Osnove obradbe signala:
Diskretna kosinusna transformacija

Tomislav Petković

Sveučilište u Zagrebu

prosinac 2021.

Rastav signala i spektar

U problemu predstavljanja signala $x(t)$ i $x[n]$ preko odabranih baznih funkcija, bilo kao prebrojivih linearnih kombinacijih

$$x(t) = \sum_k s_k \cdot \phi_k(t) \quad \text{i} \quad x[n] = \sum_k s_k \cdot \phi_k[n], \tag{1}$$

ili kao neprebrojivih kombinacija

$$x(t) = \int_\kappa s(\kappa) \cdot \phi(t, \kappa)\, d\kappa \quad \text{i} \quad x[n] = \int_\kappa s(\kappa) \cdot \phi(n, \kappa)\, d\kappa, \tag{2}$$

smo podproblem odabira baznih funkcija $\phi(\cdot)$ ostavili po strani.

U ovoj cjelini ćemo razmotriti jedan mogući pristup konstrukciji **sustava baznih funkcija**. Rezultat odabranog postupka konstrukije će biti diskretna kosinusna transformacija.

Diskretna kosinusna transformacija

Kada govorimo o diskretnoj kosinusnoj transformaciji u obradbi signala uobičajeno podrazumijevamo takozvanu DCT-II transformaciju čije bazne funkcije $\phi_k[n]$ su diskretne kosinusoide.

Bazne funkcije $\phi_k[n]$ diskretne kosinusne transformacije su čisto realne funkcije, odnosno

$$\phi_k[n] : \{0, 1, \ldots, N-1\} \to \mathbb{R} \tag{3}$$

te zato njima možemo prikazati samo čisto realne signale konačnog trajanja (i njihova parna periodička proširenja).

Definirajmo sada DCT-II transformaciju na isti način kao što smo definirali razne Fourierove transformacije.

Definicijski izrazi

Bazne funkcije za DCT-II su $\phi_k[n] = \sqrt{\frac{2-\delta[k]}{N}} \cos \frac{(2n+1)k\pi}{2N}$.

Signal $x[n]$ konačne duljine N predstavljamo kao aditivnu kombinaciju diskretnih kosinusoida $\phi_k[n]$:

$$x[n] = \sum_{k=0}^{N-1} X[k] \sqrt{\frac{2-\delta[k]}{N}} \cos \frac{(2n+1)k\pi}{2N} \tag{4}$$

Spektar $X[k]$ računamo kao:

$$X[k] = \langle x[n], \phi_k[n] \rangle = \sum_{n=0}^{N-1} x[n] \sqrt{\frac{2-\delta[k]}{N}} \cos \frac{(2n+1)k\pi}{2N} \tag{5}$$

Kako smo došli do DCT-II transformacije?

Definiranje DCT-II transformacije preko izraza (4) i (5) u potpunosti definira transformaciju, no ne odgovara nam na neka važna pitanja koja želimo razmotriti:

- Zašto baš te navedene bazne funkcije $\phi_k[n]$?
- Kako smo došli do baznih funkcija $\phi_k[n]$?
- Gdje primjenjujemo DCT-II?

Sva ta pitanja su međusobno povezana, a da bi odgovorili na njih započinjemo s uvođenjem **modela**.

Rastav signala konačnog trajanja

DCT-II transformacija se koristi u kompresiji signala s gubitkom.

Zato započinjemo naše razmatranje s modelom rastava signala konačnog trajanja od N uzoraka na $M \leq N$ baznih funkcija,

$$x[n] = e[n] + \sum_{k=0}^{M-1} s_k \phi_k[n]. \tag{6}$$

Za rastav (6) znamo da greška $e[n]$ iščezava za $M = N$ kada bazne funkcije $\phi_k[n]$ razapinju cijeli prostor.

Želimo odrediti dobre bazne funkcije $\phi_k[n]$ takve da za svaki odabrani $M < N$ greška reprezentacije signala $e[n]$ bude uvijek minimalna u normi dva, odnosno želimo riješiti

$$\underset{\phi_k[n]}{\arg\min} \left\| x[n] - \sum_{k=0}^{M-1} s_k \phi_k[n] \right\|_2. \tag{7}$$

Minimizacijski problem za kompresiju signala

U izrazu

$$\left|\left|e[n]\right|\right|_2 = \left|\left|x[n] - \sum_{k=0}^{M-1} s_k \phi_k[n]\right|\right|_2 \tag{8}$$

znamo izračunati spektar s_k jednom kada su zadani $x[n]$ i $\phi_k[n]$ pa je prema tome spektar zavisna varijabla.

Bazne funkcije $\phi_k[n]$ su nepoznate, odnosno njih želimo dobiti kao rješenje ovog minmizacijskog problema.

Prema tome da bi riješili (7) moramo nekako specificirati $x[n]$.

To možemo napraviti na barem dva različita načina:

1. Prikupimo jako puno **primjera** stvarnih signala $x[n]$ koje želimo kompresirati.
2. Postuliramo neki **model signala** za kojeg očekujemo da dobro opisuje stvarne signale.

Načelni odgovori na postavljena pitanja

Sve navedeno do sada definira **model** unutar kojeg možemo načelno odgovoriti na postavljena pitanja:

- Zašto baš te navedene bazne funkcije $\phi_k[n]$?

  Zato jer minimizraju grešku kod kompresije signala s gubitkom.

- Kako smo došli do baznih funkcija $\phi_k[n]$?

  Rješavanjem minizacijskog problema (7).

- Gdje primjenjujemo DCT-II?

  U kompresiji s gubitkom čisto realnih signala konačnog trajanja.

Pokažimo sada kako modeliramo signal $x[n]$.

Model signala

Navedene bazne funkcije DCT-II su previše elegantne te sigurno nisu rezultat analize jako puno **primjera** stvarnih signala $x[n]$.

Umjesto toga one su rezultat teorijske analize stohastičkog modela signala u kojem signal modeliramo kao snažno korelirani stacionarni Markovljev lanac prvog reda.

Jednostavnije rečeno postuliramo da su susjedni uzorci u signalu uvijek međusobno jako slični.

Stacionarni Markovljev lanac prvog reda

Svaki uzorak signala $x[n]$ je slučajna varijabla poznatih parametara.

Ne smanjujući općenitost neka su srednje vrijednosti uzoraka nula i neka su varijance jedinične, odnosno

$$\mathbb{E}\big[x[n]\big] = 0 \quad \text{i} \quad \mathrm{Var}\big[x[n]\big] = \mathbb{E}\big[x^2[n]\big] = 1. \tag{9}$$

Lanac je Markovljev ako je dovoljno definirati prijelazne vjerojatnosti između susjednih uzoraka.

Za razmatrani problem kompresije signala je dovoljno definirati korelaciju između dva susjedna uzorka

$$\mathbb{E}\big[x[n]x[n+1]\big] = \rho, \tag{10}$$

gdje je $\rho < 1$ jako blizak jedinici (jaka korelacija).

Korelacijska matrica

Stacionarnost i Markovljevo svojstvo nam sada omogućavaju da definiramo korelacije između svih uzoraka signala, odnosno da konstruiramo korelacijsku matricu R (koja je ujedno i kovarijancijska matrica jer su uzorci signala centrirane slučajne varijable).

Element korelacijske matrice u i-tom retku i j-tom stupcu jest

$$\mathbb{E}\left[x[i-1]x[j-1]\right]. \tag{11}$$

Korelacijska matrica Markovljevog lanca prvog reda jest Toeplitzova matrica dimenzija $N \times N$:

$$\mathrm{R} = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{N-1} \\ \rho & 1 & \rho & \cdots & \rho^{N-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{N-1} & \rho^{N-1} & \rho^{N-3} & \cdots & 1 \end{bmatrix} \tag{12}$$

Korelacijska matrica i bazni vektori

Kako nam poznavanje korelacijske matrice pomaže u određivanju baznih vektora i rješavanju problema (7)?

Navedimo prvo par važnih tvrdnji koje opisuju postupak (bez dokaza koje ostavljamo za samostalan rad):

- Produkt očekivanja $\mathbb{E}[XY]$ dvije slučajne varijable X i Y jest skalarni produkt, odnosno

$$\langle X, Y \rangle = \mathbb{E}[XY] \tag{13}$$

- Korelacijska i kovarijancijska matrica su Gramove matrice.
- Svojstveni vektori kovarijancijske matrice su rješenje minimzacijskog problema (7).
- Kada promatramo limes $\rho \to 1$ onda svojstveni vektori kovarijancijske matrice R stacionarnog Markovljevog lanca prvog reda teže prema baznim vektorima DCT-II.

Rješenje minimizacijskog problema

Navdene tvrdnje ukratko opisuju bitne točke rješavanja minimizacijskog problema kojim određujemo bazne funkcije.

Potpuno rješenje sa svim svojim koracima možete naći u poglavlju 9.2.7. udžbenika "Introduction to Orthogonal Transforms: With Applications in Data Processing and Analysis" autora Ruya Wanga.

Zbog duljine i složenosti postupka sve nenavedene detalje rješenja minimizacijskog problema ostavljamo kao samostalni rad za zainteresirane studente.

Napomena: NIJE ih potrebno naučiti za ispit ☺.

Rješenje minimizacijskog problema

Ključni koncept kojeg morate zapamtiti jest da svojstveni vektori kovarijancijske matrice (koja je Gramova matrica) definiraju bazu za optimalnu reprezentaciju nekog signala (ili skupa podataka) obzirom na kvadratnu grešku (norma $p = 2$).

Postupak transformacije stohastičkog signala u optimalnu bazu vezano uz kvadratnu grešku reprezentacije zove se **Karhunen-Loéveova transformacija** (KLT).

U statističkoj analizi podataka isti postupak se zove **analiza glavnih komponenti** ili skraćeno PCA, od engl. *principal component analysis*.

O nazivu DCT-II

U kontinuiranom slučaju za signale $x : \mathbb{R} \to \mathbb{R}$ postoji samo jedna kosinusna i samo jedna sinusna transformacija.

Kontinuiranu kosinusnu transformaciju koristimo za prikaz parnih signala, a sinusnu za prikaz neparnih.

Za vremenski diskretne signale konačnog trajanja $x : \{0, 1, \ldots, N-1\} \to \mathbb{R}$ defincija parnosti i neparnosti signala nije jednoznačna.

Ukupno postoji 8 različitih izbora proširenja signala konačnog trajanja u periodički simetričan signal $x : \mathbb{Z} \to \mathbb{R}$ što znači da postoji osam diskretnih kosinusnih transformacija, od DCT-I do DCT-VIII, od kojih je eto baš DCT-II jako dobra za kompresiju.

Slično tome postoji 8 diskretnih sinusnih transforamcija, od DST-I do DST-VIII.

Nedostatci DCT-II u kompresiji

Bazne funkcije DCT-II su bliske svojstvenim vektorima korelacijske matrice stacionarnog Markovljevog lanca prvog reda i to ih čini dobrim za kompresiju u slučaju kada stvarni signal odgovara tom modelu.

No u stvarnom signalu možemo očekivati prekide koji se ne uklapaju u model signala u kojem su susjedni uzorci jako korelirani jer svaki prekid lokalno prekida korelaciju.

To je razlog zašto JPEG kompresija nije dobar izbor za pohranu skeniranih tekstualnih dokumenata u kojima postoji jako puno naglih prijelaza između crne i bijele boje.

Nedostatci DCT-II u kompresiji

U stvarnoj primjeni signal nikada nije fiksne duljine od N uzoraka.

Stoga se signal razdvaja na nepreklapajuće blokove konačne duljine od N uzoraka te se onda DCT-II koristi za kompresiju pojedinačnih blokova.

No ako svaki nepreklapajući blok kompresiramo zasebno onda ne možemo garantirati neprekinutost na granicama blokova što dovodi do vidljivih artifakata (engl. *blocking artifacts*).

Zbog toga se u kompresiji zvuka gotovo uvijek koristi modificirana diskretna kosinusna transformacija (MDCT) koja osigurava neprekinutost kompresiranog signala na granicama blokova.

Preporučeno čitanje

- Ruye Wang, "Introduction to Orthogonal Transforms: With Applications in Data Processing and Analysis", Cambridge University Press, dio 7.2. i poglavlje 9., posebno dio 9.2.7.
- M. Vetterli, J. Kovačević, V. K. Goyal, "Foundations of Signal Processing" (`https://fourierandwavelets.org/`), dio 6.4.2.
- S. J. Orfanidis, "SVD, PCA, KLT, CCA, and All That", `https://www.ece.rutgers.edu/~orfanidi/ece525/svd.pdf`, dijelovi 15. i 16.