

Numerička matematika, 9. predavanje

Pina Milišić, Ana Žgaljić Keko

Skripta

Sadržaj

1	Uvod	5
1.1	Greške	5
1.2	Prikaz realnih brojeva u računalu	5
1.2.1	Preciznost, strojni epsilon i greška zaokruživanja	5
1.2.2	Propagiranje grešaka kroz aritmetičke operacije	5
1.3	Stvarne katastrofe uzrokovane greškom	5
2	Linearni sustavi	7
2.1	Gaussove eliminacije i LU faktORIZACIJA	7
2.2	Pivotiranje	7
2.3	Neki posebni tipovi linearnih sustava	7
2.3.1	Simetrične pozitivno definitne matrice	7
2.3.2	Rješavanje tridijagonalnog sustava. Thomasov algoritam	7
2.4	Analiza greške rješenja	7
3	Interpolacija i aproksimacija funkcija	9
3.1	Polinomijalna interpolacija	9
3.1.1	Primjena: numerička integracija	9
3.2	Interpolacija splajnovima	9
3.2.1	Linearni splajn	9
3.2.2	Kubični splajn	9
3.3	Trigonometrijska interpolacija	10
3.3.1	Kompleksne matrice i kompleksni vektori	10
3.3.2	Diskretna Fourierova transformacija	13
3.3.3	Brza Fourierova transformacija (FFT)	24
	Bibliography	29
4	Dodatak	31
4.1	Neki važni pojmovi iz linearne algebre i matematičke analize	31
4.1.1	”Veliko” O i ”malo” o notacija	31
4.1.2	Vektorski prostori	32
4.1.3	Neprekidnost, ograničenost, integrabilnost	33

4.1.4	Teoremi srednje vrijednosti	33
-------	---------------------------------------	----

1

Uvod

1.1 Greške

1.2 Prikaz realnih brojeva u računalu

1.2.1 Preciznost, strojni epsilon i greška zaokruživanja

1.2.2 Propagiranje grešaka kroz aritmetičke operacije

1.3 Stvarne katastrofe uzrokovane greškom

2

Linearni sustavi

2.1 Gaussove eliminacije i LU faktORIZACIJA

2.2 PIVOTIRANJE

2.3 NEKI POSEBNI TIPOLI LINEARNIH SUSTAVA

2.3.1 Simetrične pozitivno definitne matrice

FaktORIZACIJA CHOLESKOG

2.3.2 Rješavanje tridijagonalnog sustava. THOMASOV ALGORITAM

2.4 ANALIZA GREŠKE RJEŠENJA

3

Interpolacija i aproksimacija funkcija

3.1 Polinomijalna interpolacija

Greška interpolacije.

3.1.1 Primjena: numerička integracija

Newton-Cotesove formule.

Ocjene greške formula numeričke integracije

Kompozitne formule

3.2 Interpolacija splajnovima

3.2.1 Linearni splajn

3.2.2 Kubični splajn

Konstrukcija kubičnog splajna.

Minimalno svojstvo prirodnog kubičnog splajna.

3.3 Trigonometrijska interpolacija

Mnoge fizikalne pojave, poput rotacija, oscilacija i valova mogu se opisati se pomoću periodičnih funkcija pa je stoga za interpolaciju takvih veličina prirodno koristiti trigonometrijske funkcije.

U ovom poglavlju bavit ćemo se trigonometrijskom interpolacijskom zadaćom koja se svodi na rješavanje linearnog sustava sa kompleksnom matricom posebne strukture. Spomenutu matricu nazivamo Fourierova matrica. Kao što ćemo vidjeti, posebna struktura Fourierove matrice omogućava da se rješavanje dobivenog linearnog sustava svede na množenje vektora desne strane kompleksnom matricom dimenzije $n \times n$, za što je potrebno $O(n^2)$ aritmetičkih operacija. No, ovdje tek počinje najzanimljiviji dio. Naime, pokazuje se da je moguće konstruirati algoritam za rješavanje pripadnog linearnog sustava koji zahtijeva svega $O(n \log_2 n)$ računskih operacija, pri čemu je n dimenzija sustava. Ovaj algoritam naziva se brza Fourierova transformacija (eng. *Fast Fourier Transformation (FFT)*), a zasniva se na pametnom rastavu Fourierove matrice.

Kako smo dosad radili uglavnom u prostoru \mathbb{R}^n sa realnim vektorima i realnim matricama, uvest ćemo najprije neke osnovne pojmove vezane uz prostor \mathbb{C}^n . U ovom poglavlju slovo i označava kompleksnu jedinicu i neće se koristiti u notaciji indeksa.

3.3.1 Kompleksne matrice i kompleksni vektori

Uzmimo za početak kompleksni vektor $z \in \mathbb{C}^2$

$$z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

gdje su $z_j \in \mathbb{C}$, $j = 1, 2$ i razmislimo kako bismo definirali njegovu duljinu. U prostoru \mathbb{R}^2 duljinu (modul) vektora x računali smo pomoću skalarnog produkta na sljedeći način:

$$\|x\|^2 = (x, x) = x^T x.$$

Pogledajmo je li gornja definicija skalarnog produkta odgovarajuća za prostor \mathbb{C}^2 . Uzmimo na primjer kompleksni vektor $z = \begin{bmatrix} z_1 & z_2 \end{bmatrix}^T = \begin{bmatrix} 1 & i \end{bmatrix}^T$, gdje smo sa i označili kompleksnu jedinicu i izračunajmo $z^T z$. Imamo

$$\begin{bmatrix} 1 & i \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = 1 + i^2 = 0.$$

Ovaj primjer pokazuje da definicija skalarnog produkta $z^T z$ u prostoru \mathbb{C}^2 nije dobra. Prisjetimo se da za kompleksni broj $z_1 \in \mathbb{C}$ vrijedi $|z_1|^2 = \overline{z_1} z_1$ gdje je $\overline{z_1}$ broj kompleksno-konjugiran broju z_1 . Dakle, da bismo dobili dobru definiciju

skalarnog produkta u \mathbb{C}^2 moramo ne samo transponirati vektor z nego ga uz to još i kompleksno-konjugirati. Za naš primjer sada dobivamo

$$z^T \bar{z} = \begin{bmatrix} 1 & i \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} = 1 - i^2 = 2.$$

Dakle, euklidska norma vektora $z \in \mathbb{C}^n$ dobro je definirana izrazom

$$\|z\|_2^2 = (z, z) = z^T \bar{z} = \begin{bmatrix} z_1 & z_2 & \cdots & z_n \end{bmatrix} \begin{bmatrix} \bar{z}_1 \\ \bar{z}_2 \\ \vdots \\ \bar{z}_n \end{bmatrix} = |z_1|^2 + |z_2|^2 + \cdots + |z_n|^2.$$

Skalarni produkt vektora $x, y \in \mathbb{C}^n$ je općenito kompleksan broj, a definira se kao

$$(x, y) = \sum_{j=1}^n x_j \bar{y}_j. \quad (3.1)$$

Dakle, osim što, kao u realnom slučaju, transponiramo, ovdje još i odgovarajuće elemente trebamo i kompleksno konjugirati.

Napomena 3.1 Radi usklađenosti sa matričnim računom u numeričkoj linearnoj algebri često se koristi drugačija definicija skalarnog produkta u \mathbb{C}^n dana izrazom

$$(x, y) = \bar{x}^T y,$$

te se koristi oznaka $x^* = \bar{x}^T$.

U realnom prostoru \mathbb{R}^n za kvadratne matrice $A \in \mathbb{R}^{n \times n}$ vrijedi

$$\forall x, y \in \mathbb{R}^n \quad (Ax, y) = (x, A^T y) \quad (3.2)$$

gdje je A^T transponirana matrica. Kako smo u prostoru \mathbb{C}^n promijenili definiciju skalarnog produkta u odnosu na prostor \mathbb{R}^n , vidimo da za $A \in \mathbb{C}^{n \times n}$ neće vrijediti svojstvo (3.2). Da bismo u kompleksnom slučaju zadržali izraz (3.2) potrebno je pored transponiranja matrice kompleksno-konjugirati njene elemente. Na taj način dobivamo da za kvadratne matrice $A \in \mathbb{C}^{n \times n}$ vrijedi

$$\forall x, y \in \mathbb{C}^n \quad (Ax, y) = (x, A^* y) \quad (3.3)$$

gdje je $A^* = \bar{A}^T$ hermitski adjungirana (kompleksno transponirana) matrica. Sada možemo definirati pojam kompleksno-simetrične, odnosno hermitske matrice. Kažemo da je matrica $A \in \mathbb{C}^{n \times n}$ hermitska ako vrijedi

$$A^* = A. \quad (3.4)$$

Pokazuje se da vrijedi sljedeća propozicija.

Propozicija 3.1 Za sve matrice $A, B \in \mathbb{C}^{n \times n}$ i za svako $\alpha \in \mathbb{C}$ vrijede sljedeće tvrdnje:

- (i) $(A + B)^* = A^* + B^*$
- (ii) $(AB)^* = B^*A^*$
- (iii) $(\alpha A)^* = \bar{\alpha}A^*$.

Pokazuje se da hermitske matrice imaju realne svojstvene vrijednosti. Nadalje, različitim svojstvenim vrijednostima pripadaju međusobno ortogonalni svojstveni vektori. Dokazi ovdje spomenutih rezultata mogu se naći na primjer u [21].

Primjer 3.1 Zadana je matrica

$$A = \begin{bmatrix} 2 & 3 - 3i \\ 3 + 3i & 5 \end{bmatrix}.$$

Pokažimo da je zadana matrica hermitska, te odredimo njene svojstvene vrijednosti i svojstvene vektore.

Pripadne svojstvene vrijednosti označimo sa λ_1 i λ_2 , a njima odgovarajuće svojstvene vektore x , odnosno y . Standardni račun daje svojstvene vrijednosti $\lambda_1 = 8$, $\lambda_2 = -1$ te svojstvene vektore

$$x = \begin{bmatrix} 1 \\ 1 + i \end{bmatrix}, \quad y = \begin{bmatrix} 1 - i \\ -1 \end{bmatrix}.$$

Očito, svojstveni vektori x i y su međusobno ortogonalni, to jest vrijedi

$$(x, y) = x^T \bar{y} = \begin{bmatrix} 1 & 1 - i \end{bmatrix} \begin{bmatrix} 1 - i \\ -1 \end{bmatrix} = 0.$$

Kažemo da su zadani vektori $q_j \in \mathbb{C}^n$, $j = 1, \dots, n$ međusobno ortogonalni (ortonormirani) ako vrijedi

$$(q_k, q_j) = \sum_{l=1}^n (q_k)_l (\bar{q_j})_l = \begin{cases} 0, & k \neq j, \\ 1, & k = j \end{cases} \quad (3.5)$$

Formirajmo sada matricu Q tako da vektori q_j , $j = 1, \dots, n$ čine njene stupce. Dobivamo da vrijedi

$$Q^*Q = I. \quad (3.6)$$

Matrice za koje vrijedi izraz (3.6) zovu se unitarne matrice. Lako se vidi da za unitarne matrice $Q \in \mathbb{C}^{n \times n}$ vrijedi $Q^*Q = QQ^* = I$.

3.3.2 Diskretna Fourierova transformacija

Podsjetimo se najprije nekih osnovnih pojmova i formula iz kompleksne analize. Veza trigonometrijskih funkcija $\sin x$ i $\cos x$ sa kompleksnim brojevima dana je Eulerovim identitetom koji glasi

$$e^{ix} = \cos x + i \sin x, \quad (3.7)$$

gdje je $i = \sqrt{-1}$ imaginarna jedinica. Nadalje, vrijede sljedeći identiteti

$$\cos x = \frac{e^{ix} + e^{-ix}}{2}, \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i}. \quad (3.8)$$

Prema osnovnom teoremu algebre jednadžba

$$z^N - 1 = 0$$

ima N rješenja u skupu \mathbb{C} . Očito je $1 = \cos 0 + i \sin 0$. Nadalje, Moivreova formula daje

$$(\sqrt[N]{z})_k = \cos\left(\frac{0 + 2k\pi}{N}\right) + i \sin\left(\frac{0 + 2k\pi}{N}\right), \quad k = 0, \dots, N-1, \quad (3.9)$$

odnosno, zapisano pomoću Eulerovog identiteta dobivamo

$$(\sqrt[N]{z})_k = e^{i\frac{2\pi k}{N}}, \quad k = 0, \dots, N-1. \quad (3.10)$$

Primijetimo da $(\sqrt[N]{z})_k$, $k = 0, \dots, N-1$ u kompleksnoj ravnini čine vrhove pravilnog n -terokuta. Sa w označimo pripadni korijen jedinice kojeg dobivamo za $k = 1$, odnosno neka je

$$w = e^{i\frac{2\pi}{N}}. \quad (3.11)$$

Trigonometrijska interpolacijska zadaća. Zadanu 2π -periodičku funkciju f želimo interpolirati trigonometrijskim polinomom

$$\psi(x) = a_0 + \sum_{k=1}^M [a_k \cos(kx) + b_k \sin(kx)], \quad |a_M| + |b_M| \neq 0 \quad (3.12)$$

kroz $N = 2M + 1$ interpolacijska čvora zadane mreže na segmentu $[0, 2\pi]$. Dakle, da bismo odredili trigonometrijski interpolacijski polinom ψ treba odrediti N nepoznatih koeficijenata.

Napomena 3.2 Primijetimo najprije da transformacija

$$\xi = \frac{b-a}{2\pi}x + a$$

preslikava segment $[0, 2\pi]$ na $[a, b]$, odakle je jasno da je razmatranja dovoljno vršiti na segmentu $[0, 2\pi]$.

Nadalje, uočimo da ovdje trigonometrijski interpolacijski problem promatramo za slučaj kada imamo neparan broj točaka mreže, budući iz oblika interpolacijskog polinoma (3.12) vidimo da treba odrediti neparan broj koeficijenata. U slučaju parnog broja interpolacijskih točaka obično se izraz (3.12) nadopuni još jednim članom. Na primjer, za $N = 2M$ traženi interpolacijski polinom je oblika

$$\psi(x) = a_0 + \sum_{k=1}^{M-1} [a_k \cos(kx) + b_k \sin(kx)] + a_M \cos(Mx).$$

Za detalje može se pogledati na primjer u [9].

Koristeći formule (3.8) direktno slijedi da interpolacijski polinom $\psi(x)$ možemo zapisati u obliku

$$\begin{aligned} \psi(x) &= a_0 + \sum_{k=1}^M \left[a_k \frac{e^{ikx} + e^{-ikx}}{2} + b_k \frac{e^{ikx} - e^{-ikx}}{2i} \right] \\ &= a_0 + \sum_{k=1}^M \left[\frac{a_k - ib_k}{2} e^{ikx} + \frac{a_k + ib_k}{2} e^{-ikx} \right] \\ &= c_0 + \sum_{k=1}^M [c_k e^{ikx} + c_{-k} e^{-ikx}] \\ &= \sum_{k=-M}^M c_k e^{ikx} \end{aligned} \tag{3.13}$$

pri čemu su veze između koeficijenata a_k , b_k sa koeficijentima c_k dane formulama

$$c_0 = a_0, \quad c_k = \frac{a_k - ib_k}{2}, \quad c_{-k} = \frac{a_k + ib_k}{2}, \quad 1 \leq k \leq M.$$

Neka je na segmentu $[0, 2\pi]$ zadana ekvidistantna mreža sa čvorovima

$$x_j = \frac{2\pi j}{N}, \quad j = 0, 1, \dots, N-1. \tag{3.14}$$

Uočimo da, zbog periodičnosti, točka $x_N = 2\pi$ ne daje novi interpolacijski uvjet. Jednolika raspodjela interpolacijskih čvorova omogućava nam da problem nalaženja koeficijenata svedemo na određivanje koeficijenata takozvanog **faznog interpolacijskog polinoma** $p(x)$ oblika

$$p(x) = \sum_{k=0}^{N-1} c_k e^{ikx}, \quad c_k \in \mathbb{C}, \tag{3.15}$$

pri čemu koeficijente c_j dobivamo iz uvjeta interpolacije

$$p(x_j) = y_j, \quad j = 0, 1, 2, \dots, N-1.$$

Naime, sljedeći identiteti

$$\begin{aligned} e^{-ikx_j} &= e^{-\frac{i2\pi kj}{N}} = e^{-\frac{i2\pi kj}{N}} \cdot e^{\frac{i2\pi Nj}{N}} \\ &= e^{\frac{i2\pi(N-k)j}{N}} = e^{i(N-k)x_j} \end{aligned}$$

nam omogućavaju da interpolacijski polinom (3.13) izračunat u točkama mreže možemo zapisati u obliku

$$\begin{aligned} \psi(x_j) &= c_0 + \sum_{k=1}^M \left[c_k e^{ikx_j} + c_{-k} e^{i(N-k)x_j} \right] \\ &= \sum_{k=0}^{N-1} c_k e^{ikx_j} = p(x_j), \quad j = 0, \dots, N-1, \end{aligned} \tag{3.16}$$

pri čemu smo uveli oznaku $c_{-k} = c_{N-k}$. Direktni račun daje vezu koeficijenata a_k i b_k sa koeficijentima c_k :

$$a_0 = c_0, \quad a_k = c_k + c_{N-k}, \quad b_k = i(c_k - c_{N-k}), \quad k = 1, \dots, M. \tag{3.17}$$

Trigonometrijsku interpolacijsku zadaća sada formuliramo u terminima faznog interpolacijskog polinoma (3.16) na mreži (3.14) na sljedeći način:

Za zadane čvorovi interpolacije x_j zadane izrazom (3.14), te za zadane pripadne vrijednosti $f(x_j) = y_j \in \mathbb{C}$, $j = 0, \dots, N-1$ treba odrediti koeficijente c_k , $k = 0, \dots, N-1$ tako da vrijede uvjeti interpolacije

$$\sum_{k=0}^N c_k e^{ikx_j} = y_j, \quad j = 0, \dots, N-1. \tag{3.18}$$

Jednadžbu (3.18) zapišimo u matričnoj notaciji.

Uvažimo li uvjete interpolacije (3.18) dolazimo do linearnog sustava

$$F_N c = y, \tag{3.19}$$

sa nepoznicama $c = [c_0 \ \dots \ c_{N-1}]^T$, pri čemu je desna strana sustava dana vektorom $y = [y_0 \ \dots \ y_{N-1}]^T$, dok je pripadna matrica sustava F_N oblika

$$F_N = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & w & w^2 & \dots & w^{N-1} \\ 1 & w^2 & w^4 & \dots & w^{2(N-1)} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & w^{N-1} & w^{2(N-1)} & \dots & w^{(N-1)^2} \end{bmatrix} \tag{3.20}$$

Primijetimo da je element matrice $(F_N)_{jk}$ koji se nalazi u j -tom retku i k -tom stupcu dan formulom

$$(F_N)_{jk} = w^{jk}, \quad j, k = 0, \dots, N-1,$$

pri čemu je w definirano izrazom (3.11), a prvi stupac, odnosno redak matrice F_N računamo kao nulti. Matrica F_N zove se **Fourierova matrica**. Dakle, rješavanje zadaće trigonometrijske interpolacije svodi se na rješavanje linearnog sustava sa Fourierovom matricom. Štoviše, struktura Fourierove matrice omogućava nam da vektor nepoznatih koeficijenata c odredimo eksplicitno iz zadanog vektora y pomoću jednostavne formule, umjesto da rješavamo pripadni linearni sustav. U tu svrhu promotrimo detaljnije što možemo reći o strukturi matrice F_N . Uzmimo na primjer $N = 4$ (vidi Napomenu 3.2 o (ne)parnosti čvorova interpolacije). U tom slučaju je $w = e^{i\frac{2\pi}{4}} = i$, a ostali korijeni jedinice su $i^2 = -1$, $i^3 = -i$, te $i^4 = 1$. Odatle slijedi da je matrica F_4 oblika

$$F_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix}. \quad (3.21)$$

Uočimo da su stupci matrice F_4 međusobno ortogonalni. Nadalje, stupci matrice $\tilde{F}_4 = \frac{1}{2}F_4$ su ortonormalni, odakle slijedi da je matrica \tilde{F}_4 unitarna, odnosno vrijedi

$$\tilde{F}_4^* \tilde{F}_4 = I,$$

pa je $\tilde{F}_4^{-1} = \tilde{F}_4^*$. Pokažimo da gornja razmatranja vrijede i za proizvoljni $N \in \mathbb{N}$. U tu svrhu stupce matrice F_N označimo sa q_j , $j = 0, \dots, N-1$. Tada imamo da vrijedi

$$\begin{aligned} (q_j, q_k) &= q_j^T \bar{q}_k = \sum_{l=0}^{N-1} f_{lj} \bar{f}_{lk} = \sum_{l=0}^{N-1} w^{lj} \bar{w}^{lk} \\ &= \sum_{l=0}^{N-1} e^{i\frac{2\pi}{N}lj} e^{-i\frac{2\pi}{N}lk} = \sum_{l=0}^{N-1} e^{i\frac{2\pi}{N}l(j-k)} \\ &= \begin{cases} N, & k = j, \\ 0, & k \neq j \end{cases} \end{aligned}$$

Dobivamo da vrijedi $F_N^* F_N = NI$ odakle zaključujemo da je matrica $\frac{1}{\sqrt{N}}F_N$ unitarna, pa stoga vrijedi

$$F_N^{-1} = \frac{1}{N}F_N^* = \frac{1}{N}\bar{F}_N.$$

Time smo ustvari dobili formulu za rješavanje sustava (3.19). Ovdje je važno uočiti sljedeću činjenicu: da bismo riješili sustav (3.19) potrebno je pomnožiti matricu $\frac{1}{N}\overline{F}_N$ vektorom desne strane y , što zahtijeva $O(N^2)$ operacija (množenja kompleksnih brojeva). Vektor nepoznatih koeficijenata $c = [c_1 \ c_2 \ \dots \ c_{N-1}]^T$ računa se na sljedeći način:

$$c = F_N^{-1}y = \frac{1}{N}\overline{F}_N y.$$

Raspišemo li gornji izraz koordinatno, dobivamo formulu za računanje koordinata vektora c :

$$c_j = \frac{1}{N} \sum_{k=0}^{N-1} y_k w^{-jk}, \quad j = 0, \dots, N-1. \quad (3.22)$$

Nadalje, uočimo da je prema (3.19)

$$y_k = \sum_{j=0}^{N-1} c_j w^{jk}, \quad k = 0, \dots, N-1. \quad (3.23)$$

Dakle, za zadane podatke (x_j, y_j) , na mreži $x_j = 2\pi j/N$, $j = 0, \dots, N-1$, te $y_j = f(x_j)$ pri čemu je f 2π -periodička funkcija, formula (3.22) daje koeficijente pripadnog faznog polinoma danog izrazom (3.16), dok koeficijente traženog interpolacijskog polinoma (3.12) dobivamo iz formula (3.17). Naglasimo da se u točkama mreže interpolacijski polinom i njegov pripadni fazni polinom podudaraju, to jest vrijedi

$$\psi(x_j) = p(x_j), \quad j = 0, \dots, N-1. \quad (3.24)$$

Naravno, van točaka mreže vrijednosti spomenutih polinoma se razlikuju, odnosno za $x \neq x_j$ vrijedi $\psi(x) \neq p(x)$.

Da bismo potpuno riješili interpolacijsku zadaću preostaje nam još odrediti vrijednosti interpolacijskog polinoma $\psi(x)$ na mreži finijoj od zadane mreže (3.14). Radi jednostavnosti uzimamo da su čvorovi nove mreže također međusobno jednako udaljeni. Preostaje nam dakle, izračunati interpolacijski polinom $\psi(x)$ u točkama $\tilde{N} = 2\tilde{M} + 1$, pri čemu je $\tilde{M} > M$. U tu svrhu interpolacijski polinom $\psi(x)$ zapišimo u obliku

$$\psi(x) = a_0 + \sum_{k=1}^{\tilde{M}} \left[a_k \cos(kx) + b_k \sin(kx) \right],$$

pri čemu je

$$a_k = b_k = 0, \quad k = M+1, \dots, \tilde{M}. \quad (3.25)$$

Čvorovi nove mreže dani su formulom

$$\tilde{x}_j = \frac{2\pi}{\tilde{N}}j, \quad j = 0, \dots, \tilde{N} - 1,$$

a trigonometrijski interpolacijski polinom $\psi(x)$ i pripadni fazni polinom $\tilde{p}(x)$ podudaraju se u točkama mreže, odnosno vrijedi

$$\psi(\tilde{x}_j) = \tilde{p}(\tilde{x}_j) = \sum_{k=0}^{\tilde{N}-1} \tilde{c}_k e^{ik\tilde{x}_j},$$

pri čemu su koeficijenti vezani kao i u izrazu (3.17):

$$a_0 = \tilde{c}_0, \quad a_k = \tilde{c}_k + \tilde{c}_{\tilde{N}-k}, \quad b_k = i(\tilde{c}_k - \tilde{c}_{\tilde{N}-k}), \quad k = 1, \dots, \tilde{M}.$$

Iz gornjih formula te uvažavajući (3.25) slijedi

$$\tilde{c}_k = 0, \quad \tilde{c}_{\tilde{N}-k} = 0, \quad k = M + 1, \dots, \tilde{M},$$

to jest

$$\begin{aligned} \tilde{c}_{M+1} &= \dots = \tilde{c}_{\tilde{M}} = 0, \\ \tilde{c}_{\tilde{N}-\tilde{M}} &= \dots = \tilde{c}_{\tilde{N}-M-1} = 0. \end{aligned}$$

Kako je $\tilde{N} - \tilde{M} = 2\tilde{M} + 1 - \tilde{M} = \tilde{M} + 1$ slijedi

$$\tilde{c}_{M+1} = \dots = \tilde{c}_{\tilde{N}-(M+1)}.$$

Nadalje, $a_0 = \tilde{c}_0$, dok za $k = 1, \dots, M$ vrijedi

$$a_k = \tilde{c}_k + \tilde{c}_{\tilde{N}-k}, \quad b_k = i(\tilde{c}_k - \tilde{c}_{\tilde{N}-k}),$$

odnosno

$$\tilde{c}_k = c_k, \quad \tilde{c}_{\tilde{N}-k} = c_{N-k}, \quad k = 1, \dots, M.$$

Konačno, zapišemo li vektor \tilde{c} u obliku

$$[c_0 \quad \dots \quad c_M \quad 0 \quad \dots \quad 0 \quad c_{M+1} \quad \dots \quad c_{N-1}]^T,$$

te uvažimo li uvjete

$$\psi(\tilde{x}_j) = \tilde{p}(\tilde{x}_j)$$

vidimo da se izvrednjavanje trigonometrijskog polinoma $\psi(x)$ u točkama finije mreže \tilde{x}_j , $j = 0, \dots, \tilde{N} - 1$ svodi na računanje izraza

$$\tilde{y}_j = \sum_{k=0}^{\tilde{N}-1} \tilde{c}_k w^{jk},$$

gdje je $w = e^{2\pi/\tilde{N}}$.

Formule (3.22) i (3.23) prema svojoj važnosti nadilaze problem trigonometrijske interpolacije te imaju široku primjenu u inženjerstvu (na primjer u teoriji signala). U skladu s time imamo sljedeće važne definicije.

Definicija 3.1 Preslikavanje koje vektoru $[y_0 \ y_1 \ \dots \ y_{N-1}]^T$ pridružuje vektor $[c_0 \ c_1 \ \dots \ c_{N-1}]^T$ prema formuli

$$c_j = \frac{1}{N} \sum_{k=0}^{N-1} y_k w^{-jk}, \quad j = 0, \dots, N-1 \quad (3.26)$$

naziva se diskretna Fourierova transformacija (DFT) vektora y .

Slično, preslikavanje koje vektoru

$$[c_0 \ c_1 \ \dots \ c_{N-1}]^T$$

pridružuje vektor

$$[y_0 \ y_1 \ \dots \ y_{N-1}]^T$$

$$y_k = \sum_{j=0}^{N-1} c_j w^{jk}, \quad k = 0, \dots, N-1. \quad (3.27)$$

naziva se inverzna diskretna Fourierova transformacija (IDFT) vektora c .

Napomena 3.3 Formula diskretne Fourierove transformacije može se dobiti primjenom (kompozitne) trapezne formule (vidi Sekciju ??) na računanje koeficijenta Fourierovog reda 2π -periodičke funkcije f . Pripadni Fourierov red glasi

$\sum_{k=-\infty}^{\infty} c_k e^{ikx}$ a njegovi koeficijenti dani su formulom

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx.$$

Zaista, primijenimo li trapezno pravilo za aproksimaciju integrala $(2\pi)^{-1} \int_0^{2\pi} F(x) dx$, uz $h = 2\pi/N$ dobivamo

$$\frac{1}{2\pi} \int_0^{2\pi} F(x) dx \approx \frac{1}{2\pi} \frac{\pi}{N} [Y_0 + 2(Y_1 + \dots + Y_{N-1}) + Y_N],$$

gdje je $Y_j = F(hj) = F(2\pi j/n)$, $j = 0, \dots, N$. Nadalje, ako je $F(x)$ 2π -periodička, onda je $Y_0 = Y_N$ pa gornja formula prelazi u

$$\frac{1}{2\pi} \int_0^{2\pi} F(x) dx \approx \frac{1}{N} \sum_{j=0}^{N-1} Y_j.$$

Primijenimo li dobivenu formulu za aproksimaciju integrala u k -tom koeficijentu c_k Fourierovog reda, dobivamo

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx \approx \frac{1}{N} \sum_{j=0}^{N-1} f\left(\frac{2\pi j}{N}\right) e^{-i2\pi jk/N}.$$

Konačno, vrijedi

$$c_k \approx \frac{1}{N} \sum_{j=0}^{N-1} y_j w^{-jk}, \quad (3.28)$$

pri čemu je $y_j = f(2\pi j/N)$, te $w = e^{2\pi i/N}$. Napomenimo ovdje da unatoč činjenici da je $w^{-jk} = w^{-j(k+N)}$, izraz (3.28) ne aproksimira koeficijente c_k za $k \geq N$ iz jednostavnog razloga što koeficijenti c_k ne čine periodički niz. Ustvari, izraz (3.28) aproksimira c_k samo za indekse k koji su relativno mali u odnosu na N , odnosno kada je $h = 2\pi/N$ malo obzirom na frekvenciju k .

Napomena 3.4 U literaturi se mogu naći različite formulacije diskretne i njoj inverzne Fourierove transformacije u kojima se umjesto $w = e^{i\frac{2\pi}{N}}$ koristi \bar{w} , odnosno faktor $1/N$ dolazi u definiciji IDFT, a ne kao kod nas u izrazu za DFT. Tako je na primjer DFT niza (x_k) duljine N u Matlabu implementirana naredbom `X = fft(x)` pri čemu je

$$X(k) = \sum_{j=1}^N x(j) \bar{w}^{(j-1)(k-1)},$$

dok pripadna IDFT dobivamo naredbom `x = ifft(X)` koja se temelji na formuli

$$x(j) = \frac{1}{N} \sum_{k=1}^N X(k) w^{-(j-1)(k-1)}.$$

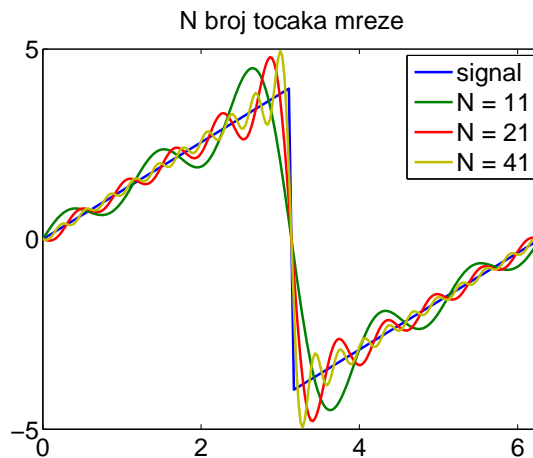
Sumiranje u gornjim formulama pomaknuto je za indeks 1 budući da se vektori u Matlabu indeksiraju od 1, a ne od 0.

Funkcija `fft` vrši DFT zadanog niza koristeći algoritam brze Fourierove transformacije koji omogućava da se broj potrebnih operacija smanji sa $\mathcal{O}(N^2)$ na $\mathcal{O}(N \log_2 N)$ gdje je N duljina zadanog vektora. Implementaciju spomenutog algoritma obradili smo u sekciji 3.3.3.

Trigonometrijska interpolacija u Matlabu je implementirana funkcijom `interpft`. Vrlo pojednostavljeno rečeno, funkcija `interpft` uzima zadani niz y_k , te broj čvorova *finije* ekvidistantne mreže. Zatim algoritmom brze Fourierove transformacije određuje pripadne koeficijente c_k faznog interpolacijskog polinoma $p(x)$, te konačno izvrednjuje traženi interpolacijski polinom $\psi(x)$ na finijoj ekvidistantnoj mreži. Ustvari, problem trigonometrijske interpolacije svodi se na kombiniranje funkcija `fft` i `ifft`. Pogledajmo na nekoliko primjera.

Primjer 3.2 Pilasti (*eng. sawtooth*) signal perioda π može se u Matlabu zadati sljedećim naredbama

```
y = inline('rem(t + pi, 2*pi)/pi*4-4');
```



Slika 3.1: Trigonometrijska interpolacija pilastog signala.

Aproksimacija gore spomenutog signala trigonometrijskim polinomima prikazana je na slici 3.1 a dobiva se u Matlabu sljedećim nizom naredbi

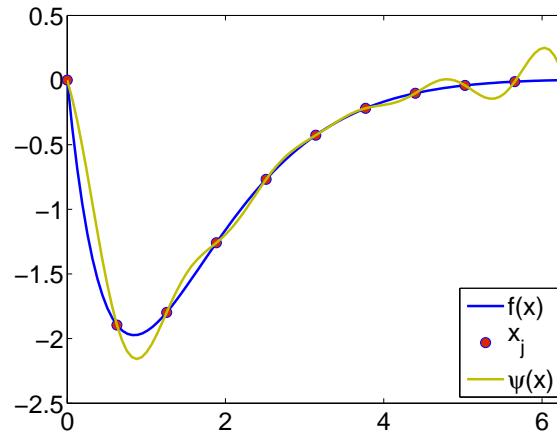
```
t = linspace(0,2*pi);
plot(t,y(t),'LineWidth',2);
hold on;

M = [5 10 20];
N = 2.*M+1;
%-----%
for i = N
    xi = linspace(0,2*pi-2*pi/i,i);
    yi = y(xi);

    xii = linspace(0,2*pi-2*pi/(81*3),81*3);
    yii = interpft(yi,81*3);
    plot(xii, yii);
    pause;
end
hold off;
```

Primjer 3.3 Zadanu funkciju $f(x) = x(x - 2\pi)e^{-x}$, $x \in [0, 2\pi]$ perioda 2π treba interpolirati trigonometrijskim polinomom na ekvidistantnoj mreži koja se sastoji od 10 zadanih točaka. koristeći Matlabovu funkciju `interpft`.

```
y = inline('x.*(x-2*pi).*exp(-x)');
%-----%
x = linspace(0,2*pi);
```



Slika 3.2: Trigonometrijska interpolacija funkcije $f(x) = x(x - 2\pi)e^{-x}$

```
plot(x,y(x),'LineWidth',2);
hold on;

xi = linspace(0,2*pi-2*pi/10,10);
plot(xi,y(xi),'bo');

zz = interpft(y(xi),100);
xz = linspace(0,2*pi-2*pi/100,100);

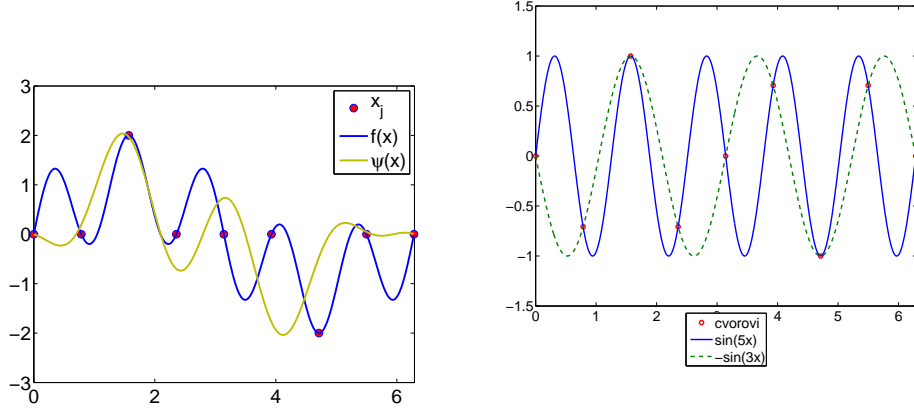
plot(xz,zz)
hold off;
```

Rezultat je prikazan na slici 3.2.

Aliasing. Odabrani broj čvorova mreže može biti nedostatan za interpolaciju članova sa visokim frekvencijama. U tu svrhu promotrimo sljedeći primjer.

Primjer 3.4 Funkciju $f(x) = \sin(x) + \sin(5x)$, $x \in [0, 2\pi]$ interpolirajmo na ekvidistantnoj mreži koristeći 8 točaka interpolacije.

Rješenje je prikazano na slici 3.3 (lijevo). Funkcija koju interpoliramo je zbroj funkcija različitih frekvencija. Broj čvorova interpolacije nije dovoljan za prikazivanje najviše frekvencije. Na slici ?? vidimo da točke interpolacije leže na grafu funkcije $\sin(5x)$ ali i na grafu funkcije $-\sin(3x)$. Funkcija koju smo ustvari interpolirali je $\sin(x) - \sin(3x)$, a ne zadana funkcija $\sin(x) + \sin(5x)$. Čvorove mreže kao i funkcije $\sin(5x)$ i $-\cos(3x)$ prikazani su na slici 3.3, desno. Efekt aliasinga u ovom slučaju se može izbjeći povećanjem broja čvorova mreže.

Slika 3.3: Aliasing kod trigonometrijske interpolacije $f(x) = \sin(x) + \sin(5x)$

Filtriranje signala. Diskretna Fourierova transformacija, odnosno algoritam FFT nadilazi problem trigonometrijske interpolacije te ima široke primjene, na primjer u teoriji signala i kompresiji slika. Ovdje navodimo primjer vezan uz filtriranje signala, preuzet iz knjige [4]. Pojednostavljeno rečeno, filtriranje signala svodi se na poništavanje koeficijenta pripadnog Fourierovog reda koji se nalaze uz članove koji odgovaraju visokim frekvencijama.

Primjer 3.5 Zadani signal

$$f(t) = e^{-\cos(t)^2} (\sin(2t) + 2 \cos(4t) + 0.4 \sin(t) \sin(50t)), \quad t \in [0, 2\pi]$$

želimo filtrirati od šumova prouzrokovanih visokim frekvencijama. Najprije diskretiziramo signal na intervalu $[0, 2\pi]$ na mreži

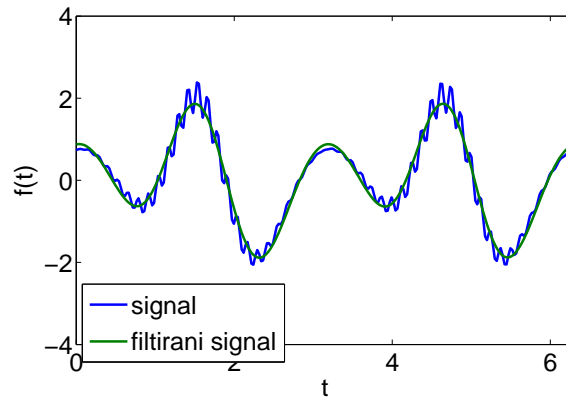
$$x_j = \frac{2\pi}{N}j, \quad j = 0, \dots, N-1,$$

za $N = 256$. Zatim napravimo diskretnu Fourierovu transformaciju (vidi formulu (3.26)) čime dobijemo diskretne Fourierove koeficijente c_k , $k = 0, \dots, 256$. Sa slike 3.4 vidimo da prisutni šum ima frekvencije veće od 5 pa signal filtriramo tako da zadržimo koeficijente c_k , $0 \leq k \leq 5$. Radi ekvidistantne mreže na $[0, 2\pi]$ vrijedi

$$e^{-ikx_j} = e^{i(N-k)x_j},$$

odakle slijedi da moramo zadržati i koeficijente c_k , $k = 251, \dots, 255$. Ostale koeficijente stavimo na 0, to jest imamo $c_k = 0$, $6 \leq k \leq 128$ slijedi da je $c_k = 0$, $128 \leq k \leq 250$. Dakle, napravili smo filtrirani diskretni niz

$$[c_0 \quad \dots \quad c_5 \quad 0 \quad \dots \quad 0 \quad c_{251} \quad \dots \quad c_{256}]$$



Slika 3.4: Primjena FFT u filtriranju signala.

Na tako dobiveni filtrirani diskretni niz frekvencija primijenimo inverznu diskretnu Fourierovu transformaciju čime dobivamo filtrirani signal prikazan na slici 3.4.

```
t = linspace(0,2*pi,2^8);
y = exp(-(cos(t).^2)).*(sin(2*t)+2*cos(4*t) +0.4*sin(t).*sin(50*t));
fy = fft(y); % computes fft of y
filterfy = [fy(1:6) zeros(1,2^8-12) fy(2^8-5:2^8)] ;
filtery = ifft(filterfy);
plot(t, y, t, filtery)
```

3.3.3 Brza Fourierova transformacija (FFT)

U prethodnoj sekciji smo vidjeli da je za računanje diskretne Fourierove transformacije izrazom (3.26) potrebno $O(N^2)$ operacija¹. Na primjer, za računanje diskretnog Fourierovog transformata niza c_k , $k = 0, \dots, N - 1$ gdje je $N = 1024 = 2^{10}$ bit će napravljeno 1024×1024 operacija, što je još uvijek bez problema izvedivo na današnjim računalima. No, kako je u primjenama obično DFT potrebno ponoviti više tisuća puta, javlja se potreba za nalaženjem algoritma koji bi zahtijevao manje operacija od vec spomenutih $O(N^2)$. Brza Fourierova transformacija (FFT) za pronalaženje DFT zadanog niza od N elemenata treba svega $O\left(\frac{1}{2}N \log_2 N\right)$ operacija. To dakle znači da smo u našem primjeru za $N = 1024$ broj potrebnih operacija reducirali sa 1024×1024 na 5×1024 , što daje smanjenje broja operacija za približno faktor 200.

U ovoj sekciji označavat ćemo n -ti korijen iz jedinice sa w_n . Glavna ideja na kojoj se zasniva FFT proizlazi iz sljedećeg zapažanja. Neka su $M, N \in \mathbb{N}$ takvi

¹Pod operacijom podrazumijevamo kompleksno množenje

da vrijedi $N = 2M$. Tada je

$$w_N^2 = w_M. \quad (3.29)$$

Činjenica (3.29) nam omogućuje da komponente vektora $y = F_N c \in \mathbb{C}^N$ dobijemo iz komponenti dvaju vektora y' i y'' koji imaju $M = N/2$ elemenata. U tu svrhu krenimo od samog vektora $c = [c_0 \ c_1 \ \dots \ c_{N-1}]^T$ i odvojimo njegove parne i neparne komponente:

$$\begin{aligned} c' &= [c_0 \ c_2 \ \dots \ c_{n-2}]^T \\ c'' &= [c_1 \ c_3 \ \dots \ c_{n-1}]^T \end{aligned}$$

Nadalje, označimo

$$y' = F_M c' \text{ i } y'' = F_M c''.$$

Sljedeći korak je pokazati da vektor y možemo zapisati pomoću vektora y' i y'' . Zapišimo izraz $y = F_N c$ po komponentama. Imamo

$$y_j = \sum_{k=0}^{N-1} w_N^{kj} c_k, \quad j = 0, \dots, N-1. \quad (3.30)$$

Rastavimo izraz (3.30) na parne i neparne komponente:

$$y_j = \sum_{k=0}^{M-1} w_N^{2kj} c_{2k} + \sum_{k=0}^{M-1} w_N^{(2k+1)j} c_{2k+1}, \quad j = 0, \dots, N-1. \quad (3.31)$$

Zapišimo izraz (3.31) pomoću vektora c' i c'' . U tu svrhu iskoristimo činjenicu (3.29) i dobivamo

$$y_j = \sum_{k=0}^{M-1} w_M^{kj} c'_k + w_N^j \sum_{k=0}^{M-1} w_M^{kj} c''_k, \quad j = 0, \dots, N-1. \quad (3.32)$$

Uočimo sada da je za $j = 0, \dots, M-1$ prva suma u izrazu (3.32) upravo jednaka y'_j , dok druga suma predstavlja y''_j . Dakle, prvih M komponenti vektora y izrazili smo pomoću vektora y' i y'' na sljedeći način:

$$y_j = y'_j + w_N^j y''_j, \quad j = 0, \dots, M-1. \quad (3.33)$$

Preostaje nam još na sličan način izraziti komponente y_M, \dots, y_{N-1} . U tu svrhu u izrazu (3.32) napravimo pomak indeksa $l = j - M$. Kako je $j = M, M+1, \dots, N-1$ i $N = 2M$ slijedi da je $l = 0, 1, \dots, M-1$. Dobivamo

$$\begin{aligned} y_{M+l} &= \sum_{k=0}^{M-1} w_M^{k(M+l)} c'_k + w_N^{M+l} \sum_{k=0}^{M-1} w_M^{k(M+l)} c''_k \\ &= \sum_{k=0}^{M-1} w_M^{kl} c'_k - w_N^l \sum_{k=0}^{M-1} w_M^{kl} c''_k, \end{aligned} \quad (3.34)$$

pri čemu smo iskoristili da vrijede sljedeće tvrdnje

$$w_M^{k(M+l)} = w_M^{kl} \text{ i } w_N^{M+l} = -w_N^l.$$

Konačno, izraz (3.34) možemo zapisati na sljedeći način:

$$y_{j+M} = y'_j - w_N^j y''_j, \quad j = 0, \dots, M-1. \quad (3.35)$$

Zapišimo formule (3.33) i (3.35) u matičnom obliku. Dobivamo

$$\begin{aligned} y &= \begin{bmatrix} F_M(Px)_1 + D_M F_M(Px)_2 \\ F_M(Px)_1 - D_M F_M(Px)_2 \end{bmatrix} = \begin{bmatrix} F_M & D_M F_M \\ F_M & -D_M F_M \end{bmatrix} Px \\ &= \begin{bmatrix} I & D_M \\ I & -D_M \end{bmatrix} \begin{bmatrix} F_M & \\ & F_M \end{bmatrix} Px \end{aligned} \quad (3.36)$$

Sama ideja FFT algoritma zasniva se na upravo izvedenoj faktorizaciji matrice F_N . Pokažimo sada kako gore opisani algoritam funkcionira na primjeru $N = 4$, $M = 2$. Grafički, korake FFT algoritma možemo prikazati na sljedeći način:

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} \rightarrow \begin{bmatrix} c_0 \\ c_2 \\ c_1 \\ c_3 \end{bmatrix} \rightarrow \begin{bmatrix} F_2 x' \\ F_2 x'' \end{bmatrix} \rightarrow [y]. \quad (3.37)$$

Uvjerimo se u to na primjeru $N = 4$. Na temelju izraza (3.33) i (3.35) matricu F_4 možemo rastaviti na sljedeći način:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & i \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -i \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (3.38)$$

Dakle, matricu F_4 dimenzije 4×4 možemo zapisati pomoću blok-matrice dimenzije 2×2 na sljedeći način:

$$F_4 = \begin{bmatrix} I & D_2 \\ I & -D_2 \end{bmatrix} \begin{bmatrix} F_2 & 0 \\ 0 & F_2 \end{bmatrix} P \quad (3.39)$$

gdje je $D = \text{diag}(1, i)$, a P je permutacijska matrica. Permutacijska matrica P rastavlja vektor c na c' i c'' , dok prva matrica u izrazu (3.39) dolazi iz formula (3.33) i (3.35).

Konačno, pogledajmo koliku uštedu nosi sa sobom FFT algoritam u odnosu na rješavanje linearnog sustava (3.19) formulom (??) što zahtijeva N^2 kompleksnih množenja. Neka je $N = 2^l$. S $A(N)$ označimo broj operacija potrebnih da se napravi FFT. Najprije je, pomoću formula (3.33) i (3.35) potrebno formirati vektore y' i y'' što zahtijeva $2A\left(\frac{N}{2}\right)$ operacija. Također, y'' se množi s vrijednostima

w_N^j što zahtijeva još $N/2$ množenja. Napomenimo da imamo još $N/2$ zbrajanja i $N/2$ oduzimanja koja u ovom računu zanemarujemo. Ukupni broj kompleksnih množenja izračunat ćemo rekursivno. Imamo

$$\begin{aligned} A(N) &= 2A\left(\frac{N}{2}\right) + \frac{N}{2} = 2\left[2A\left(\frac{N}{4}\right) + \frac{N}{4}\right] + \frac{N}{2} \\ &= 2^2A\left(\frac{N}{4}\right) + \frac{N}{2} + \frac{N}{2} = 2^2\left[2A\left(\frac{N}{8}\right) + \frac{N}{8}\right] + \frac{N}{2} + \frac{N}{2} \\ &= 2^3A\left(\frac{N}{2^3}\right) + \frac{N}{2} + \frac{N}{2} + \frac{N}{2} = \dots \end{aligned}$$

Gornji račun vodi nas na sljedeći izraz:

$$A(N) = 2^{l-1}A\left(\frac{N}{2^{l-1}}\right) + (l-1)\frac{N}{2}.$$

Iskoristimo li sada da je $N = 2^l$ i $A(2) = 4$ dobivamo

$$\begin{aligned} A(N) &= 2^{l-1}A(2) + (l-1)\frac{N}{2} = 2^{l+1} + (l-1)\frac{N}{2} \\ &= 2N + (l-1)\frac{N}{2} = \frac{1}{2}lN + \frac{3}{2}N \\ &= \frac{1}{2}N \log_2 N + \frac{3}{2}N. \end{aligned}$$

Konačno, zaključujemo da broj kompleksnih množenja potrebnih za izvođenje FFT iznosi

$$\frac{1}{2}N \log_2 N + O(N).$$

Bibliografija

- [1] M. T. Heath. *Scientific Computing: An Introductory Survey*. McGrawHill, New York, 2002.
- [2] D. Kressner. *Numerische Methoden*.
<http://www.math.ethz.ch/education/bachelor/lectures/fs2008/math/nm>,
ETH Zuerich 2008.
- [3] L. Erdos. *Numerical Mathematics I*.
<http://www.mathematik.uni-muenchen.de/~lerdos/Notes>,
Georgia Tech 2002 and TU Muenchen 2008.
- [4] A. Baggers, F.J. Narcowich. *A First Course in Wavelets with Fourier Analysis.*, Texas AM University, Texas 200?.
- [5] C. Moler. *Numerical Computing with MATLAB and Experiments with MATLAB*.
<http://www.mathworks.com/company/aboutus/founders/clevemoler.html>
- [6] A. Quarteroni, R. Sacco, F. Saleri. *Numerical Mathematics*. Texts in Applied Mathematics, Springer, Berlin, Heidelberg 2007.
- [7] J. L. Buchanan, P. R. Turner. *Numerical Methods and Analysis*. McGraw-Hill, Inc., New York, 1992.
- [8] D. Goldberg. *What Every Computer Scientist Should Know About Floating-Point Arithmetic*. ACM Computing Surveys, Vol. 23, No. 1, March 1991.
- [9] I. Ivanšić. *Numericka matematika*. Element, Zagreb, 2005.
- [10] KinChen D. Kincaid, W. Cheney. *Numerical Analysis, Mathematics of Scientific Computing*. Brooks/Cole Publishing Company. Pacific Grove, California, 1991.
- [11] M. L. Overton. *Numerical Computing with IEEE Floating Point Arithmetic*. SIAM, Philadelphia, 2001.
- [12] G. Dahlquist, A. Björck. *Numerical Methods in Scientific Computing*. SIAM, 2007.
- [13] Z. Drmač, V. Hari, M. Rogina, S. Singer, S. Singer, I. Slapničar.
Numericka analiza, predavanja i vježbe.
<http://web.math.hr/~rogina/2001096/numan.html>,
PMF-MO, Zagreb, 2003.

- [14] E. Sulli, D.F. Mayers. *An Introduction to Numerical Analysis*. Cambridge University Press, London, 2003.
- [15] N.J. Higham. *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia, 1995.
- [16] A. Jüngel. *Quasi-hydrodynamic Semiconductor Equations*. Birkhäuser, Basel, 2001.
- [17] A. Jüngel. *Transport Equations for Semiconductors*. Lecture Notes, Johannes Gutenberg-Universität Mainz, 2004.
- [18] N. Klusdahl, A. Krizan, D. Ferry, and C. Ringhofer. Self-consistent study of the resonant-tunneling diode. *Phys. Rev. B* 39 (1989), 7720-7735.
- [19] S. Singer. *Numerička matematika*. Skripta, PMF-MO, Zagreb 2008.
- [20] R. Pinnau.
- [21] G. Strang. *Linear Algebra*. Brooks/Cole Thompson Learning, 1986.
- [22] A. Quarteroni, F. Saleri. *Scientific Computing with MATLAB and Octave*. Springer, Milano 2006.

4

Dodatak

4.1 Neki važni pojmovi iz linearne algebre i matematičke analize

4.1.1 "Veliko" O i "malo" o notacija

Često je potrebno uspoređivati asimptotsko ponašanje zadanih realnih funkcija $f(x)$ i $g(x)$ kada varijabla x teži ka x_0 . U tu svrhu koristimo notaciju "veliko" i "malo" O^1 s kojom smo se upoznali u Matematici 2. Ponovimo ovdje, radi potpunosti, definiciju spomenutih simbola O i o. Neka su $f, g: X \mapsto \mathbb{R}$, $X \subset \mathbb{R}$ te x_0 točka gomilišta skupa X (x_0 može biti jednako ∞). Simbole $\mathcal{O}(\cdot)$ i $o(\cdot)$ definiramo na sljedeći način:

$$f = \mathcal{O}(g) \quad (x \rightarrow x_0) \iff \limsup_{x \rightarrow x_0} \left| \frac{f(x)}{g(x)} \right| < \infty \quad (4.1)$$

$$f = o(g) \quad (x \rightarrow x_0) \iff \lim_{x \rightarrow x_0} \left| \frac{f(x)}{g(x)} \right| = 0. \quad (4.2)$$

Uočimo najprije da je izraz (4.1) ekvivalentan tvrdnji da postoji konstanta $C > 0$ takva da vrijedi $\left| \frac{f(x)}{g(x)} \right| < C$, za svako x iz neke okoline točke x_0 . Nadalje, važno je naglasiti da se znak jednakosti u izrazima (4.1), (4.2) koristi u simboličkom smislu. Tako naprimjer iz tvrdnji $f_1(x) = o(g(x))$ i $f_2(x) = o(g(x))$ ne slijedi nužno da je $f_1 = f_2$. Ono što u tom slučaju možemo zaključiti je naprimjer da vrijedi $f_1(x) - f_2(x) = o(g(x))$, odnosno $f_1(x) + f_2(x) = o(g(x))$.

Primjer 4.1 Uvjerite se da vrijede sljedeće tvrdnje:

- $f(x) = 7x + 6x^2$, $g(x) = x^2$. Vrijedi: $f = \mathcal{O}(g)$, $x \rightarrow 1$

¹Landauova notacija

- $f(x) = x \ln(1+x)$, $g(x) = x$. Vrijedi: $f = o(g)$, $x \rightarrow 0+$
- $e^x = 1 + x + \frac{x^2}{2} + \mathcal{O}(x^3)$, $x \rightarrow 0$
- $\ln(x) = o(x^\alpha)$, $x \rightarrow \infty$, $\alpha > 0$
- $\frac{1}{1-h} = 1 + h + \mathcal{O}(h^2)$, $h \rightarrow 0$ ($0 < h < 1$)
- $2n^4 + 3n^3 + 5n^2 + n = \mathcal{O}(n^4)$, $n \rightarrow \infty$

4.1.2 Vektorski prostori

Jedan od najvažnijih koncepata u matematici je pojam vektorskog prostora. Ugrubo rečeno, vektorski prostor je svaki skup u kome znamo zbrajati elemente i množiti ih s realnim, odnosno kompleksnim brojevima. Pri tome zbrajanje i množenje zadovoljava uobičajena pravila zbrajanja i množenja realnih brojeva i ono nas ne vodi izvan promatranog skupa. Preciznije, pogledajmo sljedeću definiciju.

Definicija 4.1 Za neprazni skup X kažemo da je **vektorski prostor** ukoliko su zadovoljeni sljedeći aksiomi.

A1 Suma svaka dva elementa $x, y \in X$ jedinstveno određuje element $z = x+y \in X$ pri čemu vrijede sljedeća pravila

- (i) $x + y = y + x$ (komutativnost)
- (ii) $(x + y) + z = x + (y + z)$ (asocijativnost)
- (iii) Postoji element $0 \in X$ takav da vrijedi $x + 0 = x$, $\forall x \in X$.
- (iv) Za svaki element $x \in X$ postoji element $-x$ takav da vrijedi $x + (-x) = 0$.

A2 Za svaki realan broj $\alpha \in \mathbb{R}$ te za svaki element $x \in X$ jedinstveno je određen element $\alpha x \in X$ tako da vrijedi

- (i) $\alpha(\beta x) = (\alpha\beta)x$
- (ii) $1x = x$

A3 Zbrajanje i množenje definirano aksiomima A1 i A2 zadovoljava zakone distributivnosti:

- (i) $(\alpha + \beta)x = \alpha x + \beta x$
- (ii) $\alpha(x + y) = \alpha x + \alpha y$

4.1.3 Neprekidnost, ograničenost, integrabilnost

4.1.4 Teoremi srednje vrijednosti

Propozicija 4.1 *Neka su funkcije $f(x)$ i $g(x)$ integrabilne na segmentu $[a, b]$, te neka je $g(x) \geq 0, \forall x \in [a, b]$. Tada vrijedi*

$$m \int_a^b g(x)dx \leq \int_a^b f(x)g(x)dx \leq M \int_a^b g(x)dx,$$

gdje je $m = \inf f(x)$, $M = \sup f(x)$.

Teorem 4.1 *Integralni teorem srednje vrijednosti Neka su funkcije f i g integrabilne na segmentu $[a, b]$ i neka je $m = \inf f(x)$ i $M = \sup f(x)$. Nadalje, neka je $g(x) \geq 0 \forall x \in [a, b]$. Tada postoji $c \in \mathbb{R}$, $m \leq c \leq M$ takav da vrijedi*

$$\int_a^b f(x)g(x)dx = c \int_a^b g(x)dx. \quad (4.3)$$

Posebno, ako je g neprekidna na $[a, b]$ onda postoji $\tau \in [a, b]$ takav da vrijedi

$$\int_a^b f(x)g(x)dx = f(\tau) \int_a^b g(x)dx.$$

Dokaz.