Obrada informacija: Podaci, signali, sustavi te invarijante i simetrije

T. Petković

Sveučilište u Zagrebu

prosinac 2021.

Informacija i podaci

Što je informacija, što su podaci, a što su signali?

Informacija je apstraktni pojam čija definicija ovisi o kontekstu, npr. informacije u svakodnevnom životu su obavijesti i novosti dok u kontekstu teorije informacija govorimo o razrješenju neodređenosti te uvodimo pojam entropije kao mjere neodređenosti.

Informacija je neraskidivo povezana s **podacima**.

U kontekstu **obrade signala** umjesto o informaciji češće govorimo o **podacima** i o **signalima**.

Podaci

Podaci mogu biti:

1. **numerički**, npr. brojevi
2. **kategorički**, npr. osobna imena

Numerički podaci mogu biti **kontinuirani** (npr. realni i kompleksni brojevi) ili **diskretni** (npr. prirodni i cijeli brojevi).

Kategorički podaci su u osnovi vrijednosti iz nekog konačnog skupa oznaka.

No ako te oznake možemo međusobno *uspoređivati*, odnosno ako postoji *uređaj* na skupu oznaka, onda govorimo o **ordinalnim** podacima. Ordinalne podatke je moguće *sortirati*.

Predstavljanje podataka

Računala su izuzetno efikasna i brza u obradi (binarnih) brojeva.

Numerički podaci su stoga prirodan izbor za računalnu obradu.

Podatke koji nisu numerički moramo **kodirati**.

Iako kodiramo kategoričke (ili ordinalne) podatke pomoću brojeva moramo biti svjesni da standardne aritmetičke operacije ($+$, $-$, $\cdot$, $:$) nad pridijeljenim brojčanim kodovima uglavnom nemaju smisla.

Predstavljanje podataka

Ranije ste nadam se vidjeli jedno proizvoljno kodiranje za nukleotide na primjeru problema poravnavanja:

$$\begin{aligned} A\,(adenin) &\quad\mapsto\quad 1 \\ T\,(timin) &\quad\mapsto\quad 2 \\ G\,(gvanin) &\quad\mapsto\quad 3 \\ C\,(citozin) &\quad\mapsto\quad 4 \end{aligned}$$

Je li ta reprezentacija dobra za rješavanje problema poravnavanja križnom korelacijom?

Predstavljanje podataka

Izuzetno poučan primjer predstavljanja podataka jest kodiranje riječi u obradi teksta.

Želimo riječima pridružiti višedimenzionalne vektore na način da struktura vektorskog prostora u kojem predstavljamo riječi dobro opisuje sintaktičke i semantičke pravilnosti jezika.

Jedno od najpoznatijih predstavljanja jest `word2vec`.

1. T. Mikolov, Y. Wen-tau, Z. Geoffrey. "Linguistic regularities in continuous space word representations.", https://www.aclweb.org/anthology/N13-1090.pdf
2. T. Mikolov, K. Chen, G. Corrado, J. Dean. "Efficient Estimation of Word Representations in Vector Space.", https://arxiv.org/abs/1301.3781
3. T. Mikolov, I. Sutskever, K. Chen, G. Corrado, J. Dean. Distributed Representations of Words and Phrases and their Compositionality, https://arxiv.org/abs/1310.4546

Signali

Što je signal?

Koja je razlika između običnog podatka i signala?

Podatke čini više varijabli zajedno koje poprimaju vrijednosti iz nekog skupa. U pravilu ne očekujemo uređeni odnos između varijabli!

Signali su pak **funkcije** (preslikavanja) koje iskazuju kako se zavisna varijabla ponaša ovisno o nezavisnoj varijabli.

Nezavisna varijabla (domena) je parametar o kojem ovisi zavisna varijabla (kodomena).

Nezavisne varijable su često vrijeme ili prostor.

Jednodimenzionalni signali

Do sada ste vidjeli više primjera jednodimenzionalnih signala: vodostaj rijeke Save, redoslijed nukleotida u molekuli DNK, redoslijed boja u nasumičnom izvlačenju šarenih bombona.

Signali su bili diskretni po nezavisnoj varijabli pa ste ih mogli predstaviti kao nizove podataka.

Signal je funkcija ili preslikavanje iz domene u kodomenu.

Domena je $\mathbb{N}$ ili $\mathbb{Z}$ ili neki njihov podskup.

Kodomena je $\mathbb{R}$ ili neki konačan skup vrijednosti (nukleotidi, boje).

Jednodimenzionalni signali

Kada govorimo o kontinuiranim ili o diskretnim signalima moramo dodatno pojasniti je li kontinurana domena ili kodomena.

Pod kontinuranim signalom uobičajeno podrazumijevamo preslikavanje $s : \mathbb{R} \to \mathbb{R}$, odnosno i domena i kodomena su kontinuirane.

Nazivamo ga još i analogni signal.

Pod diskretnim signalom uobičajeno podrazumijevamo preslikavanje $s : \mathbb{Z} \to \mathbb{R}$, odnosno domena je diskretna, a kodomena kontinuirana.

Ako su i domena i kodomena diskretne onda govorimo o digitalnom signalu.

Višedimenzionalni signali

Je li višedimenzionalna domena ili kodomena?

I domena i kodomena mogu biti višedimenzionalne.

Razmotrimo to na primjeru dvije slike:

Slika

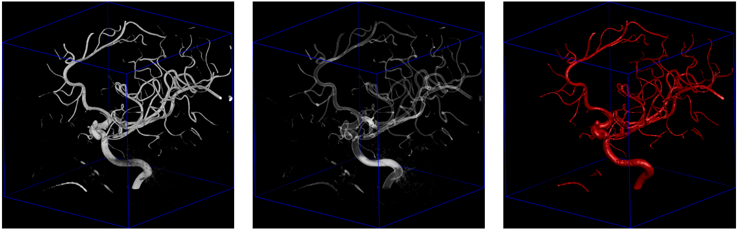
Slika u boji ima dvije nezavisne varijable, prostorne koordinate x i y, te tri zavisne varijable, intenzitet crvene, zelene i plave boje.

Video

Video signal je niz slika u vremenu pa prema tome ima tri nezavisne varijable, dvije prostorne koordinate x i y te vrijeme t.

Prostorni 3D podaci

Prostorni 3D signali imaju tri nezavisne varijable i barem jednu zavisnu varijablu.



Koordinatni sustavi

Česti problem jest kako parametrizirati prostornu domenu, odnosno koji koordinatni sustav se koristi.

Česti izbori su:

1. pravokutni ili Kartezijev koordinatni sustav (x, y, z)
2. sferni koordinatni sustav (r, θ, ϕ)
3. cilindrični koordinatni sustav (ρ, ϕ, z)

Neki izbori su bolji od ostalih za pojedine primjene.

Koordinatni sustavi

Koji koordinatni sustav je prirodan za prikaz podataka?

To je otvoreno pitanje čiji odgovor ovisi o podacima i o primjeni.

Analiza glavnih komponenti (eng. *prinicpal component analysis*, PCA) nam pomaže u konstrukciji jednog takvog koordinatnog sustava.

Binokularni vid

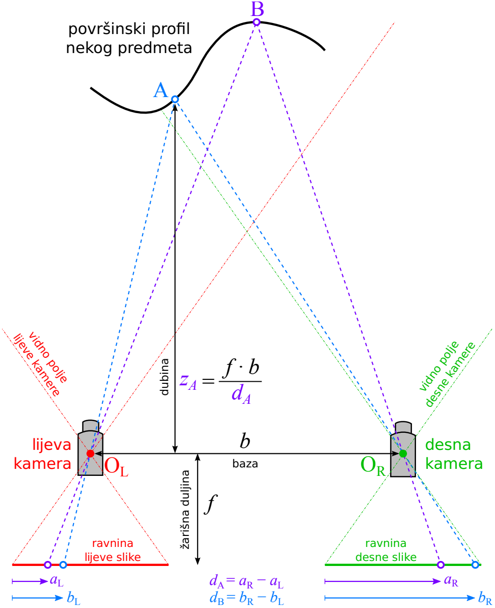
Što je to binokularni ili stereskopski vid?

Dvije kamere mjere dubinu.

Koraci obrade:

1. uklanjanje distorzije;
2. stereo rektifikacija;
3. nalaženje korespondentnih točaka;
4. triangulacija.

Nedostatak: ne radi za površine bez teksture.



2D ili 3D podaci

Binokularni ili steroskopski vid nam daje informaciju o dubini pojedine točke slike.

Prema tome kada govorimo o binokularnom vidu signal koje dobivamo nije samo čista 2D slika, no nije niti volumni 3D signal.

Za svaku točku slike znamo i dubinu (udaljenost točke od kamere), no ta dubina (ni)je nezavisna varijabla (zbog epipolarnog ograničenja) te stoga ne možemo takav signal zvati 3D signalom.

Ponekad se koristi naziv $2\frac{1}{2}$D ili 2.5D.

Obrada podataka

Kako obrađujemo podatke?

$$u \longrightarrow \boxed{\text{obrada podataka}} \longrightarrow y = f[u]$$

Sustav obrade podataka možemo promatrati kao crnu kutiju koja iz ulaznog signala u daje izlazni signal y.

Govorimo o ulazno-izlaznom modelu sustava i preslikavanju f između signala:

$$y(t) = f\big[u(t)\big] \quad \text{ili} \quad y[n] = f\big[u[n]\big]$$

Preslikavanje f obavlja obradu koju želimo. No postoje li neka **poželjna** svojstva koja preslikavanje f mora zadovoljiti?

Poželjna svojstva sustava

Ima li smisla modificirati način obrade signala kroz vrijeme za 1D signale koje ste do sada vidjeli, npr. vodostaj?

Ako nema onda preslikavanje f mora biti **vremenski nepromjenjivo**.

Je li poželjno da sustav obrade signala bude invarijantan na odabir kodiranja kategoričkog signala?

Ako jest onda MAFFT uz odabir kodiranja $A \mapsto 1$, $T \mapsto 2$, $G \mapsto 3$ i $C \mapsto 4$ nije dobar jer je korelacija $C - C$ značajno izraženija od korelacije $A - A$.

Simetrije i invarijante

Prema tome neka od svojstava funkcije obrade f su izuzetno važna.

Primjer iz klasične fizike: Neka je ulazni signal x sila na neko tijelo te neka je izlazni signal y trajektorija tog tijela. Sustav opisan funkcijom obrade f mora poštovati fizikalne zakone koji su neovisni od izbora koordinatnog sustava te koji moraju biti konzistentni obzirom na promjene jedinica mjere.

U fizici općenito govorimo o bažardnim simetrijama koje su povezane s invarijanostima fizikalnih zakona te nas zanimju npr. vrijednosti koje su invarijantne obzirom na promjene koordinatnih sustava.

Slično tome u obradi informacija govorimo o **invarijanosti** sustava obrade obizrom na određenu **manipulaciju** ulaznih podataka.

Vremenska promjenjivost/nepromjenjivost

$$u(t) \longrightarrow \boxed{f} \longrightarrow y(t) = f\big[u(t)\big]$$

Sustav je vremenski nepromjenjiv (eng. *time-invariant*) ako:

$$\forall\big(u(t), y(t)\big)\ \forall t_0 \in \mathbb{R} : f\big[u(t + t_0)\big] = y(t + t_0)$$

Riječima: Sustav je vremenski nepromjenjiv ako za svaki ulazno-izlazni par signala $u(t)$ i $y(t)$ te za svaki vremenski pomak vrijedi da je odziv sustava na pomaknutu pobudu $u(t + t_0)$ jednak pomaknutom odzivu $y(t + t_0)$.

Napomena: Jedan od smjerova pomaka signala nije moguć u praksi zbog kauzalnosti. Koji?

Prostorna promjenjivost/nepromjenjivost

$$u(x) \longrightarrow \boxed{f} \longrightarrow y(x) = f\left[u(x)\right]$$

Sustav je prostorno nepromjenjiv (eng. *time-invariant*) ako:

$$\forall\big(u(x), y(x)\big)\ \forall x_0 \in \mathbb{R} : f\left[u(x + x_0)\right] = y(x + x_0)$$

Riječima: Sustav je prostorno nepromjenjiv ako za svaki ulazno-izlazni par signala $u(x)$ i $y(x)$ te za svaki prostorni pomak vrijedi da je odziv sustava na pomaknutu pobudu $u(x + x_0)$ jednak pomaknutom odzivu $y(x + x_0)$.

Kauzalnost/nekauzalnost

Sustav obrade može biti **kauzalan** i **nekauzalan**.

Kod obrade vremenskih signala u stvarnom vremenu možemo koristiti samo prošle vrijednosti signala, odnosno ograničeni smo smjerom vremena. U tom slučaju obradu ograničavamo na kauzalne sustave.

Sustav je kauzalan ako:

$$\forall t_0 \ \forall y(t_0); \nexists t > t_0 : y(t_0) = f\left[\cdots, u(t), \cdots\right]$$

Napomena: Kauzalnost nije ograničavajuća ako je nezavisna varijabla prostor. U slici možemo očitati vrijednosti lijevo, desno, gore i dolje od trenutne. U nizu nukleotida također možemo očitati prethodne i sljedeće nukleotide.

Izotropnost/anizotropnost

Kada obrađujemo 2D, 3D i višedimenzionalne signale želimo li da obrada ovisi o prostornom smjeru?

Ako ne želimo onda govorimo o **izotropnosti**, a ako želimo onda govorimo o **anizotropnost**.

Izotropni sustavi ne preferiraju neki prostorni smjer u odnosu na sve ostale.

Anizotropni sustavi imaju preferirani smjer.

Je li ljudski vizualni sustav anizotropan?

Je li ljudski vizualni sustav anizotropan?

Je li ljudski vizualni sustav anizotropan?

Je li ljudski vizualni sustav anizotropan?

P. Thompson. “Margaret Thatcher: A New Illusion.” Perception 9 (1980): 483-484.

`https://doi.org/10.1068/p090483`

Invarijantnosti sustava obrade slike

Želimo da sustav obrade slike bude invarijantan na određene transformacije slike.

Neke transformacije od interesa su:

1. translacija
2. rotacija
3. refleksija
4. skaliranje

Koje od njih su važne?

Kako izgleda translacija i rotacija ravnine?

Invarijantnosti sustava obrade slike

Slike koje obrađujemo su snimljene kamerom koja 3D svijet preslikava u 2D ravninu.

To preslikavanje najčešće opisujemo preko modela kamere s točkastim otvorom (eng. *pinhole camera model*).

Kada govorimo o invarijantnostima zanima nas invarijantnost na translaciju i rotaciju u 3D prostoru te na skaliranje koje se javlja kao posljedica odabira skale prilikom preslikavanja i 3D svijeta u 2D ravninu.

Invarijatnost na translaciju i rotaciju u 2D ravnini je prihvatljiva samo za male pomake.

Invarijantnost i konvolucijske neuronske mreže

Konvolucijske neuronske mreže sastoje se od konvolucijski slojeva.

Operacija 2D konvolucije je

$$b(x,y) = a(x,y) ** h(x,y) = \sum_{\chi} \sum_{\nu} a(\chi,\nu) h(x-\chi, y-\nu),$$

gdje je $a(x,y)$ ulaz, $b(x,y)$ izlaz, te $h(x,y)$ jezgra konvolucije (ili impulsni odziv sustava).

Operacija 2D konvolucija je invarijatna na pomak, odnosno vrijedi

$$a(x+x_0, y+y_0) ** h(x,y) = b(x+x_0, y+y_0).$$

Prema tome konvolucijske neuronske mreže su prirodno neosjetljive na 2D translaciju.
Što s 2D rotacijom? I što s 3D translacijom i rotacijom?

Obrada skupova

Zamislimo sustav obrade podataka koji radi na skupovima.

Ulaz u sustav je skup $\mathrm{u} = \{u_1, u_2, \ldots, u_n\}$ od n elemenata u kojem je redoslijed nevažan. Prema tome ne obrađujemo signale već obrađujemo nepovezane varijable.

Jedan ulazno/izlazni model za koji je ulaz skup, a izlaz jedna vrijednost je:

$$\mathrm{u} = \{u_1, u_2, \ldots, u_n\} \longrightarrow \boxed{f} \longrightarrow y = f\left[u_1, u_2, \ldots, u_n\right]$$

Želimo da sustav odnosno njegova funkcija f daje isti y bez obzira na redoslijed ulaza.

Permutacijska invarijantnost

Na primjer za $n = 3$ treba biti:

$$f[u_1, u_2, u_3] = f[u_1, u_3, u_2] = f[u_2, u_1, u_3] =$$
$$= f[u_2, u_3, u_1] = f[u_3, u_1, u_2] = f[u_3, u_2, u_1]$$

Prema tome za obradu skupova nužno svojstvo funkcije f jest neosjetljivost na redoslijed ili permutacijska invarijantnost (eng. *permutation invariance*).

Ako s π označimu funkciju permutacije skupa onda općenito mora vrijediti

$$f\left[\pi(u_1, u_2, \ldots, u_n)\right] = f\left[u_1, u_2, \ldots, u_n\right],$$

odnosno za svaku permutaciju π ulaza sustav mora dati isti odziv.

Problem reprezentacije

Sada znamo da treba vrijediti $f\left[\pi(\mathbf{u})\right] = f\left[\mathbf{u}\right]$, no kako općenito realizirati funkciju koja to zadovoljava?

To je općeniti problem reprezentacije. Pokušavamo odgovoriti na pitanje postoji li neko (jednostavno) razlaganje funkcije f koje bi nam pojednostavnilo njenu konstrukciju?

Pokazuje se da je jedan mogući rastav

$$f = \rho\Big(\sum\nolimits_{u_i} \phi(u_i)\Big),$$

gdje su ρ i ϕ odgovarajuća preslikavanja (transformacije).

M. Zaheer et al. “Deep sets.” Advances in neural information processing systems 30 (2017): 3391-3401. `https://arxiv.org/abs/1703.06114`