogi istraživači su prijavili sličan empirijski uspjeh. Postoji *de facto* poznato četiri-na-jedan praktično pravilo koje kaže da je za točnu rekonstrukciju potrebno oko 4 inkoherentna uzorka za svaki nepoznati ne-nul uzorak [7].

3.4 Sažeto očitavanje u šumovitom okruženju

Pokazano je da sažimajuće očitavanje ima moć rekonstruirati signale iz relativno malo uzoraka, međutim još jedno bitno svojstvo je da to isto može napraviti i iz okruženja u kojem je prisutan šum.

Općeniti objekti kojima se bavimo nisu u potpunosti rijetki već približno rijetki pa se postavlja pitanje je li moguće postići točnu rekonstrukciju iz veoma poduzorkovanih mjerenja.

U bilo kojim realnim primjenama, podaci će biti komprimirani barem malom količinom šuma jer osjetilni uređaji nemaju beskonačnu preciznost. Zbog toga je važno da tehnika sažimajućeg očitavanja bude otporna na takve nepravilnosti.

U najgorem slučaju, male preturbacije u zadanim uzorcima trebaju uzrokovati male pretpurbacije u rekonstruiranim podacima.

Zbog lakšeg izračuna, ovi problemi mogu se predočiti jednostavnom jednadžbom:

$$y = Ax + z, \quad (22)$$

gdje je $x \in \mathbb{R}^n$ rekonstruirani vektor, A je $m \times n$ osjetilna matrica koja daje informaciju o x te je z stohastička ili deterministička pogreška. Treba imati na umu da ovaj zapis ima smisla ako podrazumijevamo da je x slijed koeficijenata objekta u prikladnoj bazi ϕ .

Važno svojstvo koje je potrebno da bi se dokazala otpornost teorije sažimajućeg očitavanja je ograničeno izometrijsko svojstvo (eng. restricted isometry property), tzv. RIP.

Definicija 3. Za svaki cijeli broj $S = 1, 2, 3, \dots$ definirana je izometrijska konstanta δ_S matrice A kao najmanji broj za kojega vrijedi da:

$$(1 - \delta_S) \|x\|_{\ell_2}^2 \leq \|Ax\|_{\ell_2}^2 \leq (1 + \delta_S) \|x\|_{\ell_2}^2 \quad (23)$$

sadržava sve S -rijetke vektore x .

Ne tako restriktivno može se reći da matrica A podliježe RIP pravilu reda S ako δ_S nije tako blizu vrijednosti jedan. Kad ovo svojstvo vrijedi, A približno čuva euklidsku duljinu S -rijetkih signala, što zauzvrat implicira da S -rijetki vektori ne mogu biti u nul-prostoru matrice A , što je korisno jer se inače ne bi mogli rekonstruirati.

Ekvivalentan opis svojstva RIP je da su svi podskupovi S -stupaca uzetih iz A gotovo ortogonalni. Ne mogu biti u potpunosti ortogonalni jer matrica A nije kvadratna.

Da bi veza između RIP svojstva i teorije sažimajućeg očitavanja bila očita, zamislimo da želimo pribaviti S -rijetke signale pomoću A . Pretpostavimo da je δ_{2S} značajno manji od jedan.

To implicira da sve udaljenosti parova između S -rijetkih signala moraju biti dobro očuvane u mjernom prostoru. Za sve S -rijetke vektore x_1, x_2 Vrijedi:

$$(1 - \delta_{2S})\|x_1 - x_2\|_{\ell_2}^2 \leq \|Ax_1 - Ax_2\|_{\ell_2}^2 \leq (1 + \delta_{2S})\|x_1 - x_2\|_{\ell_2}^2 \quad (24)$$

U nastavku će biti prikazano kako ova činjenica ima utjecaja na stvaranje učinkovitog i otpornog algoritma.

3.4.1 Rekonstrukcija signala iz malog skupa podataka

Ako RIP svojstvo vrijedi, tada sljedeći linearni program omogućava točnu rekonstrukciju:

$$\min_{\tilde{x} \in \mathbb{R}^n} \|\tilde{x}^*\|_{\ell_1} \quad \text{podložno} \quad Ax^* = y = (Ax) \quad (25)$$

Teorem 3. Pretpostavimo da je $\delta_{2S} < \sqrt{2} - 1$. Tada za rješenje x^* iz jednadžbe (25) vrijedi:

$$\|x^* - x\|_{\ell_2} \leq \frac{C_0 \|x - x_S\|_{\ell_1}}{\sqrt{S}} \text{ i } \|x^* - x\|_{\ell_2}^2 \leq C_0 \|x - x_S\|_{\ell_1} \quad (26)$$

za neku konstantu C_0 , gdje je x_S vektor x kojemu su sve osim najvećih S komponenti postavljene u nulu [12].

Zaključci ovog teorema su jači od zaključaka teorema 2. Ako je x S -rijedak, onda je $x = x_S$ i rekonstrukcija je točna. Ali teorem 3 se bavi svim signalima. Ako x nije S -rijedak, tada (26) procjenjuje da će kvaliteta rekonstruiranog signala biti takva kao da je netko a-priori znao lokaciju najvećih S -komponenti i izravno samo njih mjerio. Još jedna razlika je da ovaj teorem daje deterministički rezultat, tj. ne uključuje vjerojatnost. Ako izabrana matrica A podliježe hipotezi iz teorema, možemo biti sigurni da ćemo točno rekonstruirati sve S -rijetke vektore te S -najvećih vrijednosti svih vektora i nema vjerojatnosti neuspjeha.

U ovom trenutku još nedostaje veza između vrijednosti S koja označava broj komponenti koje se mogu učinkovito rekonstruirati i vrijednosti m koja predstavlja broj mjerenja ili redaka matrice.

Da bi se postigli zadovoljavajući rezultati, potrebno je pronaći matrice koje zadovoljavaju RIP svojstvo s vrijednostima S bliskima m . Izgradnja tih matrica bit će prikazana u poglavlju 4.

3.4.2 Rekonstrukcija signala iz šumovitih podataka

Dani su nam podaci koji sadrže šum kao što je prikazano u jednadžbi (22) i koristimo ℓ_1 -minimizaciju s opuštenim uvjetima za rekonstrukciju:

$$\min \|x^*\|_{\ell_1} \quad \text{podložno} \quad \|Ax - y\|_{\ell_2} \leq \varepsilon \quad (27)$$

gdje ε ograničava količinu šuma u podacima.

Osim tog rješenja, u literaturi su predloženi tzv. *Dantzig selector* ili kombinatorički optimizacijski program predložen od Haupta i Novaka. Pokazalo se da ti algoritmi daju dobra rješenja u slučaju Gaussovog šuma s ograničenom varijancom.

Problem (27) se često naziva LASSO, što označava metodu regresijske analize i može se učinkovito riješiti. Više informacije može se pronaći u literaturi [11].

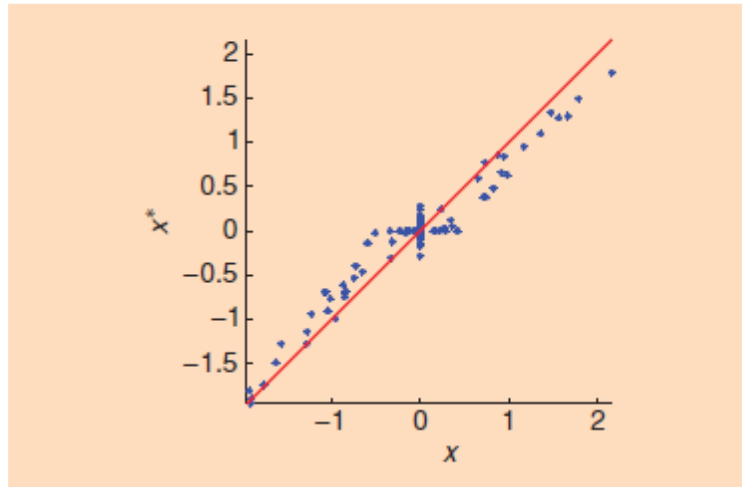
Teorem 4. Pretpostavimo da je $\delta_{2S} < \sqrt{2} - 1$. Tada za rješenje x^* iz jednadžbe (27) vrijedi:

$$\|x^* - x\|_{\ell_2} \leq \frac{C_0 \|x - x_S\|_{\ell_1}}{\sqrt{S}} + C_1 \varepsilon \quad (28)$$

za neke konstante C_0 i C_1 [12].

Teorem 4 pokazuje da je rekonstrukcijska pogreška ograničena zbrojem dva uvjeta. Prvi uvjet je greška koja bi se mogla dogoditi kad koristimo podatke bez šuma, a drugi je proporcionalan razini šuma. Nadalje konstante C_0 i C_1 su tipično male. Na slici 11 prikazani su rezultati rekonstrukcije iz šumovitih podataka za konstante $\delta_{2S} = 0.25$, $C_0 \leq 5.5$ i $C_1 \leq 6$.

Taj rezultat pokazuje da je metoda sažimajućeg očitavanja praktičan i otporan osjetilni mehanizam. Radi sa svim vrstama signala, koji ne moraju nužno biti rijetki, i dobro se nosi sa šumom.



Slika 11 Rekonstrukcija signala iz šumovitih podataka [7]

Na slici 11 prikazan je signal na horizontalnoj osi te signal x^* na vertikalnoj osi. U ovom primjeru korišten je ulazni skup podataka veličine $n = 512$ te skup mjerenja $m = 256$. Korišteni signal je 64-rijedak. U modelu (22), osjetilna matrica ima $N(0, 1/m)$ nezavisnih jednoliko distribuiranih koeficijenta i z je vektor bijelog Gaussovog šuma prilagođen tako da vrijedi $\frac{\|Ax\|_{\ell_2}}{\|z\|_{\ell_2}} = 5$.

Također, vrijedi $\|x^* - x\|_{\ell_2} \approx 1.3\varepsilon$.

3.5 Sažeto uzorkovanje

Akvizicija podataka se obično radi tako da se prikupi ogromna količina podataka i onda se veliki dio odbaci u fazi kompresije kako bi se olakšalo čuvanje i prijenos. Taj proces se može usporediti s tim da netko uzme polje piksela visoke rezolucije f , izračuna sve koeficijente transformacije, kodira najveće koeficijente i odbaci ostale te na kraju završi sa f_s . Proces akvizicije masovnih podataka praćen kompresijom je iznimno nepotreban i usporediv s korištenjem mekapikselne kamere te zatim kodiranja slike u veličinu od samo par kilobajta.

Sažeto očitavanje radi potpuno drugačije i ponaša se kao da je moguće prikupiti samo važne informacije o objektu koji nas zanima. Uzimajući otprilike $O(S \log n/S)$ slučajnih projekcija, moguće je rekonstruirati signal s točnošću najmanje dobro kao f_S , najbolja S -aproximacija.

Drugim riječima, CS mjerni protokoli u biti prevode analogne podatke u već komprimirani digitalni oblik tako da se mogu dobiti signali visoke rezolucije od samo par senzora. Sve što je potrebno nakon akvizicije je „dekompresija” mjerenih podataka [7].

3.6 Primjene sažimajućeg očitavanja

Činjenica da kompresibilni signal može biti efikasno uhvaćen koristeći broj inkoherentnih mjerenja koji je proporcionalan informacijskoj razini $S \ll n$ ima implikacije koje su dalekosežne i tiču se velikog broja mogućih primjena. Neke od tih primjena su:

- Kompresija podataka – U nekim situacijama, rijetka baza ψ može biti nepoznata koderu ili nepraktična za implementaciju za kompresiju podataka. Međutim, slučajno dizajnirana baza ϕ može biti smatrana univerzalnom strategijom za kodiranje jer je njeno dizajniranje neovisno o bazi ψ . Poznavanje i sposobnost implementacije ψ je potrebno samo za dekodiranje ili rekonstrukciju f . Ta univerzalnost može biti posebno od pomoći kad je riječ o kodiranju distribuiranih izvora u okruženju više signala, kao što su senzorske mreže.
- Kanalsko kodiranje – Principi sažimajućeg očitavanja uključujući rijetkost, slučajnosti i konveksnu optimizaciju mogu biti preokrenuti i primijenjeni za dizajniranje brzih kodova za ispravljanje grešaka kako bi se spriječile greške tijekom prijenosa.
- Inverzni problemi – U nekim situacijama, jedini način za prikupljanje f može biti korištenje mjernog sustava ϕ . Pretpostavljajući da rijetka baza ψ postoji za f i inkoherentna je s ϕ , efikasno uzorkovanje može biti moguće. Jedna takva primjena uključuje angiografiju magnetske rezonance, gdje ϕ snima podskup Fourierove transformacije i željena slika f je rijetka u vremenskoj ili valičnoj domeni.

- Akvizicija podataka – U nekim važnim situacijama potpuno prikupljanje n diskretnih vremenskih uzoraka analognog signala može biti teško za izvesti ili naknadno komprimirati. U tom slučaju bi od pomoći moglo biti dizajniranje fizičkog uređaja za uzorkovanje koji izravno snima diskretna inkohherentna niskofrekvencijska mjerenja analognog signala.

Ta primjena sugerira da matematičke i računske metode mogu imati veliki utjecaj u područjima gdje konvencionalan dizajn hardvera ima značajna ograničenja. Primjerice konvencionalni uređaji za snimanje kao što su CCD ili CMOS tehnologija su ograničeni na vidljivi spektar. Međutim, CS kamera koja prikuplja inkohherentna mjerenja koristeći polje digitalnog mikrozrcala (potreban je samo jedan fotoosjetljivi element umjesto milijuna) bi mogla značajno proširiti te mogućnosti [7].

Na slikama 12, 13 i 14 prikazani su neki od primjera korištenje sažimajućeg očitavanja u realnim aplikacijama.

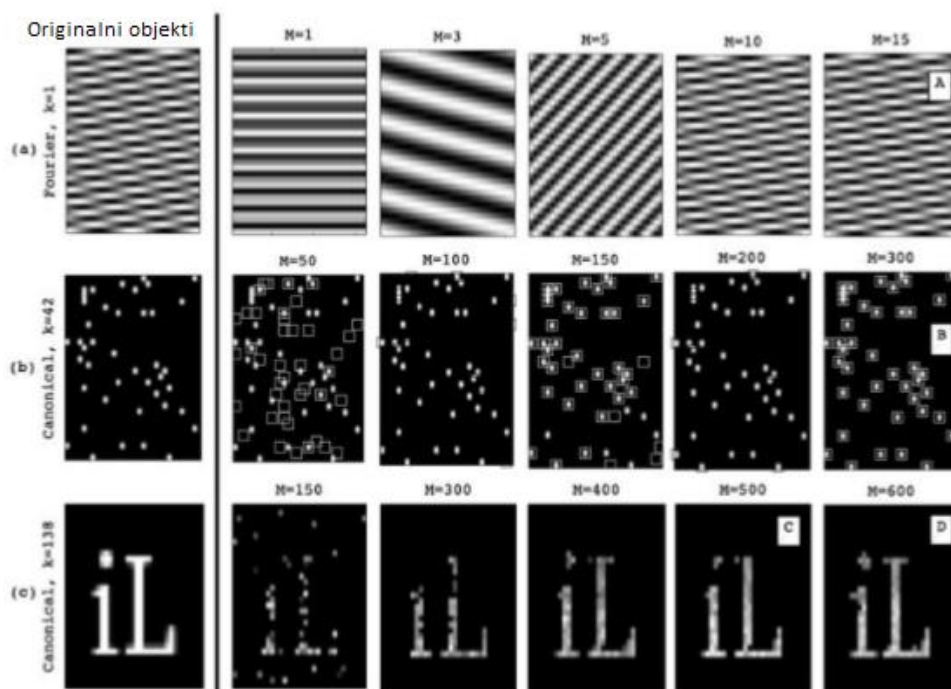
Na slici 12 prikazana je rekonstrukcija slike iz sto tisuća slučajnih mjerenja (desno), dok je originalna slika imala milijun piksela (lijevo).

Na slici 13 u lijevom stupcu su prikazani originalni objekti veličine $N = 1024$ piksela. K predstavlja broj ne-nul koeficijenata, a m broj uzetih mjerenja. S desne strane prikazani su rekonstruirani objekti.

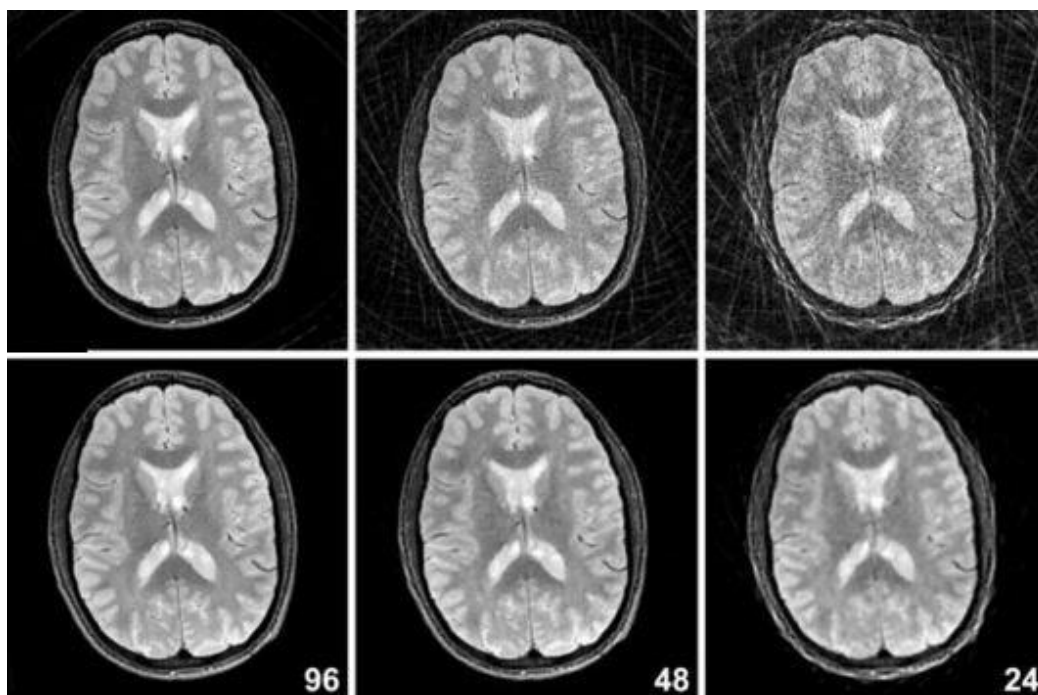
Na slici 14 prikazana je rekonstrukcija ljudskog mozga iz 96, 48 i 24 radijalnih presjeka kao jedan od primjera primjena sažimajućeg očitavanja. Slike u gornjem redu su dobivene koristeći konvencionalne rekonstrukcijske tehnike, dok je donji red dobiven koristeći paralelne slike i sažimajuće očitavanje. Više informacija moguće je pronaći u literaturi [8].



Slika 12 Primjena sažimajućeg očitavanja na slike [13]



Slika 13 Rekonstrukcija objekata iz različitog broja mjerenja [14]



*Slika 14 Rekonstrukcija slike ljudskog mozga iz različitog broja radijalnih presjeka
[8]*

4. Primjena sažimajućeg očitavanja

Da bismo koristili tehniku sažimajućeg očitavanja, potrebno je pronaći mjernu matricu i rijetku bazu pogodnu za korištenje pri rekonstrukciji ulaznog skupa podataka te izabrati rekonstrukcijski algoritam. U nastavku će biti predloženo više mogućih opcija.

4.1 Primjeri rijetkih baza

Općenito govoreći, bilo koja $N \times N$ ortogonalna matrica može biti rijetka baza. Primjeri nekih češće korištenih su Fourierova baza, diskretna kosinusna transformacija (DCT), Radonova transformacija i valićna baza (DWT). U nastavku su prikazane usporedbe nekih rijetkih baza na primjerima slika.



Slika 15 Usporedba algoritama za kodiranje JPEG i JPEG2000 [15]

Na slici 15 prikazana je usporedba algoritma JPEG koji je standardno implementiran na digitalnim kamerama i kao bazu koristi DCT na 8×8 blokovima te algoritma JPEG2000 koji koristi valićnu bazu. Ona je mnogo prikladnija za slike s rubovima jer daje manje S koeficijenata pa je bolja za faktor 2 od JPEG-a.

Na slici 16 prikazana je usporedba pri korištenju DCT linearne i nelinearne aproksimacije te valićne nelinearne. Linearna podrazumijeva zadržavanje S koeficijenata na fiksnim lokacijama, dok nelinearna zadržava S najvećih koeficijenata.



Slika 16 Usporedba valićne i DCT baze [15]

4.2 Primjeri osjetilnih matrica

Osjetilne matrice mogu biti slučajne i determinističke. U većini primjena kao što su obrada medicinskih slika, analiza geofizikalnih podataka, komunikacije i ostalo problemi obrade podataka koriste se slučajne osjetilne matrice. Međutim, upotreba determinističkih matrica je jako poželjna na strukturama koje omogućuju brzu implementaciju sa smanjenim zahtjevima pohrane. U nastavku su prikazani razni primjeri obiju vrsta osjetilnih matrica.

4.1.1 Slučajne osjetilne matrica

Nastavno na potpoglavlje 3.4.1, potrebno je pronaći matrice koje posjeduju RIP svojstvo s vrijednostima ne-nul elemenata S bliskima broju mjerenja m , tj. želimo pronaći matrice sa svojstvom da su stupci matrice iz proizvoljnih podskupova gotovo ortogonalni. Što su veći podskupovi, to je bolje.

Za to je ponovno bitna vjerojatnost, tj. slučajan odabir. Slučajne osjetilne matrice su jednostavne za korištenje jer su inkohherentne s fiksiranim rijetkim bazama.

Takve matrice mogu se izgraditi na sljedeće načine:

- 1) slučajnim uniformnim uzorkovanjem n stupaca na jediničnoj sferi \mathbb{R}^m
- 2) uzorkovanjem nezavisnih jednoliko distribuiranih ulaza iz normalne razdiobe sa srednjom vrijednošću 0 i varijancom $1/m$
- 3) uzorkovanjem slučajne projekcije P kao u poglavlju 3.3.2 i normalizacijom

$$A = \sqrt{\frac{n}{m}} P$$

- 4) uzorkovanjem nezavisnih jednoliko distribuiranih ulaza iz simetrične Bernoullijeve razdiobe $P\left(A_{i,j} = \pm \frac{1}{\sqrt{m}}\right) = \frac{1}{2}$ ili neke druge gausovske razdiobe

S velikom vjerojatnošću, sve ove matrice posjeduju RIP svojstvo pod uvjetom da vrijedi:

$$m \geq C S \log \frac{n}{s} \quad (29)$$

gdje je C konstanta različita u svakoj instanci.

U svim navedenim primjerima vjerojatnost da osjetilne matrice ne zadovoljavaju RIP pravilo pod uvjetom da vrijedi (29) pada eksponencijalno sa m .

Zanimljivo je to da ne postoje osjetilne matrice i rekonstrukcijski algoritmi koji mogu dati zaključke kao u teoremu 3 sa značajno manjim brojem uzoraka od onog koji je dan u (29). U tom smislu korištenje slučajnih matrica zajedno s ℓ_1 minimizacijom je skoro optimalna osjetilna strategija.

RIP svojstvo se može postići i korištenjem para ortobaza kao u poglavlju 3.3.1. Ako vrijedi $A = R \phi \psi$, gdje R uniformno jednoliko izvlači m koordinata, dovoljno je imati:

$$m \geq C S (\log n)^4 \quad (30)$$

mjerjenja da bi RIP svojstvo bilo održano s velikom vjerojatnošću. Ako netko želi još manju vjerojatnost neuspjeha koja pada eksponencijalno, potrebno je koristiti eksponent 5. To dokazuje da se mogu stabilno i točno rekonstruirati gotovo rijetki signali iz jako malog broja uzoraka u inkohherentnoj domeni.

Naposljetku, RIP pravilo može biti održano i za osjetilne matrice $A = \phi \psi$, gdje je ψ proizvoljna ortobaza i ϕ je $m \times n$ mjerna matrica dobivena slučajno iz prikladne razdiobe. Ako se ψ fiksira i izabere ϕ između navedenih opcija 1) do 4), tada će RIP pravilo vrijediti za matricu $A = \phi \psi$ ako vrijedi (29), gdje je C ponovno neka konstanta koja ovisi o instanci.

Ovakve slučajne mjerne matrice ϕ su univerzalne jer nije potrebno poznavati rijetku bazu prilikom dizajniranja osjetilnog sustava [7].

4.1.2 Determinističke osjetilne matrice

Iako slučajne osjetilne matrice omogućuju rekonstrukciju s visokom vjerojatnošću, također imaju neke mane kao što su pretjerana kompleksnost u rekonstrukciji, značajni zahtjevi za pohranu te nepostojanje efikasnog algoritma za provjeru zadovoljava li osjetilna matrica RIP svojstvo uz malu konstantu δ_S . Problem rekonstrukcije može postati težak ako se radi o signalu velikih dimenzija jer je potrebno konstruirati mjernu matricu koja zadovoljava RIP uz malu konstantu δ_S . Zbog toga se koriste specifične strukture determinističkih osjetilnih matrica za rješavanje tih problema.

Determinističke matrice se dijele u dvije kategorije: one temeljene na koherenciji, primjerice deterministička konstrukcija osjetilnih matrica korištenjem algebarskih krivulja preko konačnih polja i one temeljene na RIP pravilu, primjerice *chirp* osjetilne matrice, Reed-Mullerovi kodovi drugog reda, binarni Bose-Chaudhuri-Hocquenghem (BCH) kodovi i osjetilne matrice sa statističkim RIP pravilom. U nastavku će biti opisane *chirp* osjetilna matrica i osjetilna matrica sa statističkim restriktivnim izometrijskim svojstvom, ostale spomenute matrice se mogu pronaći u literaturi [15].

4.1.2.1 Chirp osjetilna matrica

Diskretni chirp duljine m ima sljedeći oblik:

$$A_{r,w} = \frac{1}{\sqrt{m}} \exp \left\{ \frac{2\pi i}{m} \omega l + \frac{2\pi i}{m} r l^2 \right\}, l, \omega, r \in \mathbb{Z}_m,$$

gdje je r frekvencija chirpa i ω osnovna frekvencija.

Puna chirp osjetilna matrica A dimenzija $m \times m^2$ može biti zapisana kao:

$$A_{chirp} = [U_{r1} U_{r2} \dots U_{rk}]$$

Svaka matrica U_{rt} ($t = 1, 2, \dots, \omega$) je $m \times m$ matrica sa stupcima koji su dobiveni od chirp signala koji imaju fiksnu chirp frekvenciju r_t s osnovnom frekvencijom ω koja varira od 0 do $m - 1$. Primjerice, za $k = 2$, $r, m, l \in \{0, 1\}$ dobivamo:

$$U_{r1} = \begin{bmatrix} 1 & 1 \\ 1 & e^{i\pi} \end{bmatrix}, \quad U_{r2} = \begin{bmatrix} 1 & 1 \\ e^{i\pi} & e^{i\pi+i2\pi} \end{bmatrix}$$

Konačno, deterministička chirp matrica A dimenzija 2×4 izgleda:

$$A_{chirp} = [U_{r1} \ U_{r2}] = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & e^{i\pi} & e^{i\pi} & e^{i\pi+i2\pi} \end{bmatrix}$$

Primijetimo da za $r = r_1 = 0$, matrica U_{r1} koja odgovara chirp frekvenciji r_1 postaje matrica diskretne Fourierove transformacije (DFT).

Većina chirp osjetilnih matrica podliježe brzom rekonstrukcijskom algoritmu složenosti $O(km \log m)$.

4.1.2.2 Osjetilna matrica sa statističkim restriktivnim izometrijskim svojstvom

Matrica A dimenzija $m \times n$ nazvana je StRIP reda k s konstantom δ i vjerojatnošću $1 - \varepsilon$ ako vrijedi:

$$P(\|Ax\|^2 - \|x\|^2) \geq 1 - \varepsilon, \quad (31)$$

s tim da x ima uniformu razdiobu između svih k -rijetkih vektora na prostoru \mathbb{R}^n s istom normom.

Matrica A dimenzija $m \times n$ nazvana je UStRIP (jedinostveni StRIP) reda k s konstantom δ i vjerojatnošću $1 - \varepsilon$ ako zadovoljava (31) i vrijedi:

$$\{z \in \mathbb{R}^n : Az = Ax\} = \{x\},$$

s vjerojatnošću koja premašuje $1 - \varepsilon$ te da x ima uniformu razdiobu između svih k -rijetkih vektora na prostoru \mathbb{R}^n s istom normom.

U literaturi [15] je pokazano da ovakve matrice mogu biti konstruirane koristeći chirp matrice, Reed-Mullerove kodove drugog reda i BCH kodove.

4.3 Rekonstrukcijski algoritmi za ℓ_1 – minimizaciju

ℓ_1 – minimizacija odnosi se na pronalaženje minimalnog rješenja ℓ_1 norme za pod-određeni linearni sustav $b = Ax$. U zadnje vrijeme predloženo je puno algoritama povećane brzine za rješenje tog problema. Neki od njih su:

- Projekcija gradijenta
- Homotopija
- Iterativno smanjenje praga
- Približni gradijent
- Povećani Lagrangeov množitelj

Više detalja o predloženim metodama može se pronaći u literaturi [18].

Jedan od algoritama implementiran u Matlabu je SeDuMi (eng. Self-Dual-Minimization) [17]. Implementira osobnu-dualnu ugradbenu tehniku za optimizaciju preko osobnih-dualnih homogenih konusa. Ta tehnika predložena od strane Ye, Todda i Mizuna [19] omogućuje rješavanje određenih optimizacijskih problema u jednoj fazi i vodi do optimalnog rješenja ili dokaza nemogućnosti. Prednosti njegove implementacije u Matlabu su što dopušta upotrebu kompleksnih podataka, u potpunosti iskorištava prednosti rijetkih signala, što vodi do značajnog ubrzanja algoritma i ima teoretski dokazanu granicu iteracija u najgorem slučaju koja iznosi $O(\sqrt{n} \log \frac{1}{\varepsilon})$. Algoritam SeDuMi je dio cvx optimizacijskog paketa u Matlabu [29].

Drugi široko upotrebljavani optimizacijski paket u Matlabu je l1-MAGIC paket koji rješava sedam različitih optimizacijskih problema.

Za maksimalnu računsku učinkovitost, rješenja za svaki od sedam problema implementirana su zasebno. Sva rješenja imaju istu osnovnu strukturu. Kod se može koristiti na dva različita načina.

Jedan od njih je na tzv. nižoj razini gdje je sustav konstruiran eksplicitno te riješen egzaktno, a drugi je na tzv. višoj razini gdje se koristi iterativni algoritam bez matričnog zapisa koji aproksimativno rješava sustav, primjerice konjugirani gradijenti ili CG, koji se koristi kad je broj linearni jednažbi prevelik da bi se koristio matrični zapis (korištenjem funkcije cgsolve) [28].

4.4 Mjera točnosti rekonstrukcije

Najčešća mjera za točnost rekonstrukcije signala je srednja kvadratna pogreška (eng. mean squared error) ili MSE. Srednja kvadratna pogreška ili srednja kvadratna devijacija mjeri prosjek kvadrata grešaka ili devijacija, tj. razliku između procjenitelja i onog što se procjenjuje [20].

Definicija 4. $MSE = \frac{1}{n} \sum_{i=1}^n (x' - x)^2,$

gdje je n broj uzoraka, x' rekonstruirani signal, x ulazni signal.

Kao drugu mjeru točnosti možemo uzeti srednju apsolutnu pogrešku (eng. mean absolute error) ili MAE. Srednja apsolutna greška pokazuje koliko su blizu bila predviđanja eventualnim ishodima. Mjeri prosjek apsolutnih grešaka [21].

Definicija 5. $MAE = \frac{1}{n} \sum_{i=1}^n |x' - x|,$

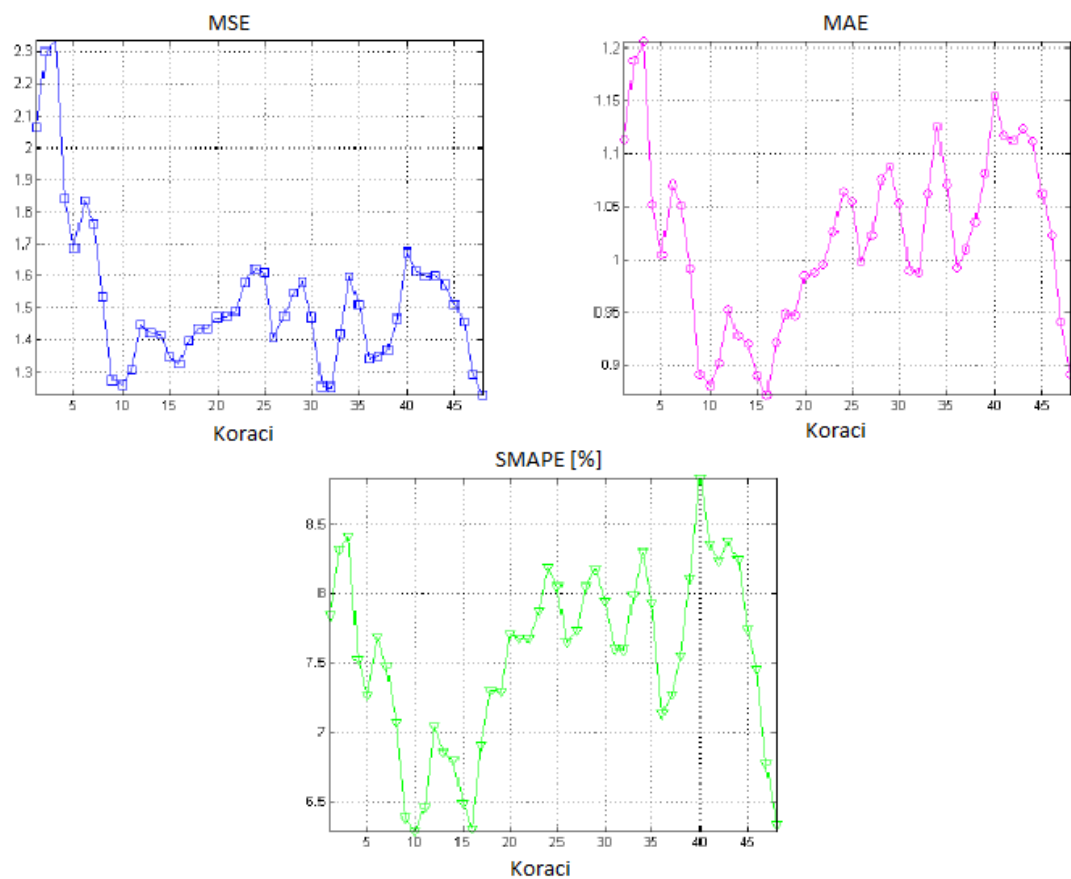
gdje je n broj uzoraka, x' rekonstruirani signal, x ulazni signal.

Treća mogućnost za mjeru pogreške je simetrična srednja apsolutna postotna pogreška (eng. Symmetric mean absolute percentage error) ili SMAPE. Simetrična srednja apsolutna postotna pogreška je mjera točnosti temeljena na relativnim pogreškama. Ako ta vrijednost iznosi 0%, prognoza je približno točna [25].

Definicija 6. $SMAPE = \frac{1}{n} \sum_{i=1}^n \frac{|x' - x|}{(|x'| + |x|)/2},$

gdje je n broj uzoraka, x' rekonstruirani signal, x ulazni signal.

Na slici 22 prikazana je usporedba tih triju metoda izračuna pogrešaka na primjeru podataka zahtjev za pitkom vodom [22].



Slika 17 Usporedba MSE, MAE i SMAPE na istom skupu podataka [22]

5. Pronalazak spline rijetke baze

U prethodnim poglavljima opisana je teorija spline funkcija i sažimajućeg očitavanja. Kako su spline funkcije same po sebi interpolacijske funkcije, tj. od manjeg broja uzoraka generiraju više uzoraka, formule kojima se izvodi taj postupak su upravo formule koje mogu poslužiti kao matrica rijetke baze za sažimajuće očitavanje na način da su s jedne strane početni podaci koje treba interpolirati, tj. rijetki uzorci, a s druge strane interpolirani podaci, tj. rekonstruirani uzorci. U nastavku će biti opisan način na koji se dolazi do matrice rijetke baze jer SeDuMi rekonstrukcijski algoritam prima matrični zapis. Međutim, algoritam l1-MAGIC može primiti funkciju umjesto matrice, ali provođenje tog postupka je potrebno doraditi.

5.1 Algoritam za generiranje spline rijetke baze

U potpoglavlju 2.3 opisani su uvjeti za spline interpolaciju čijim rješavanjem dolazimo do matričnog zapisa izračuna jednog od koeficijenata prirodnog kubnog splinea. Ostali koeficijenti se izračunavaju iterativnim formulama.

Algoritam kojim bi se došlo do matrice koja veže ulazne i interpolirane podatke kao rezultat svakog koraka treba imati matrični zapis i izgleda:

- Sažimanjem uvjeta za interpolaciju dolazimo do matričnog zapisa interpolacijske zadaće i tridijagonalne matrice
- Potrebno je invertirati tridijagonalnu matricu i doći do jednog od spline koeficijenata
- Iterativne formule kojima se dolazi do preostalih koeficijenata potrebno je zapisati u matričnom obliku
- Sve koeficijente potrebno je složiti kao stupce nove matrice
- Dobivenu matricu potrebno je pomnožiti s težinskom matricom koja predstavlja evaluaciju polinoma te tako dobivamo interpolirane uzorke
- Sve korištene matrice i operacije potrebno je kaskadno povezati kako bi se dobila jedinstvena matrica koja predstavlja rijetku bazu

5.2 Tridijagonalni matrični zapis i invertiranje

Prisjetimo se matričnog zapisa interpolacijske zadaće (14) iz poglavlja 2.3. Broj podskupova $n-1$, gdje n predstavlja dužinu ulaznog niza podataka, nad kojima interpoliramo ulazne podatke određuje broj redaka tridijagonalne matrice A koja je dimenzija $(n-2) \times (n-2)$, ostale koeficijente određuju rubni uvjeti [9].

$$h_i = x_{i+1} - x_i, \quad \gamma_0 = \gamma_n = 0, \quad v_i = \frac{6}{h_i}(y_{i+1} - y_i) - \frac{6}{h_{i-1}}(y_i - y_{i-1})$$

$$A\gamma = v$$

$$\begin{bmatrix} 2(h_1 + h_2) & h_2 & 0 & \dots & 0 \\ h_2 & 2(h_2 + h_3) & h_3 & \ddots & \vdots \\ 0 & h_3 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & h_{n-2} \\ 0 & \dots & 0 & h_{n-2} & 2(h_{n-2} + h_{n-1}) \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_{n-2} \\ \gamma_{n-1} \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n-2} \\ v_{n-1} \end{bmatrix}$$

Tridijagonalna matrica A ovisi samo o vrijednostima h_i koje definira razlika između x koordinata ulaznih uzoraka, tj. duljina podskupa. Ako su ulazni uzorci uniformni i udaljeni za 1, matrica A je konstantna i izgleda [23]:

$$A = \begin{bmatrix} 4 & 1 & 0 & \dots & 0 \\ 1 & 4 & \ddots & 0 & 0 \\ 0 & \ddots & \ddots & 1 & \vdots \\ \vdots & 0 & \ddots & 4 & 1 \\ 0 & \ddots & 0 & 1 & 4 \end{bmatrix}$$

Generiranje trodijagonalne matrice u Matlabu za bilo koji ulazni skup (x,y) izgleda:

```
n = length(x)

h = x(2:n) - x(1:n-1);
d = (y(2:n) - y(1:n-1))./h;

a=2*(h(1:(n-2))+h(2:(n-1)))
b=h(2:(n-2));

u=6*(d(2:n-1)-d(1:n-2))
A = full(gallery('tridiag',b,a,b))
```

Kod 4 Generiranje tridijagonalne matrice

Nakon što smo dobili tridijagonalnu matricu, za izračun koeficijenta koji se u polinomu nalazi uz član x^2 potrebno je invertirati tu matricu. Svojstvo tridijagonalne matrice je da se može rastaviti na gornju i donju trokutastu te zatim invertirati algoritmima unaprijedne supstitucije za donju trokutastu matricu i unatraske supstitucije za gornju trokutastu matricu. Supstitucije su dobile takav naziv jer se prvo izračuna početni korak (x_1 za donju trokutastu ili x_n za gornju trokutastu matricu) te se supstituira u sljedeći korak dok svi podaci nisu izračunati. Više detalja može se pronaći u literaturi [24].

```
function X = tridiag(L,D,U,B)

n = length(B);
for k = 2:n,
    mult = L(k-1)/D(k-1);
    D(k) = D(k) - mult*U(k-1);
    B(k) = B(k) - mult*B(k-1); %forward substitution
end
X(n) = B(n)/D(n);
for k = (n-1):-1:1,
    X(k) = (B(k) - U(k)*X(k+1))/D(k); %backward
subtitution
end

m=tridiag(b,a,b,u)
m = [ 0; m'; 0]
```

Kod 5 Invertiranje trodijagonalne matrice

```

s0=y'
s1 = (d - h.*(2*(m(1:end-1))' + (m(2:end))')/6)'
s2=m/2
s3 = (((m(2:end))'-(m(1:end-1)))'./(6*h))'
s3=[s3;0];
s1=[s1;0];
matrix=[s3';s2';s1';s0'];

```

Kod 6 Matrični zapis spline koeficijenata

Sljedeći korak je izračun preostalih spline koeficijenata. Za to je potrebno zapisati iterativne formule (15) u matričnom obliku, što se svodi na jednostavne operacije nad retcima vektora već izračunatog spline koeficijenta. Međutim, problem predstavlja sjedinjavanje svih spline koeficijenata u jedinstvenu matricu koristeći matrični operator, na čemu je potreban daljnji rad. Zasad je ta matrica dobivena jednostavnim operacijama.

5.3 Evaluacija polinoma

Nakon što smo dobili matricu spline koeficijenata koja opisuje sve polinomijalne funkcije na zadanim podskupovima, potrebno je izračunati, tj. evaluirati te polinome u novim točkama. Izabrano je da se to napravi na način ponovnog otipkavanja s većom frekvencijom, tj. da se interpoliraju novi podaci s odabranom frekvencijom.

To se može učiniti na način da se matrica spline koeficijenata množi s matricom težinskih koeficijenata koji odgovaraju potrebnim potencijama odabrane frekvencije novog uzorkovanja, tj. Δx^i , gdje i predstavlja broj retka. Primjer težinske matrice za frekvenciju 0.25 izgleda ovako:

$$A = \begin{bmatrix} 0 & 0.25^3 & 0.5^3 & 0.75^3 \\ 0 & 0.25^2 & 0.5^2 & 0.75^2 \\ 0 & 0.25 & 0.5 & 0.75 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Ako je ulazni niz podataka bio uniforman, onda će se prolaziti jednak broj puta kroz svaki stupac matrice težinskih koeficijenata. Slučaj kad je ulazni niz neuniforman je nešto kompliciraniji jer je potrebno izračunati nekoliko nepotrebnih koeficijenata na užim podskupovima kako bi uzorci iz većih podskupova mogli biti interpolirani. Zato se ti nepotrebni uzorci naknadno nuliraju.

```
rate=0.5;
xs=[x(1):rate:x(n)];

for i=1:length(x)-1,
    h(i)=x(i+1)-x(i)
end

broj_stupaca = max(h)/rate;
transform1=[0 0 0 1];

for i=1:broj_stupaca-1,
    transform1 = [transform1, (i*rate)^3
(i*rate)^2 (i*rate) 1]
end

transform_new = reshape(transform1,[4,broj_stupaca]);
ys = matrix * transform_new;

for i=1:length(x)-1,
    if h(i) < max(h)
        L=(max(h)-h(i))/rate
        ys(i,L+1:length(ys(1,:)))=0
    end
end

ysa = reshape(ys.',1,[]);
ysa(ysa==0)=[]

ysa=ysa(1:length(xs))

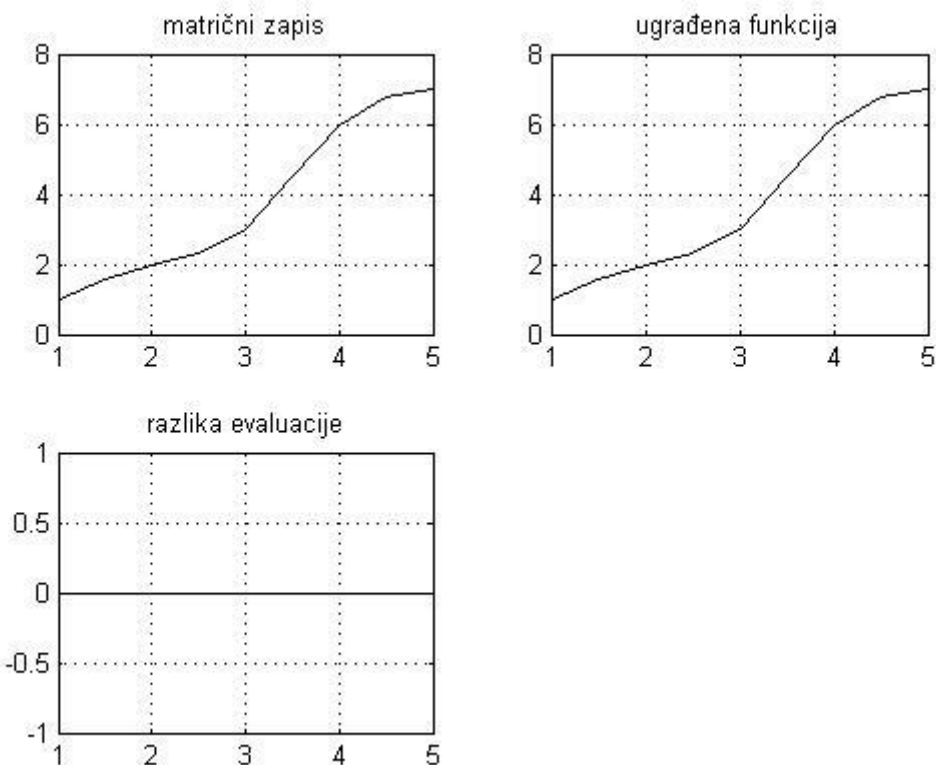
figure
ax1 = subplot(2,2,1)
plot(xs,ysa,'k-')
title('matrični zapis')
```

Kod 7 Izračun težinske matrice i provjera interpoliranih podataka

Da bismo provjerili točnost dosad korištenih matrica, usporedit ćemo tako dobivene vrijednosti polinoma s ugrađenom funkcijom *csape* u Matlabu ($x = [1 \ 2 \ 3 \ 4 \ 5]$, $y = [1 \ 2 \ 3 \ 6 \ 7]$, frekvencija 0.5).

```
pp = csape(x,y,'second');  
ax2 = subplot(2,2,2)  
plot(xs, fnval(pp,xs),'k-')  
  
title('ugrađena funkcija')  
  
ax3 = subplot(2,2,3)  
plot(xs, (fnval(pp,xs)-ysa),'k-')  
title('razlika evaluacije')  
  
grid(ax1,'on')  
grid(ax2,'on')  
grid(ax3,'on')
```

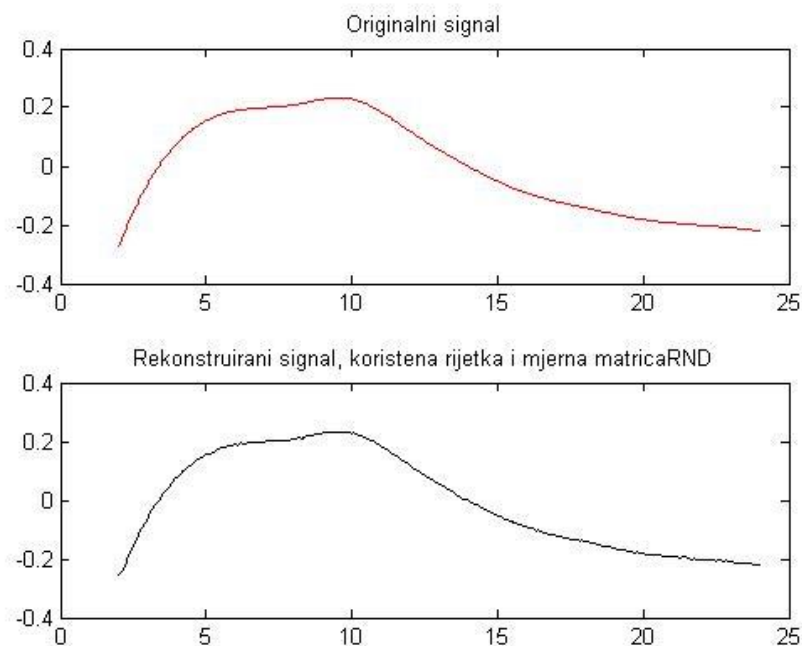
Kod 8 Usporedba ugrađene funkcije i matričnog zapisa



Slika 18 Usporedba ugrađene funkcije i matričnog zapisa

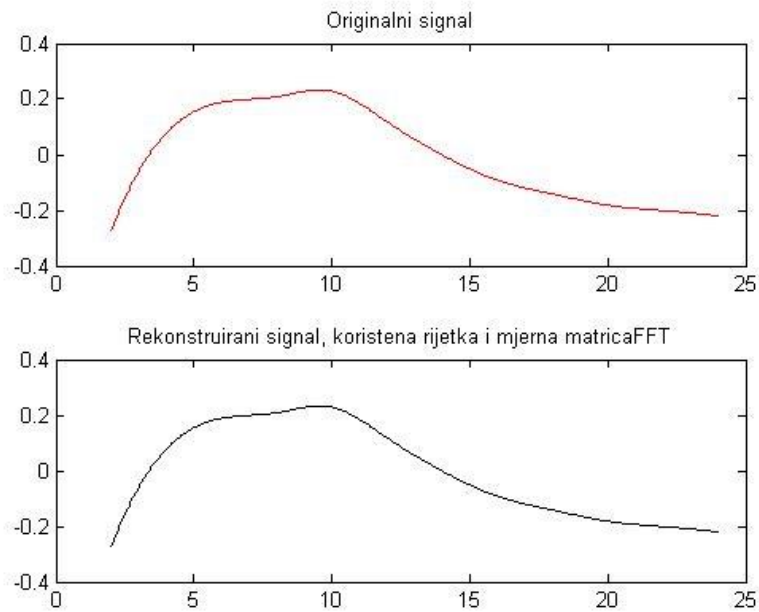
6. Provedba sažimajućeg očitavanja na spline funkcijama

Iako ne koristimo spline rijetku bazu, moguće je provesti tehniku sažimajućeg očitavanja na spline funkcijama koristeći druge rijetke baze kao što je DCT baza. Za mjerne matrice korištene su slučajne matrice s Gaussovom razdiobom, FFT koeficijentima i DWT koeficijentima. Kod korištenja za provedbu ove tehnike može se pronaći u literaturi [26]. Kao mjera odstupanja rekonstruirane i originalne karakteristike koristi se metoda srednje kvadratne pogreške. Referentni signal čini spline funkcija s 441 koeficijentom, od kojih se 50% odbacuje da bi se dobio rijetki signal. Prikazan je i primjer s odbacivanjem 75% koeficijenata.

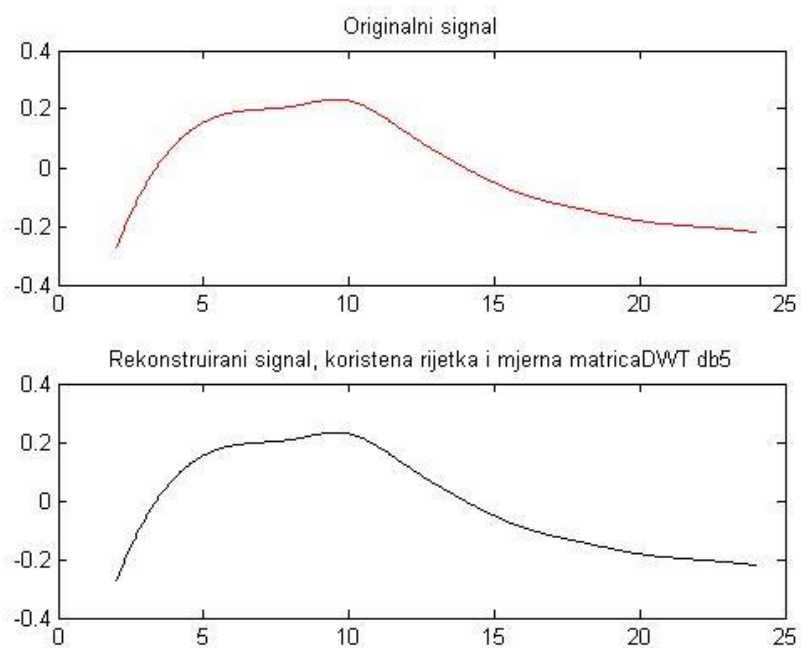


Slika 19 CS tehnika koristeći slučajnu mjernu matricu

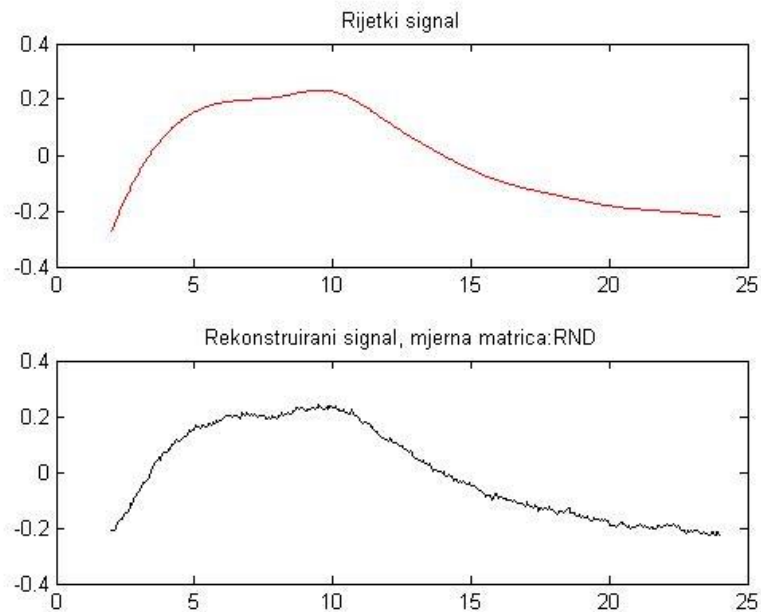
Pogreška na primjeru sa slike 19 iznosi -41 dB, na slici 20 iznosi -192 dB, na slici 21 -190 dB te na slici 22 -25 dB.



Slika 20 CS tehnika koristeći slučajnu mjernu matricu s FFT koeficijentima

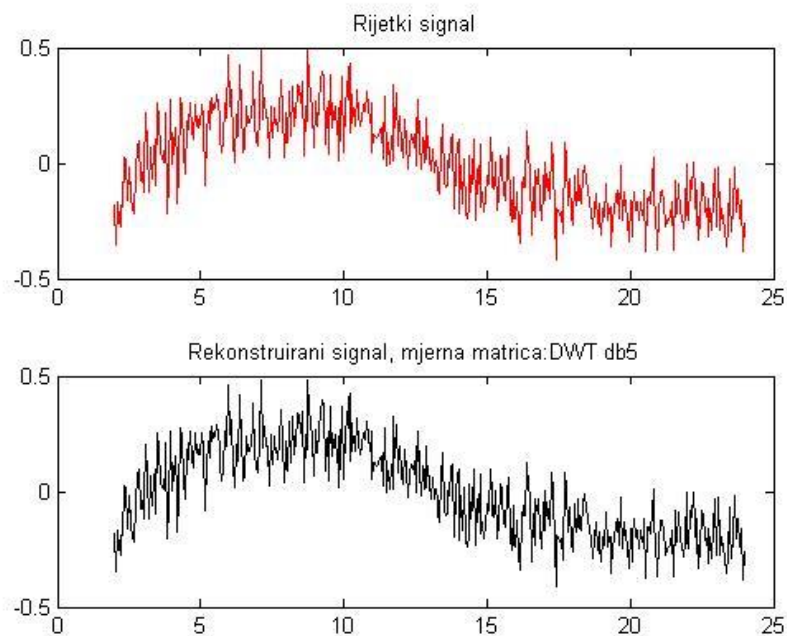


Slika 21 CS tehnika koristeći slučajnu mjernu matricu s DWT koeficijentima



Slika 22 Odbacivanje 75% koeficijenata

Ako na isti referentni signal dodamo bijeli Gaussov šum tako da omjer $\frac{\|Ax\|_{\ell_2}}{\|z\|_{\ell_2}}$ iznosi približno 5 te primijenimo algoritam (27) za redukciju šuma, rezultat je odstupanje od -6 dB u okruženju od 2.7ε . Više rezultata u literaturi [27].



Slika 23 CS tehnika u šumovitom okruženju

7. Zaključak

Spline funkcije imaju široku primjenu u današnjoj matematici, pogotovo kao tehnika interpolacije. Zbog svog interpolativnog svojstva da ograničen skup uzoraka prošire na veći skup, pogodne su za primjenu sažimajućeg očitavanja. Sažimajuće očitavanje je tehnika obrade signala koja omogućuje rekonstrukciju signala iz mnogo manje uzoraka nego što to zahtijeva Nyquistov teorem. Počiva na dva svojstva, da je signal u nekoj bazi rijedak, tj. nema mnogo ne-nul uzoraka u toj bazi te inkohherenciji između rijetke i mjerne matrice koja utječe na broj uzoraka potreban za rekonstrukciju.

Da bi se rekonstrukcija uspješno provela, potrebno je pažljivo odabrati rijetku i mjernu bazu koje zajedno čine osjetilnu matricu. U slučaju spline funkcija, rijetka matrica se lako dobiva zbog same njihove prirode. Međutim, proces dobivanja potpunog matričnog zapisa nije jednostavan te su napravljeni svi koraci osim spajanja vektora spline koeficijenata u jedinstvenu matricu koristeći matrični operator. Zbog toga što algoritmi optimizacije ℓ_1 norme zahtijevaju matrični zapis ili cjelovitu jednadžbu kao ulazni parametar, sama tehnika sažimajućeg očitavanja nije provedena sa spline rijetkom bazom. Pretpostavlja se da bi tehnika trebala biti uspješno provedena s jako malim brojem uzoraka, čak i ako se oni nalaze u samo nekim podintervalima spline funkcija zbog uvjeta interpolacije koji spaja rubove pojedinih podintervala pa su svi uzorci međusobno povezani.

CS tehnika je provedena na slučaju DCT rijetke baze i nekoliko primjera slučajnih mjernih matrica sa SeDuMi rekonstrukcijskim algoritmom, u prisustvu i bez dodanog šuma. Rezultati su zadovoljavajući, CS tehnika se pokazala otpornom i učinkovitom pri rekonstrukciji smanjenog skupa podataka.

Ovaj rad je podložen nadopunama i budućim saznanjima, primjerice dovršetku konstrukcije spline rijetke baze ili provedbi I1-MAGIC algoritma, koji bi mogli doprinijeti utvrđivanju korisnosti i važnosti primjene sažimajućeg očitavanja na spline funkcije.

Literatura

- [1] Milišić, Josipa Pina, Žgaljić Keko, Ana: Uvod u numeričku matematiku za inženjere, 1.izdanje, Element, 2013.
- [2] Izenda, Inc., From Ships to Data: a History of the Spline, 16.09.2014., <https://www.izenda.com/blog/from-ships-to-data-a-history-of-the-spline/>, 20.06.2016.
- [3] Fakultet strojarstva i brodogradnje, Numerička matematika, 21.12. 2008., https://www.fsb.unizg.hr/mat-4/OldWeb/4_p.pdf , 20.06.2016.
- [4] Elk, Jim, Satin Doll and Waterwitch, <http://www.elkspar.com/satin-doll.html>, 20.06.2016.
- [5] The Boat Design and Boat building Site, 30.05.2007., <http://www.boatdesign.net/forums/education/splineweights-com-make-your-own-spline-weights-10149.html>, 20.06.2016.
- [6] TrueArt, EasySpline, <http://www2.trueart.pl/?URIType=Directory&URI=Products/Plug-Ins/EasySpline>, 20.06.2016.
- [7] Candès, Emmanuel J., Wakin, Michael B., An introduction to compressive Sampling, Signal Processing Magazine, IEEE, Volume 25, Issue 2, March 2008.
- [8] Uecker, Martin, 2015., Parallel Imaging as Approximation in a Reproducing Kernel Hilbert Space, <http://wwwuser.gwdg.de/~muecker/noir.html> ,20.06.2016.
- [9] Fomel, Sergey, 2002., Interpolation: Spline Interpolation, <http://sepwww.stanford.edu/sep/sergey/128A/answers6.pdf>, 20.06.2016.
- [10] Mahaffy, Joe, Interpolation and Polynomial Approximation, 2010., <http://www.rohan.sdsu.edu/~jmahaffy/courses/s10/math541/lectures/pdf/week06/lecture.pdf>, 20.06.2016.
- [11] R. Tibshirani, "Regression shrinkage and selection via the lasso," J. Roy. Stat. Soc. Ser. B, vol. 58, no. 1, pp. 267-288, 1996.

- [12] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.
- [13] Kadambi, Achuta, Compressed Sensing: Introduction and Apps, <http://www.slideshare.net/cameraculture/compressed-sensing-achuta-kadambi>, 20.06.2016.
- [14] Antoine Liutkus, David Martina, Sébastien Popoff, Gilles Chardon, Ori Katz, Geoffroy Lerosey, Sylvain Gigan, Laurent Daudet & Igor Carron, Imaging With Nature: Compressive Imaging Using a Multiply Scattering Medium, 09.07.2014., <http://www.nature.com/articles/srep05552>, 20.06.2016.
- [15] Thu L. N. Nguyen and Yoan Shin, 05.11.2013., Deterministic Sensing Matrices in Compressive Sensing: A Survey, <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3836388/>, 20.06.2016.
- [16] Romberg, Justin, 09.01.2012., An Overview of Sparsity with Applications to Compression, Restoration, and Inverse Problems, <http://www.ens-lyon.fr/DI/wp-content/uploads/2011/10/lecture-1-5-sparsity-overview.pdf>, 20.06.2016.
- [17] Sturm, Jos F., 01.08.1998., Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.49.6954&rep=rep1&type=pdf>, 20.06.2016.
- [18] Yang, Allen Y., Ganesh, Arvind, Zhou, Zihan, Shankar Sastry, S., Ma, Yi, A review of fast L1-minimization algorithms for robust face recognition, <http://ecovision.mit.edu/~sai/12S990/Yang10-SIAM.pdf>, 20.06.2016.
- [19] Y. Ye, M.J. Todd, and S. Mizuno. An $O(\sqrt{nL})$ -iteration homogeneous and self-dual linear programming algorithm. *Mathematics of Operations Research*, 19:53-67, 1994.
- [20] Wikipedia, 17.01.2016., Mean squared error, https://en.wikipedia.org/wiki/Mean_squared_error, 20.06.2016.
- [21] Wikipedia, 02.06.2016., Mean absolute error, https://en.wikipedia.org/wiki/Mean_absolute_error, 20.06.2016.

- [22] V. Puig, C. Ocampo-Martinez, R.R. Negenborn, 01.2015., Model Predictive Control for Combined Water Supply and Navigability/Sustainability in River Systems, https://www.researchgate.net/figure/266477965_fig3_Figure-3-Error-measurements-MSE-MAE-and-SMAPE
- [23] Brandao, Mark, 04.03.2014., Natural and Clamped Cubic Splines, <http://www.slideshare.net/mbrandao1/natural-and-clamped-cubic-splines>, 20.06.2016.
- [24] Wikipedia, 08.06.2016., Triangular matrix, Forward and back substitution, https://en.wikipedia.org/wiki/Triangular_matrix#Forward_and_back_substitution, 20.06.2016.
- [25] Wikipedia, 09.02.2016., Symmetric mean absolute percentage error, https://en.wikipedia.org/wiki/Symmetric_mean_absolute_percentage_error, 20.06.2016.
- [26] Ankica Gogić, Nikolina Očić, Janja Paliska, Antonio Zemunik, akademska godina 2014./2015., Primjena compressive sensing metode na rekonstrukciju slika, http://www.fer.unizg.hr/predmet/nmdos_a/samostalni_rad#, 20.06.2016
- [27] Emmanuel Candes, Justin Romberg, Terence Tao, veljača 2005., Stable Signal Recovery from Incomplete and Inaccurate Measurements, <http://statweb.stanford.edu/~candes/papers/StableRecovery.pdf>, 20.06.2016.
- [28] Emmanuel Candès, Justin Romberg, listopad 2005., l1-magic : Recovery of Sparse Signals via Convex Programming, <http://statweb.stanford.edu/~candes/l1magic/downloads/l1magic.pdf>, 20.06.2016.
- [29] CVX Research, Inc., lipanj 2015., CVX: Matlab Software for Disciplined Convex Programming, <http://cvxr.com/cvx/>, 20.06.2016.

Reprezentacija spline funkcija korištenjem sažimajućeg očitavanja

Sažetak

Spline funkcije su polinomijalne funkcije po intervalima koje se koriste za opis i manipulaciju krivuljama te se često upotrebljavaju, primjerice za interpolaciju. Upravo zbog svog svojstva da iz ograničenog broja uzoraka generiraju kontinuirane krivulje, spline funkcije prikladne su za korištenje tehnike sažimajućeg očitavanja. Tehnika sažimajućeg očitavanja je efikasan i jednostavan protokol akvizicije koji na, o signalu neovisan način, uzorkuje signal niskom frekvencijom i kasnije koristi moć računanja za rekonstrukciju iz nečeg što se čini kao nepotpun skup podataka. Glavni principi na kojima se temelji su rijetkost signala i inkoherencija te je stoga potrebno pažljivo odabrati mjernu i rijetku bazu koje zajedno čine osjetilnu matricu. U ovom radu pokazan je postupak primjene sažimajućeg očitavanja na spline funkcije te donesen zaključak o budućim primjenama spline rijetke baze.

Ključne riječi: spline, interpolacija, sažimajuće očitavanje, rijetkost, osjetilna matrica, šum, RIP, Matlab, SeDuMi, I1-MAGIC.

Compressed sensing representation of spline functions

Abstract

Spline functions are polynomial functions on intervals that are used for description and manipulation of curves and are often used, for example for interpolation. Because of their property that they generate continuous curves from a limited number of samples, spline functions are convenient for using compressive sensing. Compressive sensing is an efficient and simple acquisition protocol which, independent of the signal, samples a signal with low frequency and then uses computational power to reconstruct the incomplete set of data. Main principles on which it is based are signal sparsity and incoherence, so it is required to carefully choose measurement and sparse basis, which together make sensing matrix. In this thesis the process of application of compressive sensing on spline functions has been shown and future applications of spline functions are mentioned.

Keywords: spline, interpolation, compressive sensing, sparsity, sensing matrix, noise, RIP, Matlab, SeDuMi, I1-MAGIC.