

Fakulta riadenia a informatiky, Žilinská univerzita v Žiline
Fakulta prírodných vied, Univerzita Mateja Bela,
Banská Bystrica

Analýza procesov lineárnymi metódami

Ondrej Šuch, Martin Klimo, Ivan Cimrák, Katarína Bachratá



ŽILINSKÁ UNIVERZITA V ŽILINE, Fakulta riadenia a informatiky

UNIVERZITA MATEJA BELA, BANSKÁ BYSTRICA, Fakulta prírodných vied

Analýza procesov lineárnymi metódami

Autori:

Ondrej Šuch, PhD.

prof. Ing. Martin Klimo, PhD.

Dr. Ivan Cimrák

doc. RNDr. Katarína Bachratá, PhD.

Vedecký redaktor:

prof. RNDr. Ivo Čáp, CSc.

Recenzovali:

doc. RNDr. Elena Wisztová, PhD.

prof. RNDr. Ladislav Kvasz, Dr.

Vydala Žilinská univerzita v Žiline, Žilina 2012

ISBN:

© Ondrej Šuch, Martin Klimo, Ivan Cimrák, Katarína Bachratá

Úvod

Čo je proces?

Každý z nás má skúsenosť, že sa naše okolie mení. Od týchto zmien je odvodené aj naše chápanie času. Človek sa však neuspokojuje s tým, že zmeny v jeho okolí len konštatuje, ale snaží sa poznať príčiny ich vzniku, predpovedať ich následky v budúcnosti, ba dokonca ich ovplyvňovať. Človek najčastejšie registruje zmeny v jeho okolí priamo, pomocou zmyslov. Tým informačne najbohatším zmyslom je zrak a v súčasnosti vidíme veľké úsilie o spracovanie, prenos a uchovanie zrakových vnemov pomocou techniky. Autori tejto učebnice sa venujú druhému informačne bohatému zmyslu - sluchu, resp. jednému jeho účelu, ktorým je hlasová komunikácia. Dnes sa však človek nezaujíma len o zmeny, ktoré sú priamo postrehnuteľné zmyslami. Sú to zmeny, ktorých rýchlosť je mimo dosahu našich zmyslov - sú pre ne príliš pomalé, alebo príliš rýchle, prípadne naše zmysly ich nie sú schopné vôbec prijímať. Vnímame ich len sprostredkované pomocou záznamových a meracích prístrojov. Všimol si niekto z vás, že erupcie na Slnku sa vyskytujú periodicky s periódou približne 13 rokov? Počul niekto lokačné signály netopierov, ktoré okolo vás preleteli? Dnes však predsa o týchto dejoch vieme, pretože človek zostrojil prístroje, schopné tieto zmeny indikovať a zaznamenávať. Dokonca zaznamenávame aj zmeny, ktoré spôsobuje človek svojou činnosťou. Vieme, koľko pribudlo v minulom roku železa vyrobeného na Slovensku, koľko bolo vydaných kníh, koľko študentov sa zapísalo na vysokoškolské štúdium. Človeka zaujímajú zmeny najmä ekonomické, sociálne a kultúrne, hoci ich nevnímame priamo zmyslami. Určite ste si všimli, že uvádzame príklady len takých zmien, čo vieme odmerať. Zmeny, ktoré nevieme kvantifikovať sú síce tiež dôležité (a možno najdôležitejšie pre zachovanie ľudstva), predsa však nebudú obsahom tejto učebnice. Môžeme teda povedať, že predmetom tejto učebnice budú zmeny, pri ktorých vieme povedať, v akej sú súvislosti voči nejakej referenčnej periodicky sa opakujúcej zmene (napr. periodickej zmene atómov plynového amoniaku na vlnovej dĺžke 1,2599 cm), čiže hodín. Týmto hodinami vytvárame hodnotu času, v ktorom zmena nastala a akým číslom sa dá veľkosť zmeny ohodnotiť.

Okrem toho, že vnímame zmeny merateľných veličín, máme skúsenosť, že pri rovnako vnímanej fyzikálnej veličine dochádza k rovnakým kauzálnym zmenám, napr. voda začne vriieť, alebo zamrzne. Preto zavedieme určitú referenčnú hodnotu nielen pre veľkosť zmeny, ale aj pre samotnú veľkosť veličiny (čo je samozrejme veľkosť jej zmeny od referenčnej hodnoty). Napr. Anders Celsius zaviedol ako referenčnú hodnotu teploty takú teplotu, pri ktorej zamrzá voda a na meranie zmeny teploty určil

druhú referenčnú hodnotu - teplotu, pri ktorej voda vrie (všetko pri ďalších definovaných podmienkach pokusu). Dnes tieto hodnoty poznáme ako 0°C a 100°C . Iné spôsoby merania teploty zaviedli Fahrenheit a Réamour a aby to nebolo jednoduché, ďalší fyzici si zvolili ako referenčnú hodnotu tú najnižšiu teplotu, ktorá môže v prírode byť. Prečo si teda aj Vy nezvolíte svoju stupnicu pre vyjadrenie veľkosti teploty? Zrejme preto, že už nejaké máme a jednak preto, že okrem hodnoty teploty by ste museli vysvetľovať aj Vašu stupnicu (nám Európanom stačia problémy ak nám povedia teplotu vo Fahrenheitoch). Pretože však budeme pracovať s rôznymi fyzikálnymi aj nefyzikálnymi veličinami a ich zmenami, bude vhodné, ak si neskôr povieme, aké spoločné zásady pre tvorbu systému merania dodržiavať, aby sme s hodnotami veličiny mohli pracovať rovnako intuitívne, ako pracujeme so stupňom Celsia, metrom alebo sekundou. Možno práve Vy budete potrebovať vytvoriť novú veličinu (zrejme nefyzikálnu) a navrhnúť spôsob merania jej veľkosti.

Ako sme uviedli, poznaniu príčin a následkov zmien okolo nás sa venuje veda dlhodobo, preto je prirodzené, že sa špecializovala podľa prostredia, v ktorom sa tieto zmeny dejú. Zmeny v prírode študuje najmä fyzika, chémia, biológia, máme vedy o človeku ako medicína alebo psychológia a máme tiež ekonomické a spoločenské vedy. Vašu pozornosť ako študentov technickej vysokej školy by sme radi sústredili na technické vedy, ktoré upriamujú pozornosť aj na to, ako takéto zmeny riadiť alebo vytvárať pomocou umelo vytvorených zariadení (napr. počítačmi). Táto učebnica však v žiadnom prípade nechce konkurovať učebniciam fyziky, medicíny, ekonómie či programovaniu. Chce len upozorniť na to, že určité zákonitosti v zmenách sledovaného okolia môžeme odvodiť len zo samotného záznamu veľkosti týchto zmien bez ohľadu na to, čo je príčinou ich vzniku. V tomto zmysle je táto učebnica pomocníkom pri štúdiu fyziky, medicíny, atď., nie ich konkurentom, či dokonca náhradou. Je prirodzené, že ak budeme abstrahovať od príčiny vzniku zmien a budeme analyzovať len postupnosť zaznamenaných číselných hodnôt, tak hlavným pomocníkom pri tomto štúdiu bude matematika. Pri výklade sme sa snažili o to, aby sme používali čo najjednoduchšie matematické nástroje. Najskôr sú použité len tie, s ktorými máte skúsenosti zo strednej školy a základného kurzu matematiky a až neskôr sú zovšeobecnené tak, aby ste pochopili ako analýzu zmien v jednom prostredí možno bez podstatných zmien použiť aj v prostredí inom, napr. úplne novom.

Ako robiť analýzu procesov?

Predmetom skúmania tejto učebnice je súbor hodnôt niektorej merateľnej veličiny tak, ako boli zaznamenané v určitých časových okamihoch. Tento súbor hodnôt, spolu so zodpovedajúcimi časovými okamihmi budeme nazývať **proces**. Aby zreteľnejšie vynikli základné myšlienkové postupy pri analýze procesov, zameriame sa najmä na súbory, ktoré obsahujú konečný počet nameraných hodnôt. Pozície nameraných hodnôt procesu nie sú navzájom zameniteľné, pretože každá sa viaže k určitému časovému okamžiku. Príklad na trojicu čísel, ktorých hodnoty sú viazané na pozíciu v zázname poznáte - je to bod v priestore, ktorého poloha je daná tromi súradnicami. A vlastnosti bodov v priestore poznáme nielen z osobnej skúsenosti pri pohybe v priestore, ale aj zo štúdia geometrie. Základnou myšlienkou tejto učebnice je použiť predstavy,

ktoré máme o trojrozmernom priestore pri analýze procesov. Nič nám nebráni v tom, aby sme interpretovali časovú súradnicu procesu ako súradnicu bodu a aby sme si proces s tromi zaznamenanými hodnotami predstavili ako bod v priestore. Pretože však málokedy potrebujeme poznať vlastnosti procesu len s tromi zaznamenanými hodnotami, ako prvý krok zovšeobecnenia zavedieme viacrozmerný priestor. Len pre Vašu predstavu: 1 sekunda jednej stopy záznamu zvuku na Vašom CD nosiči obsahuje 44 100 hodnôt o intenzite akustického tlaku snímaného dvomi mikrofónmi pri nahrávke. Takýto zvuk si potom môžeme predstaviť ako jeden bod v 44 100-rozmernom priestore. Asi ťažko si takýto priestor predstavíme, ale všetko, čo budeme s týmto bodom robiť sa nebude odlišovať od toho, čo robíme s bodom v trojrozmernom priestore.

Aby sme mohli posudzovať súvislosti medzi polohami bodov v priestore, zavedieme pojem vektora ako objektu, ktorý má koniec v danom bode priestoru a začiatok v počiatku súradnicového systému, tj. v bode s nulovými súradnicami. Takáto predstava pomáha nielen vo fyzike, ale aj v geometrii umožňuje zaviesť pojem uhla medzi vektormi, veľkosť vektora, násobok vektora, súčet vektorov a podobne. Pojem vektora, resp. vektorového priestoru bude základným nástrojom pre analýzu procesov v tejto učebnici. Podotknime ešte, že okrem konečnorozmernosti ohraničíme naše úvahy na priestory lineárne. Intuitívne si môžeme takýto priestor predstaviť ako priestor všade rovnako hustý, v ktorom platí priama úmera, tj. že súčet dvoch rovnakých vektorov je rovnaký, ako dvojnásobok tohto vektora. Samozrejme, že tieto pojmy vymedzíme v učebnici presnejšie.

Ako nám však predstava o vlastnostiach bodov v priestore môže pomôcť pri poznaní vlastností procesu? Proces, v ktorom medzi hodnotami neexistuje žiadna súvislosť, zodpovedá bodu, ktorý sa s rovnakou možnosťou nachádza na hociktorom mieste priestoru. Ak však bude medzi hodnotami procesu určitá závislosť, zodpovedajúci bod bude ležať v blízkosti niektorej roviny, priamky, či dokonca bodu v priestore, medzi ktorými je rovnaký typ závislosti. A ak zistíme, že bod zodpovedajúci procesu leží v blízkosti určitej priamky so zadanými vlastnosťami, potom na približné určenie jeho polohy v priestore postačuje jediné číslo, ktoré určuje polohu najbližšieho bodu na priamke. Vo všeobecnosti sa dá povedať, že medzi hodnotami procesu sme našli určitú zákonitosť, ak na „akceptovateľný“ popis procesu postačí podstatne menej čísel (parametrov modelu zákonitosti), než je jeho zaznamenaných hodnôt. Napríklad ak by procesom bol záznam 1 sekundy už spomínaného zvuku a za akceptovateľný by sme považovali tento zvuk tak, ako by sme ho počuli cez mobilný telefón (GSM enhanced full rate), tak potom na jeho popis miesto 44 100 hodnôt postačuje 940 čísel.

Našou metódou pri hľadaní zákonitostí v procese bude otáčanie priestoru tak, aby sme uvideli, k akej rovine, priamke, či bodu s vopred zadanými vlastnosťami sa proces blíži. Zistíme napríklad, či proces s tromi hodnotami nemá tú vlastnosť, že všetky jeho hodnoty sú približne rovnaké. Aké miesto bodov vytvárajú v priestore body, ktorých všetky tri súradnice sú rovnaké? Je to priamka (nazvime ju bázickou) prechádzajúca počiatkom súradnicového systému a zvierajúca s každou osou uhol 45° . Ak chceme zistiť, či proces má uvedenú vlastnosť, pozrieme sa na priestor tak, aby nám „prechádzala“ bázická priamka cez sietnicu, takže ju uvidíme len ako „bázický“ bod. Ak takto pozorovaná vzdialenosť medzi „bázickým“ a analyzovaným bodom bude malá, má proces vlastnosť, že všetky jeho hodnoty sú približne rovnaké a analyzovaný proces môžeme približne nahradiť (aproximovať) procesom, ktorého všetky hodnoty

sú rovnaké, tj. bodom, ktorý leží na bázekej priamke. Kde však na nej leží? Postačí, ak zadáme jeho vzdialenosť od počiatku, prípadne jednu jeho súradnicu (pretože ostatné dve sú rovnaké). Poznanie základných vlastností vektorového priestoru umožní toto číslo vypočítať.

Pre koho je učebnica určená a čo obsahuje

Informatické fakulty na technických vysokých školách na Slovensku historicky vznikali z katedier informatiky na elektrotechnických fakultách. V matematickej príprave študentov je vďaka tomu dodnes venovaná dominantná pozornosť matematickej analýze aj v prípade, že je možné riešiť úlohy viacerými spôsobmi. Typickým príkladom je napríklad metóda lineárnej regresie. Metódy lineárnej algebry, pomocou ktorých sa dá regresia interpretovať ako priemet do podprieoru, nie sú študentom na technických informatických fakultách príliš známe. Lineárna regresia je častejšie odvodzovaná ako optimalizačná úloha a riešená metódami diferenciálneho počtu.

V tejto učebnici by sme chceli ukázať, že pomocou metód diskkrétnej matematiky je možné priamo riešiť množstvo problémov, ktoré sa zatiaľ riešia metódami spojitej matematiky, a až potom sa transformujú do diskkrétnej formy, aby mohli byť implementované na počítači. Výlučne diskrétnymi metódami síce nevyriešime všetky úlohy, ale aspoň časť z nich urobíme prístupnejšiu väčšiemu počtu študentov.

Učebnica je určená študentom posledného ročníka bakalárskeho štúdia alebo študentom inžinierskeho stupňa štúdia informatiky na technických vysokých školách. U čitateľov sa predpokladá ovládanie látky zo základných matematických kurzov ako sú lineárna algebra a pravdepodobnosť, prípadne základy diferenciálneho počtu. Táto skladba prerekvizít by podľa predchádzajúcej úvahy nemala nutne obsahovať kalkulus. Chceme však využiť súčasný stav vyučovania matematiky a v učebnici uvádzame niekoľko príkladov na odvodenie tej istej metódy nielen pomocou metód lineárnej algebry ale aj pomocou metód diferenciálneho počtu, ktoré študenti poznajú.

Hlavným cieľom učebnice je ukázať študentom postup, ako vedomosti z matematiky získané v nižších ročníkoch možno použiť na popis metód a algoritmov, ktoré sa skutočne využívajú v technickej praxi.

Prvá kapitola učebnice je motivačná a popisuje (aj keď nie úplne detailne), ako dobré porozumenie matematickej podstaty technickej aplikácie umožňuje urobiť jej vylepšenia pri zmenených technologických podmienkach. Táto časť sa zaoberá riešením jednoduchšieho problému, či proces má vopred zadanú vlastnosť, teda či jemu zodpovedajúci bod sa nachádza v blízkosti zadaného podprieoru a ak áno, tak ktorým procesom ho aproximovať.

Druhá kapitola je venovaná vektorovým priestorom a využitiu vektorov na modelovanie časových radov. Formálne definície a výpočty, ktoré študenti poznajú z matematických predmetov, sú vysvetlené na elementárnych príkladoch. Týmto čitateľ získa vhľad do vlastností a pojmov, ktoré potom použije aj v priestoroch, v ktorých si už tieto objekty predstaviť nedokáže.

Tretia kapitola je venovaná lineárnej regresii a jej interpretácii ako priemetu vektora do podprieoru menšej dimenzie.

V ďalších troch kapitolách sú popísané metódy a postupy využívané v informa-

tike, ktoré je možné popísať teoretickým aparátom z predchádzajúcich kapitol. Ide o metódy využívané pri prenose signálov digitálnymi sieťami, ako sú modulácia, kódovanie, kompresia, optimálny príjem. Ďalej je to metóda kĺzavých súčtov a diskretna Fourierova transformácia.

Posledná kapitola učebnice rieši zložitejší problém: či viacero procesov má nejakú spoločnú vlastnosť a ak áno tak akú? V geometrickej interpretácii to znamená zistiť, či viacero bodov v priestore (predpokladajme, že sú viac než tri), leží v blízkosti nejakej roviny alebo priamky. Ak áno, tak ju nájdime tak, aby po nahradení pôvodných bodov bodmi tejto roviny či priamky sme sa dopustili najmenšej chyby. Samozrejme úlohu rozšírime do viacrozmerného priestoru. My budeme v tejto učebnici predpokladať, že hodnoty procesu (súradnice bodu) sú náhodné, ale poznáme stupeň lineárnej závislosti medzi nimi. To znamená, že budeme potrebovať základné vedomosti z teórie pravdepodobnosti.

Pre Vás, študentov informatiky, je poznanie metód analýzy procesov potrebné nielen preto, aby ste ich vedeli použiť vo vedách ktoré skúmajú príčiny zmien v určitej oblasti, akými sú skôr spomínaná fyzika, ekonómia atď., ale aj preto, aby ste tieto metódy naučili počítače, tj. aby ste ich boli schopní naprogramovať. Metódy, ktoré sú vysvetlené v tejto učebnici (založené na lineárnych priestoroch) nie sú tými, ktoré by boli pri analýze malého objemu dát lepšie, než by to urobil človek na základe skúseností. Dnes je však potrebné nepretržite analyzovať veľké množstvo procesov, veľakrát v nepriaznivých podmienkach a s krátkou reakčnou dobou. Za týchto okolností je lepšie pri analýze procesov nahradiť človeka počítačom.

Veríme, že pochopenie princípov analýzy procesov pomocou jednoduchých lineárnych modelov Vám pomôže prípadne aj neskôr, ak ich budete potrebovať aplikovať v nových oblastiach, alebo ak s ich výsledkami už nebudete spokojní a siahnete po nelineárnych modeloch.

História učebnice

Štúdium matematických modelov fyzikálnych procesov patrí k základným krokom pochopenia prenosu informácie. Preto vždy bolo súčasťou predmetu *Teória oznamovania*, ktorého hlavným predstaviteľom nielen v Žiline ale na Slovensku bol prof. Kroutl. V tom čase (1960 – 1977) to boli najmä spojité signály, ktoré boli nositeľom informácie v telefónnych systémoch, preto hlavným nástrojom na ich štúdium bola matematická analýza. S príchodom digitálnych technológií sa digitalizuje prenos hlasu a objavuje sa služba prenosu dát ako nový fenomén. Na Katedre technickej kybernetiky, ktorá bola predchodcom Fakulty riadenia a informatiky na Žilinskej univerzite sa od školského roku 1979/80 začal učiť predmet *Teória informácie a prenos dát*. V rámci vyučovania tohto predmetu bol zvolený netradičný prístup k vyučovaniu modelovania informačných procesov založený na algebraických vlastnostiach signálov. Zatiaľ čo klasický postup vyučovania sa sústredil na spojité signály a modely diskretných signálov vyvodzoval z nich, algebraický prístup umožnil vysvetliť základné myšlienky modelovania signálov na jednoduchých signáloch s diskretným časom a neskôr ich zovšeobecniť aj na signály so spojitým časom. Po založení Fakulty riadenia a informatiky a vzniku študijného odboru *Informačné a riadiace systémy* bol od šk. roku

1993/94 vyučovaný povinný predmet *Analýza procesov*. Cieľom predmetu bolo naučiť študentov tvoriť základné modely procesov z nameraných dát bez ohľadu na to, čo dáta reprezentujú a aká je časová mierka. Preto mohol byť tento predmet spoločný pre zamerania *Aplikovaná informatika*, *Manažment* aj pre zameranie *Informačná a riadiaca technika*. Predmet tak vhodne dopĺňal špecializované predmety, v ktorých sa študent dozvedel, čo dáta reprezentujú a aké ďalšie poznatky o nich vieme. Text tejto učebnice odzrkadľuje naše mnohoročné skúsenosti s vyučovaním predmetu *Analýza procesov*.

Obsah

Úvod	i
1 Teoretické modely a ich využitie v technike	1
1.1 Počítačové spracovanie reči	1
1.2 Spracovanie signálu pre prenos	2
1.3 Výpočet koeficientov vektora v rôznych bázach	4
1.4 Nekonečnorozmerné vektorové priestory	7
1.5 Využitie pre IP telefóniu	11
2 Procesy a dáta ako prvky vektorových priestorov	13
2.1 Vektorové priestory	13
2.2 Báza a dimenzia vektorového priestoru	23
2.3 Súradnice vektora vzhľadom na bázu	26
2.4 Zmena bázy	27
2.5 Skalárny súčin	30
2.6 Koeficienty vektora v ortogonálnej báze	32
2.7 Norma vektora	33
2.8 Kolmý priemet vektora do podpriestoru	36
2.9 Gramova-Schmidtova ortogonalizácia	38
3 Modely využívajúce závislosť hodnôt procesu od času	43
3.1 Vektor ako časová závislosť	43
3.2 Aproximácia vektora v podpriestore	44
3.3 Metóda najmenších štvorcov	47
3.4 Chyba aproximácie	48
3.5 Priamka rovnobežná s časovou osou, $y = c$	50
3.6 Priamka prechádzajúca počiatkom, $y = c \cdot t$	52
3.7 Priamka neprechádzajúca počiatkom, $y = c_0 + c_1 \cdot t$	53
3.8 Parabola, $y = c_0 + c_1 \cdot t + c_2 \cdot t^2$	55
3.9 Lineárna regresia ako priemet do podpriestoru	59
3.10 Všeobecný zápis lineárnej regresie	60
3.11 Koeficienty v neortogonálnej báze	67
3.12 Linearizácia nelineárnych úloh	71

4	Modely využívajúce súvislosti medzi hodnotami procesu	77
4.1	Modely kľzavých súčtov	77
4.2	Všeobecná formulácia úlohy	84
5	Periodické vlastnosti procesov	91
5.1	Periodické modely	91
5.2	Harmonická báza	98
5.3	Diskrétna Fourierova transformácia (DFT)	103
5.4	Priemet do podpriestoru určeného DFT	106
5.5	Walsh-Hadamardova báza	119
5.6	Použitie Fourierovej transformácie v praxi	123
6	Analýza viacrozmerných objektov	125
6.1	Dôležité súradnice	126
6.2	Váhové koeficienty pre maximálny rozptyl	129
6.3	Analýza hlavných komponentov	136
6.4	Náhodné procesy	145
6.5	Náhodné vektory	147
6.6	Karhunen Loëvov rozklad	152
7	Lineárne systémy	159
7.1	Lineárny systém	160
7.2	Modulácia	161
7.3	Optimálny prijímač	162
7.4	Lineárne časovo invariantné transformácie	164
7.5	Spektrálny popis lineárnych, časovo invariantných transformácií	165
7.6	Gradientová metóda	167
7.7	Metóda adaptívneho hľadania koeficientov	169
8	Diskrétné lineárne systémy	171
8.1	Detekcia parity	171
8.2	Ochrana proti reverzii polarizácie	173
8.3	Hammingov (7,4) kód	174
8.4	Konvolučné kódy	179
8.5	Polynomiálne kódy	182
9	Modelovanie náhodnosti v procesoch	187
9.1	Modelovanie lineárnej regresie	188
9.2	Stacionárny proces	193
9.3	Využitie stacionarity na predikciu a interpoláciu	195
9.4	Biely šum	199
9.5	$MA(q)$ procesy	200
9.6	$AR(p)$ procesy	203
9.7	Yule-Walkerove rovnice	204
9.8	Durbin-Levinsonova rekurzia	206
9.9	Woldova dekompozícia	208

10 Matematický softvér	211
10.1 Matlab	212
10.2 R	216
10.3 Mathematica	221
Literatúra	225
Register	227

1 Teoretické modely a ich využitie v technike

Začnime citátom z knihy [20], v ktorej je milo popísaný rozdiel medzi vedou a technikou: „Technika je o vyrábaní vecí, bez toho, že by im človek musel rozumieť. Veda je o porozumení veciam bez toho, aby ich človek musel vyrábať.“ Cieľom tejto kapitoly nie je ponúknuť ucelenú teóriu, ale načrtnúť problém, ktorý bol riešený v spolupráci medzi vedou a technikou. Na riešení konkrétnej úlohy si ukážeme, že je užitočné rozumieť aj matematickému modelu, aj tomu, aké sú technické prostriedky, ktorými je model realizovaný. Strety medzi technickou a matematickou koncepciou pri riešení problémov vedú často k nedorozumeniam, prípadne zaznávaníu inžinierov matematikmi („lebo sú povrchní“) aj matematikov inžiniermi („lebo dokazujú úplné samozrejmosti a navyše vnucujú svoje dokazovanie aj 'normálnym' ľuďom“). V už citovanej knihe nájdeme aj ďalšiu vetu: „Skutočná veda vznikla až vtedy, keď si ľudia uvedomili, že teória a pokusy idú ruka v ruku a že táto kombinácia je účinný spôsob ako vyriešiť hromadu problémov a objaviť zaujímavé nové problémy.“

V tejto kapitole netreba všetkému rozumieť. Je to ilustratívna a motivačná kapitola a preto čitateľ nemusí hľbať nad detailami. Tie budú vysvetlené v nasledujúcich kapitolách.

Nasledujúca ukážka ilustruje ako hlbšie matematické pochopenie už existujúceho riešenia môže ukázať cestu novým technickým riešeniam. V situácii, keď sa zmenia podmienky, je možné model upraviť. Budeme sa zaoberať počítačovým spracovaním reči, konkrétne IP (Internet Protocol) telefóniou. V klasickej telefónii je signál spracovaný tzv. A/D (analogovo/digitálnou) konverziou, pri ktorej sa z analogového signálu vytvorí digitálny, vhodný na prenos. Dá sa dokázať, že v súčasnosti používaný A/D prevod je najlepší možný. Ale najlepší možný len pre telefóniu cez telefónne siete. Dátové siete (internet) používajú na prenos inú technológiu, pri ktorej vznikajú iné problémy ako pri klasickej telefónii. Navrhne novú spôsob A/D konverzie, pri ktorom bude zabezpečená väčšia odolnosť prenášaného obsahu (v našom prípade rečového signálu) voči stratám paketov, ktoré vznikajú pri prenose signálu IP sieťou. Tým sa dosiahne vyššia kvalita prijateľného telefónneho signálu.

1.1 Počítačové spracovanie reči

Počítačové spracovanie reči umožňuje hlasovú komunikáciu medzi dvoma ľuďmi pomocou počítača alebo medzi počítačom a človekom. Podľa toho, ako počítač reč spracuje, rozlišujeme tri hlavné úlohy spracovania reči počítačom: rozpoznávanie, syntézu a prenos reči sieťou.

Rozpoznávanie reči je proces, pri ktorom počítač identifikuje hovorené slová a reaguje na ne. Teda na rovnaký hovorený podnet dostaneme od počítača vždy rovnakú

odozvu, napríklad sa na monitore vypíše príslušný text. Iný výstup môže byť identifikácia a verifikácia osôb, bližšie sa o rozpoznávaní reči dá dočítať napríklad v [12], [13].

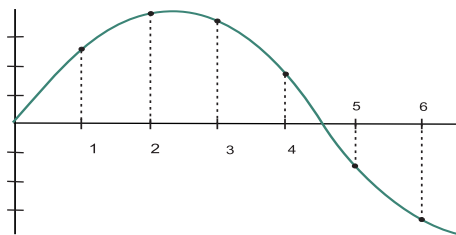
Opačným spôsobom spracovania je syntéza reči, ktorá spočíva vo vytvorení zvukového prepisu textu, ktorý nebol vopred v počítači uložený. Pre syntézu sa používajú dva postupy: buď modelovanie vokálneho traktu človeka, alebo skladanie zvuku z vopred nahovorených rečových jednotiek. Teda zo súboru nahovorených foném (resp. difém) uložených v databáze sa poskladá slovo, ktoré zaznie z reproduktora. Jeden z funkčných syntetizátorov vytvorených pre slovenský jazyk sa dá vyskúšať na internetovej adrese <http://tts.kis.fri.uniza.sk>. Tento syntetizátor vyvinula skupina pracovníkov a doktorandov katedry KIS FRI ŽU. Popis práce tohto syntetizátora a vysvetlenie ako funguje je možné nájsť napríklad v článkoch [6], [15].

Prenos hlasu telefónnou sieťou používa principiálne inú technológiu, ako je prenos zdigitalizovaného hlasu cez dátové siete. Rozdiel je v spôsobe fungovania prenosových mechanizmov telefónnej siete a dátových sietí. Telefónna sieť funguje na princípe prepájania okruhov a dátové siete fungujú na princípe prepájania paketov. To znamená, že pre klasický telefónny hovor je vytvorený samostatný okruh po celú dobu hovoru. Pre hovor cez IP siete sú krátke segmenty reči balené do paketov a prenášané sieťou. Mechanizmus prepájania okruhov, používaný v telefónnych sieťach je prispôbený rovnomernému prenosu a dokáže mu garantovať požadovanú prenosovú kapacitu, rýchlosť a pravidelnosť doručovania. Tento spôsob fungovania telefónnych sietí je však podstatne drahší ako prenos reči po dátových sieťach. V dátových sieťach (napr. v sieti internet) je používaný princíp prepájania paketov. Pri prepájaní paketov je snaha čo najefektívnejšie využiť celkovú dostupnú prenosovú kapacitu. Na druhej strane v dátových sieťach nie je garantovaný čas doručenia dát alebo ich maximálne meškanie. Je však zrejmá ekonomická výhodnosť prenosu telefónnych hovorov po dátovej sieti, pretože užívateľ platí len za pripojenie k internetu, ale neplatí za dĺžku hovoru alebo za vzdialenosť, na ktorú je hovor prenášaný. Presnejšie informácie o prenose reči internetom nájde čitateľ napríklad v práci [16].

1.2 Spracovanie signálu pre prenos

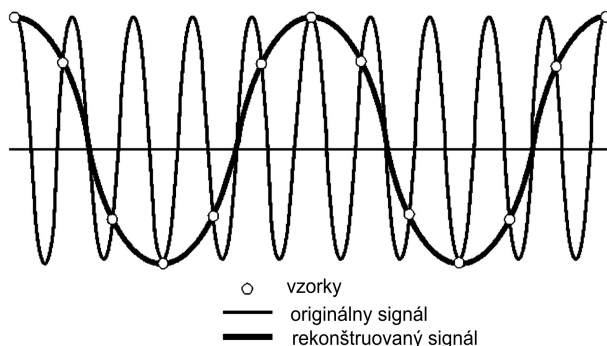
Pre prenos od vysielača k prijímaču je potrebné zmeniť analógový rečový signál na digitálny (A/D konverzia). V súčasnosti sa A/D konverzia pre klasickú aj IP telefóniu realizuje rovnakým spôsobom. V prvom kroku sa nasnímajú (odmerajú) hodnoty rečového signálu vo vopred určených časových okamihoch. Čím sú tieto okamihy k sebe bližšie, tým je signál popísaný presnejšie. Preto je jedným z parametrov určujúcich kvalitu zvuku vzorkovacia frekvencia, ktorá udáva ako často sa meria hodnota vstupného signálu. Pre telefóniu sa používa napríklad frekvencia 8 kHz, čo znamená, že za 1 sekundu je odmeraných 8000 hodnôt signálu (vzoriek). Hudobné záznamy v CD kvalite používajú frekvenciu 44,1 kHz. Vzorkovanie signálu je znázornené na obrázku 1.1.

Veľkosť intervalu medzi jednotlivými snímanými vzorkami určuje Shannonova vzorkovacia veta, ktorá hovorí, že vzorkovacia frekvencia má byť aspoň dvojnásobkom najvyššej frekvencie obsiahnutej v signáli. Ľudské ucho počuje frekvencie od 20



Obr. 1.1. Spracovanie signálu - vzorkovanie

Hz do 20 kHz, odtiaľ sa odvodilo vzorkovanie 44,1 kHz ako približne dvojnásobok najvyššej frekvencie počuteľnej človekom. Pri nedodržaní pravidla o veľkosti vzorkovacej frekvencie môže byť zrekonštruovaný signál úplne odlišný od vstupného signálu, ako vidíme na obrázku 1.2. Tento problém sa nazýva aliasing.



Obr. 1.2. Vznik aliasingu pri vzorkovaní s nesprávnou vzorkovacou frekvenciou

Namerané hodnoty signálu sú bodmi nespočítateľnej množiny, na ich vyjadrenie je však k dispozícii iba konečný počet čísel. Preto je potrebné hodnoty signálu v druhom kroku upraviť (zaokrúhliť) tak, aby množina hodnôt signálu mala len konečne veľa prvkov. Výsledné hodnoty sa nazývajú kvantovacie hladiny a čím je ich viac, tým presnejšie je signál zaznamenaný. Zvyčajne sa v telefónii používa na zakódovanie jednej hodnoty 8 binárnych čísel, presnejšie 255 hodnôt (hladín, úrovní), 128 so znamienkom „+“, 128 so znamienkom „-“, pričom 0 je v oboch skupinách. Zaokrúhľovanie v telefónii sa realizuje nerovnomerne, kvôli zachovaniu rovnakej relatívnej chyby pre veľké aj malé hodnoty. Najjednoduchším nástrojom na nerovnomerné zaokrúhľovanie je odmerka na múku alebo na lieky, ktorá je pre malé množstvá úzka a pre väčšie širšia. V telefónii sa na nerovnomerné zaokrúhľovanie nepoužíva odmerka, ale realizuje sa pomocou logaritmickej funkcie, ktorou zmeníme škálovanie úrovní (kvantovacích hladín) signálu.

Posledným krokom spracovania pre prenos je kódovanie, teda prepis hodnoty signálu do binárneho tvaru, jeho vyjadrenie pomocou núl a jednotiek. Pri klasickej te-

lefónii sú potom tieto hodnoty prenášané cez okruh, ktorý je vytvorený počas celého telefónneho hovoru. Pri IP telefónii sú z digitálnych hodnôt tvorené rámce, ktoré sú následne balené do paketov. V jednom pakete sú dáta zodpovedajúce 20 ms reči. Po prenose dátovými sieťami je z prenesených diskretných hodnôt $f(k\Delta)$ vytvorený analógový (spojitý) signál ako súčet radu

$$f(t) = \sum_{k=-\infty}^{\infty} f(k\Delta) \frac{\sin(\Omega(t - k\Delta))}{\Omega(t - k\Delta)} \quad \text{kde} \quad \Omega = \frac{\pi}{\Delta} \quad (1.2.1)$$

Problémy v IP telefónii nastávajú v prípade, že niektoré z paketov nedorazia včas. Tento jav je spôsobený náhodným charakterom oneskorenia v sieti (chvenie), alebo poškodením paketu pri prenose. Chvenie (jitter) je spôsobené náhodným zafažením siete a rôznymi cestami, ktorými pakety s dátami sieťou prechádzajú. Navyše, aj keď je možné detekovať poškodené pakety, nie je možné si vyžiadať ich opätovný prenos, pretože v tom čase už majú byť prehrávané telefonujúcemu účastníkovi. Rozklad do nekonečného radu (1.2.1) sa na počesť jeho autora nazýva Shannonov rozklad a môžeme ho chápať aj ako rozklad signálu do bázy, pričom prenášané vzorky (dáta) sú vlastne koeficienty tohto rozkladu. Čo to znamená rozklad signálu do bázy si vysvetlíme v nasledujúcej podkapitole.

1.3 Výpočet koeficientov vektora v rôznych bázach

V tomto odseku si na jednoduchých príkladoch ukážeme ako sa počítajú koeficienty rozkladu do bázy. Ako model, použijeme trojrozmerný Euklidov priestor, jednoduchšie povedané: priestor vytvorený z usporiadaných trojíc reálnych čísel.

Bázou $\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2)$ tohto priestoru nazveme systém vektorov, ktorý je úplný a jednoznačný. Teda každý vektor priestoru sa bude dať napísať ako lineárna kombinácia vektorov bázy a navyše, bude sa to dať urobiť iba jediným spôsobom.

Bázami trojrozmerného Euklidovho priestoru môžu byť napríklad systémy:

$$\mathcal{B}_0 = ((1, 0, 0), (0, 1, 0), (0, 0, 1))$$

$$\mathcal{B}_1 = ((5, 0, 2), (0, 2, 0), (1/5, 0, -1/2))$$

$$\mathcal{B}_2 = ((1, 1, 0), (0, 1, 1), (1, 0, 1))$$

Rozložme teraz vektor $\mathbf{f} = (5, 4, 2)$ postupne do troch báz $\mathcal{B}_0, \mathcal{B}_1, \mathcal{B}_2$.

Pod pojmom rozložiť vektor \mathbf{f} do bázy $\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2)$ rozumieme nájsť také koeficienty c_0, c_1, c_2 , pre ktoré platí

$$\mathbf{f} = (5, 4, 2) = c_0 \mathbf{b}_0 + c_1 \mathbf{b}_1 + c_2 \mathbf{b}_2$$

V báze

$$\mathcal{B}_0 = ((1, 0, 0), (0, 1, 0), (0, 0, 1))$$

bude mať vektor \mathbf{f} koeficienty rozkladu $c_0 = 5$, $c_1 = 4$, $c_2 = 2$.

Naozaj platí

$$\mathbf{f} = (5, 4, 2) = 5 \cdot (1, 0, 0) + 4 \cdot (0, 1, 0) + 2 \cdot (0, 0, 1)$$

Navyše sa koeficienty c_0 , c_1 , c_2 dajú vypočítať aj podľa vzorca

$$c_k = \frac{\mathbf{f} \odot \mathbf{b}_k}{\mathbf{b}_k \odot \mathbf{b}_k} \quad (1.3.1)$$

kde \odot značí skalárny súčin vektorov, napríklad

$$(5, 4, 2) \odot (1, 0, 0) = 5 \cdot 1 + 4 \cdot 0 + 2 \cdot 0 = 5$$

Teda

$$c_0 = \frac{(5, 4, 2) \odot \mathbf{b}_0}{\mathbf{b}_0 \odot \mathbf{b}_0} = \frac{(5, 4, 2) \odot (1, 0, 0)}{(1, 0, 0) \odot (1, 0, 0)} = \frac{5}{1} = 5$$

$$c_1 = \frac{(5, 4, 2) \odot (0, 1, 0)}{(0, 1, 0) \odot (0, 1, 0)} = \frac{4}{1} = 4$$

$$c_2 = \frac{(5, 4, 2) \odot (0, 0, 1)}{(0, 0, 1) \odot (0, 0, 1)} = \frac{2}{1} = 2$$

Rovnako v báze

$$\mathcal{B}_1 = ((5, 0, 2), (0, 2, 0), (1/5, 0, -1/2))$$

koeficienty c_0 , c_1 , c_2 vektora $\mathbf{f} = (5, 4, 2)$ vypočítame podľa rovnakého vzorca $c_k = \frac{\mathbf{f} \odot \mathbf{b}_k}{\mathbf{b}_k \odot \mathbf{b}_k}$. Ich hodnoty budú tentokrát

$$c_0 = \frac{(5, 4, 2) \odot \mathbf{b}_0}{\mathbf{b}_0 \odot \mathbf{b}_0} = \frac{(5, 4, 2) \odot (5, 0, 2)}{(5, 0, 2) \odot (5, 0, 2)} = \frac{29}{29} = 1$$

$$c_1 = \frac{(5, 4, 2) \odot (0, 2, 0)}{(0, 2, 0) \odot (0, 2, 0)} = \frac{8}{2} = 4$$

$$c_2 = \frac{(5, 4, 2) \odot (1/5, 0, -1/2)}{(1/5, 0, -1/2) \odot (1/5, 0, -1/2)} = \frac{0}{0.29} = 0$$

Naozaj platí

$$\mathbf{f} = (5, 4, 2) = 1 \cdot (5, 0, 2) + 4 \cdot (0, 2, 0) + 0 \cdot (1/5, 0, -1/2)$$

V báze

$$\mathcal{B}_2 = ((1, 1, 0), (0, 1, 1), (1, 0, 1))$$

už vzorec (1.3.1) pre výpočet koeficientov nefunguje. Koeficienty vypočítané podľa vzorca $c_k = \frac{\mathbf{f} \odot \mathbf{b}_k}{\mathbf{b}_k \odot \mathbf{b}_k}$ by totiž boli

$$c_0 = \frac{(5,4,2) \odot \mathbf{b}_0}{\mathbf{b}_0 \odot \mathbf{b}_0} = \frac{(5,4,2) \odot (1,1,0)}{(1,1,0) \odot (1,1,0)} = \frac{9}{2} = 4.5$$

$$c_1 = \frac{(5,4,2) \odot \mathbf{b}_1}{\mathbf{b}_1 \odot \mathbf{b}_1} = \frac{(5,4,2) \odot (0,1,1)}{(0,1,1) \odot (0,1,1)} = \frac{6}{2} = 3$$

$$c_2 = \frac{(5,4,2) \odot \mathbf{b}_2}{\mathbf{b}_2 \odot \mathbf{b}_2} = \frac{(5,4,2) \odot (1,0,1)}{(1,0,1) \odot (1,0,1)} = \frac{7}{2} = 3.5$$

Teraz ale lineárnou kombináciou báзовých vektorov s koeficientmi

$$c_0 = 4.5, \quad c_1 = 3, \quad c_2 = 3.5$$

nedostaneme vektor \mathbf{f} :

$$\mathbf{f} = (5, 4, 2) \neq 4.5 \cdot (1, 1, 0) + 3 \cdot (0, 1, 1) + 3.5 \cdot (1, 0, 1) = (8, 7.5, 6.5)$$

Odvoďme všeobecný vzorec pre výpočet koeficientov c_0, c_1, c_2 vektora \mathbf{f} v báze $\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2)$. Vyjadrenie vektora \mathbf{f} v tejto báze je

$$\mathbf{f} = c_0 \mathbf{b}_0 + c_1 \mathbf{b}_1 + c_2 \mathbf{b}_2 \quad (1.3.2)$$

Vynásobíme skalárne vzťah (1.3.2) postupne všetkými báзовými vektormi. Dostaneme systém

$$\begin{aligned} \mathbf{f} \odot \mathbf{b}_0 &= c_0 \mathbf{b}_0 \odot \mathbf{b}_0 + c_1 \mathbf{b}_1 \odot \mathbf{b}_0 + c_2 \mathbf{b}_2 \odot \mathbf{b}_0 \\ \mathbf{f} \odot \mathbf{b}_1 &= c_0 \mathbf{b}_0 \odot \mathbf{b}_1 + c_1 \mathbf{b}_1 \odot \mathbf{b}_1 + c_2 \mathbf{b}_2 \odot \mathbf{b}_1 \\ \mathbf{f} \odot \mathbf{b}_2 &= c_0 \mathbf{b}_0 \odot \mathbf{b}_2 + c_1 \mathbf{b}_1 \odot \mathbf{b}_2 + c_2 \mathbf{b}_2 \odot \mathbf{b}_2 \end{aligned} \quad (1.3.3)$$

Po dosadení jednotlivých vektorov a vypočítaní príslušných skalárnych súčinov dostaneme systém troch rovníc o troch neznámych.

$$\begin{aligned} 9 &= c_0 2 + c_1 + c_2 \\ 6 &= c_0 + c_1 2 + c_2 \\ 7 &= c_0 + c_1 + c_2 2 \end{aligned}$$

Po vyriešení dostaneme správne hodnoty koeficientov

$$c_0 = 3.5, \quad c_1 = 0.5, \quad c_2 = 1.5$$

Naozaj platí

$$\mathbf{f} = (5, 4, 2) = 3.5 \cdot (1, 1, 0) + 0.5 \cdot (0, 1, 1) + 1.5 \cdot (1, 0, 1)$$

Zároveň vidíme, že pre bázy $\mathcal{B}_0, \mathcal{B}_1$ bol vzťah na výpočet jednoduchší preto, že výsledky súčinov $\mathbf{b}_k \odot \mathbf{b}_l = 0$ pre $k \neq l$. Takéto vektory, ktorých skalárny súčin je 0 nazývame kolmé (ortogonálne) a bázy, ktoré spĺňajú túto podmienku sa nazývajú ortogonálne bázy.

Teda ortogonálna báza je taká, že skalárny súčin každých dvoch rôznych vektorov je rovný nule. V ortogonálnej báze sa koeficienty vektora \mathbf{f} počítajú podľa vzorca

$$c_k = \frac{\mathbf{f} \odot \mathbf{b}_k}{\mathbf{b}_k \odot \mathbf{b}_k}$$

V neortogonálnej báze je potrebné vypočítať koeficienty vektora ako riešenie systému rovníc (1.3.3).

Vráťme sa k báзам

$$\mathcal{B}_0 = ((1, 0, 0), (0, 1, 0), (0, 0, 1))$$

$$\mathcal{B}_1 = ((5, 0, 2), (0, 2, 0), (1/5, 0, -1/2))$$

$$\mathcal{B}_2 = ((1, 1, 0), (0, 1, 1), (1, 0, 1))$$

Báza \mathcal{B}_0 sa nazýva jednotková báza a koeficienty vektora v nej ani netreba počítať, ale rovno použiť namerané hodnoty vektora.

Bázy \mathcal{B}_0 a \mathcal{B}_1 sú ortogonálne bázy a koeficienty vektora v nich vypočítame rýchlo podľa jednoduchého vzorca (1.3.1).

Báza \mathcal{B}_2 je neortogonálna báza a koeficienty vektora v nej vypočítame zložitejšie, riešením systému rovníc.

1.4 Nekonečnorozmerné vektorové priestory

Priestor rečových signálov tvoria spojitité, hladké funkcie s ohraničenou energiou. Môžeme ich však tiež považovať za vektory, teda prvky nekonečnorozmerného vektorového priestoru. Pri spracovaní reči sa najčastejšie používajú dva spôsoby rozkladu. Oba používajú ortogonálny systém báзовých funkcií.

Pre rozpoznávanie a analýzu rečového signálu sa používa Fourierova transformácia. Fourierova (alebo tiež harmonická) báza pozostáva z funkcií

$$\mathbf{b}_n(t) = e^{j\omega_n t}$$

kde j je komplexná jednotka, ale známejší je asi tvar

$$\mathbf{b}_n(t) = \cos \omega_n t \quad \text{a} \quad \mathbf{a}_n(t) = \sin \omega_n t$$

kde $\omega_n = n\omega_0 = n\frac{2\pi}{T_0}$. Fourierovu bázu by sme mohli prirovnať k báze

$$\mathcal{B}_1 = ((5, 0, 2), (0, 2, 0), (1/5, 0, -1/2))$$

ktorá je ortogonálna, ale nevidno to na prvý pohľad. Podrobnejšie odvodenie bázy ako aj základy práce s komplexnými číslami sú obsahom kapitoly 5.1.1. Rozklad signálu $\mathbf{f}(t)$ do harmonickej bázy je

$$\mathbf{f}(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega_0 t}$$

a jeho koeficienty vypočítame podľa známeho vzorca rozkladu do ortogonálnej bázy:

$$c_n = \frac{\mathbf{f}(t) \odot \mathbf{b}_n(t)}{\mathbf{b}_n(t) \odot \mathbf{b}_n(t)}$$

Skalárny súčin je však v tomto priestore definovaný pomocou integrálu zo súčinu prvého vektora a vektora komplexne združeného k druhému vektoru:

$$\mathbf{f}(t) \odot \mathbf{g}(t) = \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} f(t) \overline{g(t)} dt$$

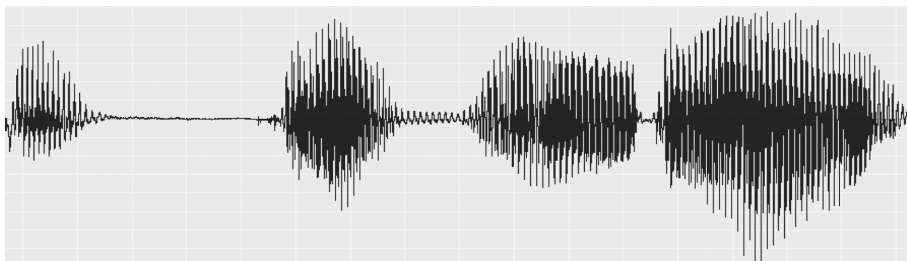
Pretože platí

$$\overline{(e^{jn\omega_0 t})} = e^{-jn\omega_0 t}$$

tak vzorec pre výpočet koeficientov má tvar

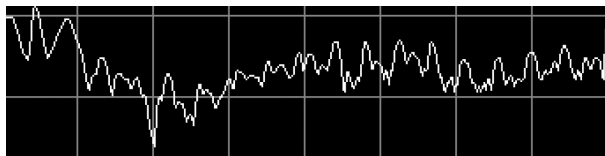
$$c_n = \frac{\mathbf{f}(t) \odot \mathbf{b}_n(t)}{\mathbf{b}_n(t) \odot \mathbf{b}_n(t)} = \frac{\int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} f(t) e^{-jn\omega_0 t} dt}{\int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} e^{jn\omega_0 t} e^{-jn\omega_0 t} dt} = \frac{1}{T_0} \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} f(t) e^{-jn\omega_0 t} dt, \quad T_0 = \frac{2\pi}{\omega_0}$$

Hodnoty $|c_n|$ sú vlastne amplitúdy jednotlivých funkcií sínus a kosínus, z ktorých sa signál skladá. Ako príklad rečového signálu je na obrázku 1.3 časový priebeh slova PYTAGORAS. Nasleduje vyjadrenie tohto slova vo Fourierovej (harmonickej) báze.



Obr. 1.3. Časový priebeh rečového signálu

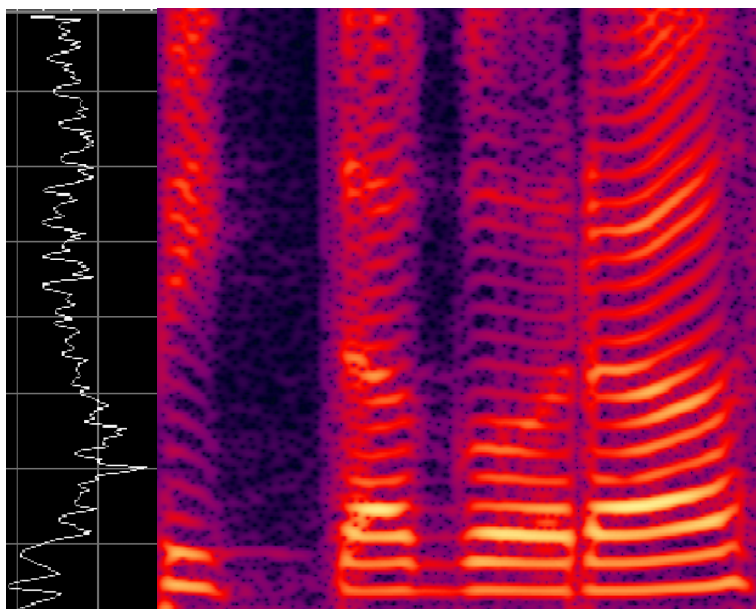
Rozklad krátkého úseku rečového signálu do Fourierovej bázy dáva koeficienty, ktorých veľkosť je znázornená na obrázku 1.4. Tieto hodnoty sa nazývajú okamžité amplitúdové spektrum signálu. Hodnoty okamžitého amplitúdového spektra sa počítajú pre krátky časový priebeh reči, ktorý sa z pôvodného signálu vyberie pomocou takzvaného okna. Pri výpočte ďalších koeficientov spektra sa okno posúva len o malý úsek, a tak sa okná prekrývajú. Niekedy sa hodnoty signálu v okne prenášobia váhovou funkciou, ktorá zabezpečí, že hodnoty na koncoch okna majú menší význam ako hodnoty v



Obr. 1.4. Priebeh amplitúdového spektra pre jedno okno

strede okna.

Na obrázku 1.5 je niekoľko takýchto priebehov vedľa seba, pričom veľkosť koeficientov je teraz udávaná farbou. Je to amplitúdové spektrum signálu z obrázku 1.3. Presnejšie je to amplitúdové Fourierove spektrum, lebo aj koeficienty v inej ako harmonickej báze sa niekedy nazývajú spektrom, alebo zovšeobecneným spektrom signálu.



Obr. 1.5. Priebeh spektra pre jedno okno a pre celý signál

Na prenos rečového signálu telefónnou sieťou, presnejšie pre A/D konverziu sa používa Shannonova transformácia. Shannonova báza pozostáva z funkcií

$$\mathbf{b}_n(t) = \frac{\sin(\Omega(t - n\Delta))}{\Omega(t - n\Delta)}, \quad \text{kde} \quad \Omega = \frac{\pi}{\Delta}$$

Rozklad signálu $\mathbf{f}(t)$ do tejto bázy je

$$\mathbf{f}(t) = \sum_{n=-\infty}^{\infty} c_n \frac{\sin(\Omega(t - n\Delta))}{\Omega(t - n\Delta)}$$

a jeho koeficienty vypočítame podľa známeho vzorca rozkladu do ortogonálnej bázy:

$$c_n = \frac{\mathbf{f}(t) \odot \mathbf{b}_n(t)}{\mathbf{b}_n(t) \odot \mathbf{b}_n(t)}$$

Skalárny súčin je aj v tomto priestore definovaný pomocou integrálu zo súčinu vektorov:

$$\mathbf{f}(t) \odot \mathbf{g}(t) = \int \mathbf{f}(t) \overline{\mathbf{g}(t)} dt$$

a po vypočítaní koeficientov v Shannonovej báze zistíme, že sa vlastne rovnajú nameraným vzorkám signálu

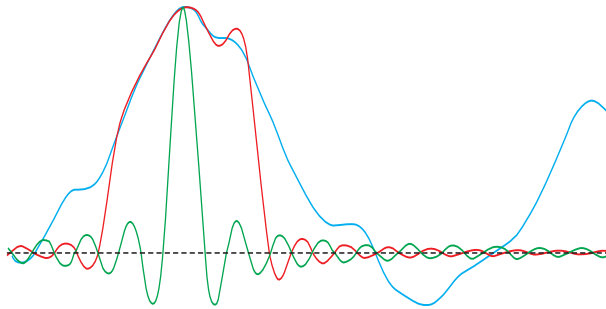
$$c_n = f(n\Delta)$$

Takže dostávame situáciu analogickú so situáciou v prípade bázy

$$\mathcal{B}_0 = ((1, 0, 0), (0, 1, 0), (0, 0, 1))$$

kedy koeficienty rozkladu nepotrebujeme počítať, lebo sú okamžite známe.

Príklad priebehu rečového signálu a jeho aproximácia pomocou jednej a siedmich Shannonových bazových funkcií je na obrázku 1.6.



Obr. 1.6. Signál vyjadrený v Shannonovej báze

Poznamenajme ešte, že aproximácia signálu pomocou konečného počtu bazových funkcií sa chová inak v prípade Fourierovej a inak v prípade Shannonovej bázy. V prípade Fourierovej bázy je chyba aproximácie rovnaká na celom časovom priebehu a s rastúcim počtom členov sa chyba rovnomerne na celej osi znižuje. V prípade Shannonovej bázy je aproximácia najpresnejšia v strede časového intervalu, a zhoršuje sa smerom k jeho okrajom, alebo inak, s rastúcim počtom členov sa rozširuje interval, v ktorom je funkcia odhadnutá presne.

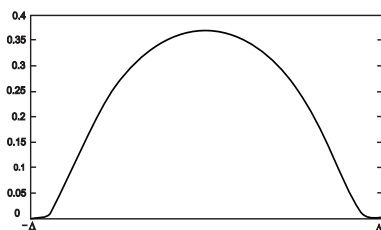
Pokúsme sa teraz nájsť analógiu k báze

$$\mathcal{B}_2 = ((1, 1, 0), (0, 1, 1), (1, 0, 1))$$

Takáto báza, pretože nie je ortogonálna, bude mať koeficienty, ktorých výpočet bude zložitejší. Bude mať aj však svoje výhody. Koeficienty v tejto báze môžu niesť aj informáciu o iných koeficientoch, ktoré umožnia rekonštrukciu koeficientov v stratenom pakete. V prípade, že báзовé funkcie budú mať tvar $\mathbf{b}_n(t) = h(t - n\Delta)$, bude mať rozklad signálu tvar

$$\mathbf{f}(t) = \sum_{n=-\infty}^{\infty} c_n h(t - n\Delta) \quad (1.4.1)$$

Pomocou teórie cyklostacionárnych náhodných procesov sa dá teoreticky dokázať, že po prenose signálu rozloženého do tvaru (1.4.1), bude mať rekonštruovaný signál rovnaké, alebo lepšie hodnoty odstupe signálu od šumu než pri v súčasnosti používanom Shannonovom rozklade. Teoretický výpočet bol urobený pre prípad straty paketov s prenášanými koeficientmi. Na obrázku 1.7 je príklad báзовej funkcie, pomocou ktorej sa naozaj dá urobiť rozklad reči do bázy.

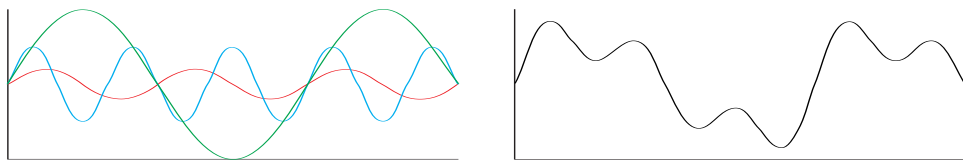


Obr. 1.7. Báзовá funkcia klobúk

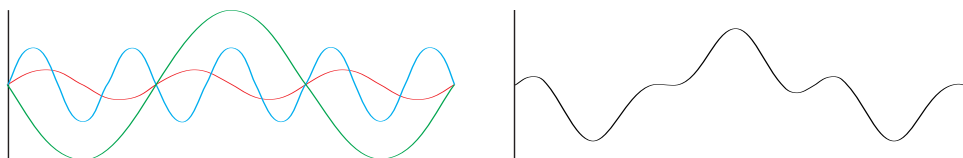
1.5 Využitie pre IP telefóniu

Zrekonštruovaný signál na strane prijímača má byť ľudským ušom počuteľný rovnako ako vyslaný signál. Nie je teda snahou, aby zrekonštruovaný signál bol čo najbližšie k pôvodnému v zmysle niektorej štandardnej metriky používanej pre vektorový priestor. Pre naše účely je potrebné, aby sa čo najmenej zmenili vlastnosti signálu, počuteľné pre ľudské ucho. Sluchom rozlišujeme amplitúdu a frekvenciu jednotlivých sínusových a kosínusových funkcií, z ktorých sa zložený rečový signál skladá. Nepočujeme fázový posun, o ktorý sú tieto funkcie posunuté. Ľudské ucho slúži ako frekvenčný analyzátor zvukových signálov. Nervové vlákna dávajú mozgu informáciu o prítomnosti jednotlivých funkcií sínus a kosínus v signáli. Mozog je informovaný o frekvencii, teda o hodnote n a amplitúde $|c_n|$ každej z nich. Fázový posun týchto funkcií však nevnímame. Na ilustráciu uvedieme dva rozdielne priebehy, ktoré ucho vníma rovnako.

Na obrázku 1.8 sú znázornené tri signály predstavujúce sínusové funkcie a zložený signál, ktorý vznikne ako ich súčet. Na obrázku 1.9 sú znázornené tri sínusové signály, ktoré majú rovnakú frekvenciu a amplitúdu, ako prvé tri. Od prvej trojice sa líšia iba o fázový posun jednej z funkcií. Zložený signál, ktorý vznikne ako súčet tejto trojice má iný časový priebeh. Vnímanie ľudským ušom je rovnaké pri zloženom signáli na obrázkoch 1.8 a 1.9. Tieto funkcie si blízke v metrike určenej počúvaním aj



Obr. 1.8. Tri sínusové signály a zložený signál, ktorý vznikne ako ich súčet



Obr. 1.9: Tri sínusové signály s posunom a zložený signál, ktorý vznikne ako ich súčet

keď v iných (klasických) metrikách sú od seba vzdialené. Nebudeme preto porovnávať priebeh pôvodného a rekonštruovaného signálu, ale zvolíme iné metódy porovnávania, konkrétne komplexný model, ktorý zahŕňa viacero parametrov, aby bolo ľudské počúvanie modelované čo najlepšie. Jedným z parametrov v tomto modeli je odstup signálu od šumu

$$SNR = 10 \log \frac{N_f}{N_s}$$

kde N_f je stredný výkon signálu a N_s je stredný výkon šumu. Hodnota podielu SNR je bezrozmerná veličina, ale zaviedli sa pre ňu jednotky, decibely $[dB]$. Logaritmická stupnica sa zavádza kvôli potrebe jemnejšieho škálovania v oblasti nižších frekvencií. Dôvodom tejto úpravy je analógia s ľudským uchom, kde sú tiež lepšie rozlišované nižšie frekvencie. Definíciu a pekné príklady na výpočet SNR je možné nájsť v publikácii [25]. V práci [2] je popísané, že rozklad a následná rekonštrukcia reči v neštandardnej báze vykazuje pri strate koeficientov lepší odstup signálu od šumu než klasicky používaný Shannonov rozklad.

Načrtnuté myšlienky ukazujú, že v podmienkach IP telefónie vznikajú principiálne iné technické problémy a ohraničenia, ako v klasickej telefónii. Hlbšie matematické pochopenie toho, ako funguje telefónia umožňuje využiť výhody jednej pri zavádzaní druhej a nechať priestor na modifikáciu toho, čo potrebné nie je. Navyše sme sa mohli presvedčiť, že matematika je prostriedkom, ktorý pomáha napredovať technickým riešeniam.

2 Procesy a dáta ako prvky vektorových priestorov

Pri spracovaní procesov je typickou úlohou rozložiť (analyzovať) daný proces alebo signál na dve alebo viacero zložiek. Je to prirodzený vedecký prístup. Geológa pri hodnotení rudného ložiska zaujíma, koľko zlata, striebra, medi sa môže z danej horniny vydolovať. Nutricionistu zaujíma koľko, tuku, cukru, vitamínov a vlákniny je v jednej porcii obedu v MacDonalde. Delenie na zložky môže mať rôzne motivácie:

- môžeme rozkladať proces na zaujímavú a nezaujímavú časť,
- môžeme rozkladať proces na známu a neznámu časť,
- môžeme rozkladať proces na podstatnú a nepodstatnú časť,
- môžeme rozkladať proces aj na viacero častí, a na zvyšok.

Tento rozklad (analýza) sa väčšinou deje vo vektorovom priestore. Cieľom tejto kapitoly je vybudovať intuitívnu predstavu o vektoroch, a súvisiacich pojmoch:

- vektorový priestor,
- lineárna kombinácia vektorov,
- báza vektorového priestoru,
- veľkosť vektora,
- skalárny súčin.

2.1 Vektorové priestory

Začnime otázkou: „Čo je to vektor?“ S týmto pojmom sme sa stretli v matematike aj fyzike a môžeme povedať, že v rôznych disciplínach sú slovom vektor popísané rôzne objekty. Napríklad vo fyzike predstavuje vektor veličinu, ktorá má okrem veľkosti aj smer. Tým sa líši od obyčajného čísla, ktoré má iba veľkosť. Príkladom vektora vo fyzike je rýchlosť, ktorá má veľkosť a smer. Ďalším príkladom vektora je zrýchlenie. V matematike na strednej škole sme sa stretli s vektorom ako usporiadanou n -ticou prvkov, označovaných ako zložky vektora. Čo majú tieto príklady vektorov spoločné?

Sčítovanie a odčítanie vektorov

V prvom rade je to možnosť sčítania a odčítania. Povedzme, že normálne bežíte rýchlosťou 20 kilometrov za hodinu. Ak vám fúka od chrbta vietor, tak dobehnete do cieľa rýchlejšie. Ak zase do tváre, tak pomalšie. Vo všeobecnosti výsledný vektor rýchlosti bežca je súčtom vektora jeho vlastnej rýchlosti a vektora rýchlosti vetra.

A ako sčítovať a odčítovať n -tice reálnych čísel? Po zložkách, teda

$$\begin{aligned}(4, 3, 7, -1) + (3, 4, 5, 6) &= (4 + 3, 3 + 4, 7 + 5, -1 + 6) = (7, 7, 12, 5) \\ (4, 3, 7, -1) - (3, 4, 5, 6) &= (4 - 3, 3 - 4, 7 - 5, -1 - 6) = (1, -1, 2, -7)\end{aligned}$$

Násobenie vektora reálnym číslom

Povedzme, že dnes nie ste vo forme a podávate len 90 % normálneho fyzického výkonu. Ak normálne bežíte rýchlosťou 20 kilometrov za hodinu, tak dnes ste schopný utekať len $0.9 \cdot 20 = 18$ kilometrov za hodinu. Je teda prirodzené hovoriť o tom, že vynásobíte váš vektor rýchlosti reálnym číslom.

Násobenie n -tíc reálnym číslom sa znova vykoná po zložkách

$$3 \cdot (2, 3, 4, 8) = (3 \cdot 2, 3 \cdot 3, 3 \cdot 4, 3 \cdot 8) = (6, 9, 12, 24)$$

podobne ako delenie nenulovým reálnym číslom

$$\frac{1}{2} \cdot (2, 3, 4, 8) = \left(\frac{1}{2} \cdot 2, \frac{1}{2} \cdot 3, \frac{1}{2} \cdot 4, \frac{1}{2} \cdot 8\right) = \left(1, \frac{3}{2}, 2, 4\right)$$

2.1.1 Veobecná definícia

Keby sme chceli presne odpovedať na otázku v úvode, možno (trocha lišiacky) povedať, že vektor je prvok vektorového priestoru. Teraz ale treba povedať, čo je vektorový priestor. **Vektorový priestor** je štruktúra obsahujúca dve množiny prvkov. Prvá z nich sú **skaláre** a tú druhú tvoria **vektory**. Prvky týchto dvoch množín možno navzájom kombinovať pomocou binárnych operácií

argumenty	možné operácie
skalár so skalárom	sčítateľ, odčítateľ, násobiť, deliť nenulovým prvkom
vektor s vektorom	sčítateľ, odčítateľ
skalár s vektorom	vynásobiť, vydeliť nenulovým skalárom

Tieto operácie musia spĺňať určité pravidlá (matematici ich nazývajú axiómy).

Komutatívnosť sčítania

Podobne ako pre reálne čísla platí $3 + 4 = 4 + 3$ musí platiť aj pre ľubovoľné dva skaláre a, b

$$a + b = b + a$$

a pre ľubovoľné dva vektory \mathbf{u}, \mathbf{v}

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$$

Ťažko si predstaviť, aký druh sčítania by nebol komutatívny. V matematike je to naozaj veľmi zriedkavý jav, pretože matematici používajú sčítanie takmer výlučne na označenie komutatívnych operácií. V chémii však napríklad nie je jedno či do skúmavky najprv dáte kyselinu sírovú a potom prilejete vodu, alebo naopak do vody prilejete kyselinu sírovú.

Asociatívnosť sčítania a násobenia

Pre ľubovoľné tri skaláre a, b, c musí platiť

$$\begin{aligned}(a + b) + c &= a + (b + c) \\ (a \cdot b) \cdot c &= a \cdot (b \cdot c)\end{aligned}$$

a podobne pre ľubovoľné tri vektory $\mathbf{u}, \mathbf{v}, \mathbf{w}$ musí platiť

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$$

Asociatívnosť operácií v matematike je ešte častejšia ako komutatívnosť. Avšak asociatívnosť je porušená v praxi pri počítaní s reálnymi číslami v počítači. Napríklad

$$(1 + 3.2 \cdot 10^{30}) + (-3.2 \cdot 10^{30}) \neq 1 + (3.2 \cdot 10^{30} + (-3.2 \cdot 10^{30}))$$

Existencia neutrálnych prvkov

Existujú dva (rôzne) skaláre $0, 1$ také, že pre ľubovoľný skalár platí

$$\begin{aligned}0 + a &= a + 0 = a \\ 1 \cdot a &= a \cdot 1 = a \\ 1 \cdot \mathbf{v} &= \mathbf{v}\end{aligned}$$

a nakoniec existuje nulový vektor $\mathbf{0}$ taký, že pre ľubovoľný vektor \mathbf{v}

$$\begin{aligned}\mathbf{0} + \mathbf{v} &= \mathbf{v} + \mathbf{0} = \mathbf{v} \\ \mathbf{0} \cdot \mathbf{v} &= \mathbf{0}\end{aligned}$$

Existencia inverzných prvkov

Pre ľubovoľný skalár a existuje prvok $-a$ taký, že

$$a + (-a) = 0$$

Podobne pre ľubovoľný vektor \mathbf{v} existuje opačný vektor $-\mathbf{v}$ taký, že

$$\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$$

Nakoniec pre ľubovoľný skalár a rôzny od 0 musí existovať inverzný prvok a^{-1} taký, že

$$a \cdot a^{-1} = 1$$

Existencia inverzných prvkov pre binárne operácie je relatívne zriedkavý jav, napríklad operácia zjednotenia množín má neutrálny prvok, ale nemá inverzný prvok. To isté platí pre sčítanie celých čísel na mikroprocesore so saturovaním.

Distributívne zákony

Tieto zákony dávajú do súvisu sčítanie s násobením.

Pre ľubovoľné skaláre a, b, c platí

$$(a + b) \cdot c = a \cdot c + b \cdot c$$

Podobne pre ľubovoľné skaláre a, b a vektor \mathbf{v} platí

$$(a + b) \cdot \mathbf{v} = a \cdot \mathbf{v} + b \cdot \mathbf{v}$$

a takisto pre ľubovoľné vektory \mathbf{v}, \mathbf{w} a skalár a platí

$$a \cdot (\mathbf{v} + \mathbf{w}) = a \cdot \mathbf{v} + a \cdot \mathbf{w}$$

Komutatívnosť násobenia skalárov

Pre ľubovoľné dva skaláre a, b musí platiť

$$a \cdot b = b \cdot a$$

Toto pravidlo je celosvetovo prijaté, s výnimkou Francúzska. My ho budeme predpokladať v celej knižke.

Všimnite si, že žiadna z axiém sa nezmieňovala o delenie. Je to preto, že delenie je definované pomocou násobenia a inverzného prvku. Teda podiel dvoch skalárov je

$$\begin{aligned} \frac{a}{b} &= a \cdot b^{-1} \\ \frac{a}{b} \cdot \mathbf{v} &= (a \cdot b^{-1}) \cdot \mathbf{v} = a \cdot (b^{-1}) \cdot \mathbf{v} \end{aligned}$$

2.1.2 Príklady vektorových priestorov

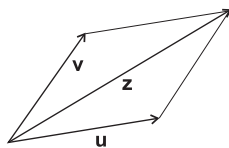
Definícia vektorového priestoru sa môže zdať komplikovaná, spĺňa však úlohu zaviesť pojem, ktorý zastreší všetky doterajšie predstavy o vektoroch. Treba si uvedomiť, že táto definícia plní ešte jednu úlohu a to, že vymedzuje aj iné, nové vektory, ktoré nemusia byť vektormi tak, ako ich chápeme v geometrii, alebo vo fyzike. Môže to byť ľubovoľný matematický objekt spĺňajúci axiomy vektorového priestoru. Napríklad polynómy stupňa najviac n , alebo spojité funkcie. Samozrejme na to, aby polynómy,

alebo funkcie mohli byť považované za vektory, musíme k uvedeným množinám objektov (vektorov) pridať operáciu sčítania, vymedziť množinu skalárov, operáciu násobenia skalárom a zabezpečiť platnosť všetkých axiém vektorového priestoru. V týchto vektorových priestoroch platia rovnaké tvrdenia a vety ako v priestoroch, ktoré budeme rozoberať podrobnejšie. Je len na vôli čitateľa, aké zovšeobecnenia si vyberie pre svoje potreby.

V ďalšom texte budeme pracovať s dvomi vektorovými priestormi $\mathbb{V}_1, \mathbb{V}_2$. Cieľom bude ukázať, že vlastnosti, ktorým rozumieme v jednom priestore sa dajú zovšeobecniť a použiť v inom priestore. Zavedieme teraz dva vektorové priestory \mathbb{V}_1 a \mathbb{V}_2 . Nech **vektorový priestor** \mathbb{V}_1 , je tvorený posunutiami roviny, teda množinu vektorov V_1 tvoria skupiny úsečiek, ktoré majú rovnaký smer, veľkosť a sú vzájomne rovnobežné. Množinu skalárov F_1 tvoria reálne čísla. Operácia $+$ na množine V_1 je definovaná pre ľubovoľné vektory $\mathbf{v}, \mathbf{u} \in \mathbb{V}_1$ nasledujúco $\mathbf{v} + \mathbf{u} = \mathbf{z}$, výsledný vektor \mathbf{z} má veľkosť a smer rovnaké ako uhlopriečka rovnobežníka so stranami \mathbf{v}, \mathbf{u} . Operácie $+$ a \cdot na množine $F_1 = \mathbb{R}$ sú definované obvyklým spôsobom. Násobenie vektora \mathbf{v} číslom c je definované tak, že výsledný vektor má rovnaký smer ako \mathbf{v} v prípade, že $c > 0$ a opačný smer ak $c < 0$. Veľkosť výsledného vektora je $|c|$ -násobok veľkosti vektora \mathbf{v} . Nulový vektor je jednoducho bod a opačný vektor je vektor, ktorý má prehodený počiatočný a koncový bod.

Vektorový priestor \mathbb{V}_2 je tvorený N -ticami reálnych čísel. Množinu skalárov F_2 tvoria znovu reálne čísla. Operácia $+$ na množine V_2 je definovaná pre ľubovoľné vektory $\mathbf{v} = (v_0, v_1, \dots, v_{N-1}), \mathbf{u} = (u_0, u_1, \dots, u_{N-1}) \in \mathbb{V}_2$ nasledujúco: $\mathbf{v} + \mathbf{u} = \mathbf{z}$ je taký vektor, pre ktorý platí $\mathbf{z} = (z_0, z_1, \dots, z_{N-1}) = (v_0 + u_0, v_1 + u_1, \dots, v_{N-1} + u_{N-1})$. Operácie $+$ a \cdot na množine skalárov $F_2 = \mathbb{R}$ sú definované obvyklým spôsobom.

Príklad 2.1.1 Na obrázku 2.1 je znázornený postup pri geometrickom sčítaní dvoch vektorov z priestoru \mathbb{V}_1 .

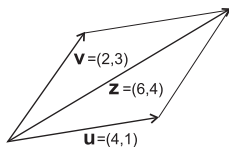


Obr. 2.1. Súčet vektorov \mathbf{v} a \mathbf{u}

Príklad 2.1.2 Vektory na obrázku 2.1 môžeme chápať ako usporiadané dvojice. Na obrázku 2.2 skutočne vidíme, že geometrický súčet v priestore \mathbb{V}_1 môžeme nahradiť algebraickým súčtom v priestore \mathbb{V}_2 .

$$\mathbf{v} + \mathbf{u} = \mathbf{z}$$

$$(2, 3) + (4, 1) = (6, 4)$$



Obr. 2.2. Súčet vektorov $\mathbf{v} = (2, 3)$ a $\mathbf{u} = (4, 1)$ je $\mathbf{z} = (6, 4)$

Úloha 2.1.1 Overte, že \mathbb{V}_1 a \mathbb{V}_2 spĺňajú axiomy vektorového priestoru.

Úloha 2.1.2 Dokážte, že ak vektor \mathbf{z} je súčtom vektorov (posunov) \mathbf{v} a \mathbf{u} v priestore \mathbb{V}_1 , tak zodpovedajúce vektory (usporiadané dvojice) \mathbf{v} a \mathbf{u} majú aj v priestore \mathbb{V}_2 súčet vektor \mathbf{z} .

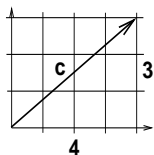
Príklad 2.1.3 Aká je veľkosť vektora $\mathbf{c} = (4, 3)$?

Riešenie:

Veľkosť vektora \mathbf{c} na obrázku 2.3 sa dá vypočítať pomocou Pytagorovej vety $c^2 = a^2 + b^2$. Teda

$$|\mathbf{c}| = \sqrt{a^2 + b^2} = \sqrt{4^2 + 3^2} = \sqrt{25} = 5$$

□



Obr. 2.3. Veľkosť vektora $\mathbf{c} = (4, 3)$.

Príklad 2.1.4 Aká je veľkosť vektora $\mathbf{d} = (a, b, c)$?

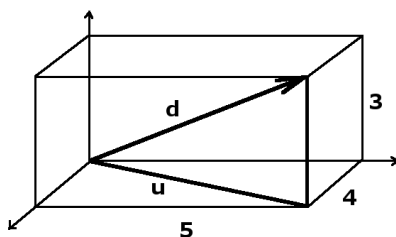
Riešenie:

Veľkosť vektora \mathbf{d} na obrázku 2.4 sa dá vypočítať pomocou dvojnásobného využitia Pytagorovej vety. Teda v trojrozmernom priestore bude veľkosť vektora

$$|\mathbf{d}| = \sqrt{a^2 + u^2} = \sqrt{a^2 + (\sqrt{b^2 + c^2})^2} = \sqrt{a^2 + b^2 + c^2}$$

□

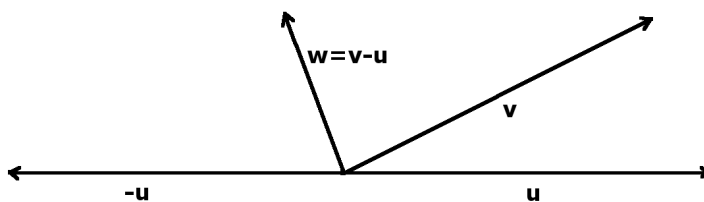
Úloha 2.1.3 Aká bude veľkosť vektora \mathbf{e} v štvorrozmernom priestore?

Obr. 2.4. Veľkosť vektora \mathbf{d} pre hodnoty $(a, b, c) = (3, 4, 5)$

Úloha 2.1.4 Môžeme vzorec pre veľkosť vektora v priestore \mathbb{V}_1 použiť v priestore \mathbb{V}_2 ? Ako bude vyzeráť vzorec pre veľkosť vektora, ktorý má N zložiek? Čo znamená rozmer priestoru?

Okrem súčtu vektorov je vo vektorovom priestore daný aj ich rozdiel.

Rozdiel vektorov $\mathbf{w} = \mathbf{v} - \mathbf{u}$ je súčet vektora \mathbf{v} a opačného vektora k vektoru \mathbf{u} . To je dobré si uvedomiť, kvôli geometrickej predstave rozdielu dvoch vektorov, ako je to vidno na obrázku 2.5.

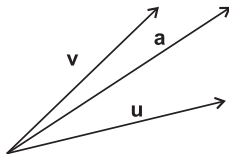
Obr. 2.5. Vektor \mathbf{w} je rozdiel vektorov \mathbf{v} a \mathbf{u}

S pojmami veľkosť vektora a rozdiel dvoch vektorov je už možné riešiť úlohu ako aproximovať jeden vektor iným vektorom a určiť chybu, ktorej sme sa pri tom dopustili.

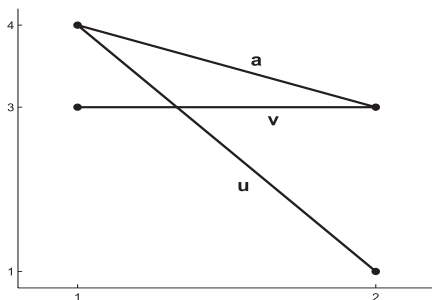
Príklad 2.1.5 Ktorý z vektorov \mathbf{u} , \mathbf{v} na obrázku 2.6 lepšie nahradí vektor \mathbf{a} ?

Riešenie:

V situácii, keď je potrebné vektor \mathbf{a} nahradiť jedným z vektorov \mathbf{u} alebo \mathbf{v} , vyberieme ten vektor, ktorý lepšie aproximuje pôvodný vektor. Vektor \mathbf{v} je k vektoru \mathbf{a} podstatne bližšie, \mathbf{v} sa viac podobá na \mathbf{a} , než vektor \mathbf{u} . Vyberieme vektor \mathbf{v} , lebo veľkosť rozdielu $\mathbf{v} - \mathbf{a}$ je menšia ako veľkosť rozdielu $\mathbf{u} - \mathbf{a}$. \square

Obr. 2.6. Aproximácia vektora \mathbf{a} bližším z vektorov \mathbf{u} , \mathbf{v}

Príklad 2.1.6 Vektory z príkladu 2.1.5 môžeme zakresliť do roviny aj tak, že jednotlivé zložky vektora budú hodnoty nejakého procesu v časoch 1 a 2, ako je to na obrázku 2.7. Takýmto spôsobom bude možné nakresliť aj viaczožkové vektory.

Obr. 2.7. Aproximácia 2-zložkového vektora \mathbf{a} bližším z vektorov \mathbf{u} , \mathbf{v}

Úloha 2.1.5 Môže sa problém z príkladu 2.1.5 riešiť aj v priestore \mathbb{V}_7 ?

Ktorý z vektorov $\mathbf{u} = (1, 4, 5, 2, 9, 3, 2)$, $\mathbf{v} = (4, 2, 1, 3, 5, 2, 3)$ je bližšie k vektoru $\mathbf{a} = (3, 5, 2, 1, 4, 4, 3)$?

Úloha 2.1.6 Ako zakresliť vektor z priestoru \mathbb{V}_7 ? Je vidno z obrázku 2.8, ktorý z vektorov \mathbf{u} alebo \mathbf{v} lepšie aproximuje vektor \mathbf{a} ?

Na to, aby sme mohli presne odpovedať na otázky v predchádzajúcich úlohách, bude potrebné zaviesť (alebo zopakovať) ďalšie pojmy z lineárnej algebry.

Nech c_0, c_1, \dots, c_{N-1} sú skaláry.

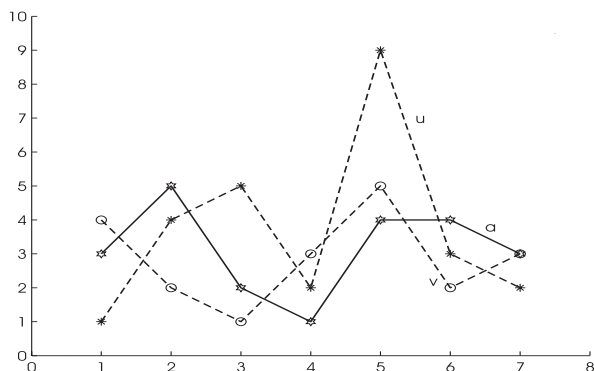
Vektor $\mathbf{v} = c_0 \cdot \mathbf{v}_0 + c_1 \cdot \mathbf{v}_1 + \dots + c_{N-1} \cdot \mathbf{v}_{N-1}$ nazveme **lineárnou kombináciou** vektorov $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{N-1}$.

Príklad 2.1.7 Vektor $(2, 4, 1)$ je lineárnou kombináciou vektorov $(1, 2, 4)$ a $(0, 0, 1)$, lebo

$$(2, 4, 1) = 2 \cdot (1, 2, 4) - 7 \cdot (0, 0, 1)$$

Rovnako vektory $(-1, -2, -4)$, $(0, 0, 0)$ aj $(1, 2, 1)$ sú lineárnymi kombináciami vektorov $(1, 2, 4)$ a $(0, 0, 1)$, pretože

$$(-1, -2, -4) = -1 \cdot (1, 2, 4) + 0 \cdot (0, 0, 1)$$

Obr. 2.8. Aproximácia 7-zložkového vektora **a** bližším z vektorov **u**, **v**

$$(0, 0, 0) = 0 \cdot (1, 2, 4) + 0 \cdot (0, 0, 1)$$

$$(1, 2, 1) = 1 \cdot (1, 2, 4) - 3 \cdot (0, 0, 1)$$

Úloha 2.1.7 Vyjadrite postupne vektor **a** = (2, 4) ako lineárnu kombináciu vektorov **u**, **v** a načrtnite príslušný obrázok.

- **u** = (1, 0), **v** = (0, 1)
- **u** = (2, 0), **v** = (0, 2)
- **u** = (1, 4), **v** = (4, 2)
- **u** = (1, 4), **v** = (2, 8)

Úloha 2.1.8 Vyjadrite postupne vektor **a** = (2, 4, 6) ako lineárnu kombináciu vektorov:

- **u** = (1, 0, 0), **v** = (0, 2, 3)
- **u** = (2, 0, 0), **v** = (0, 2, 4)
- **u** = (1, 0, 2), **v** = (0, 2, 0), **w** = (0, 0, 2)

Vektory **v**₀, **v**₁, ..., **v**_{N-1} sa nazývajú **lineárne závislé**, ak existuje vektor **c** taký, že platí

$$c_0 \mathbf{v}_0 + c_1 \mathbf{v}_1 + \dots + c_{N-1} \mathbf{v}_{N-1} = \mathbf{0}$$

$$\mathbf{c} = (c_0, c_1, \dots, c_{N-1}) \neq \mathbf{0}$$

Poznámka:

Vektor **c** je nenulový, ak aspoň jedno z čísel c_0, c_1, \dots, c_{N-1} je rôzne od nuly.

Príklad 2.1.8 Vektory $(2, 4, 8, 1)$, $(1, 2, 4, 0)$ a $(0, 0, 0, 1)$ sú lineárne závislé lebo

$$1 \cdot (2, 4, 8, 1) - 2 \cdot (1, 2, 4, 0) - 1 \cdot (0, 0, 0, 1) = (0, 0, 0, 0) = \mathbf{0}$$

teda existuje nenulový vektor koeficientov $\mathbf{c} = (1, -2, -1) \neq (0, 0, 0)$ taký, že pomocou neho je možné urobiť lineárnu kombináciu, ktorá je rovná vektoru $\mathbf{0}$.

Úloha 2.1.9 Nakreslite

- dvojicu lineárne závislých vektorov \mathbf{u} , \mathbf{v}
- trojicu lineárne závislých vektorov \mathbf{u} , \mathbf{v} , \mathbf{w}
- principiálne inú trojicu lineárne závislých vektorov \mathbf{u} , \mathbf{v} , \mathbf{w}

Úloha 2.1.10 Doplňte na

- dvojicu lineárne závislých vektorov $\mathbf{u} = (1, 0)$, $\mathbf{v} = ?$
- dvojicu lineárne závislých vektorov $\mathbf{u} = (1, 1)$, $\mathbf{v} = ?$
- dvojicu lineárne závislých vektorov $\mathbf{u} = (1, 0, 1)$, $\mathbf{v} = ?$
- trojicu lineárne závislých vektorov $\mathbf{u} = (1, 0, 1)$, $\mathbf{v} = (1, 1, 1)$, $\mathbf{w} = ?$
- trojicu lineárne závislých vektorov $\mathbf{u} = (4, 2, 4)$, $\mathbf{v} = (2, 1, 2)$, $\mathbf{w} = ?$

Vektory $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{N-1}$ sa nazývajú **lineárne nezávislé**, ak lineárna kombinácia $c_0\mathbf{v}_0 + c_1\mathbf{v}_1 + \dots + c_{N-1}\mathbf{v}_{N-1}$ je rovná nule iba pre nulový vektor $\mathbf{c} = (c_0, c_1, \dots, c_{N-1}) = \mathbf{0}$.

Príklad 2.1.9 Vektory $(2, 4, 7, 1)$, $(1, 2, 4, 0)$ a $(0, 0, 0, 1)$ sú lineárne nezávislé lebo rovnica

$$c_0 \cdot (2, 4, 7, 1) + c_1 \cdot (1, 2, 4, 0) + c_2 \cdot (0, 0, 0, 1) = (0, 0, 0, 0)$$

pre neznámu \mathbf{c} má iba jediné riešenie $\mathbf{c} = (0, 0, 0)$.

Úloha 2.1.11 Nakreslite

- dvojicu lineárne nezávislých vektorov \mathbf{u} , \mathbf{v}
- trojicu lineárne nezávislých vektorov \mathbf{u} , \mathbf{v} , \mathbf{w}
- principiálne inú trojicu lineárne nezávislých vektorov \mathbf{u} , \mathbf{v} , \mathbf{w}

Úloha 2.1.12 Doplňte na

- dvojicu lineárne nezávislých vektorov $\mathbf{u} = (1, 0)$, $\mathbf{v} = ?$
- dvojicu lineárne nezávislých vektorov $\mathbf{u} = (1, 1)$, $\mathbf{v} = ?$
- dvojicu lineárne nezávislých vektorov $\mathbf{u} = (1, 0, 1)$, $\mathbf{v} = ?$
- trojicu lineárne nezávislých vektorov $\mathbf{u} = (1, 0, 1)$, $\mathbf{v} = (1, 1, 1)$, $\mathbf{w} = ?$

- trojicu lineárne nezávislých vektorov $\mathbf{u} = (4, 2, 4)$, $\mathbf{v} = (2, 1, 2)$, $\mathbf{w} = ?$

Pri riešení úloh môže pomôcť nasledujúce tvrdenie (ktorého dôkaz čitateľ nájde v učebniciach lineárnej algebry alebo ho urobí sám). Pre lepšiu orientáciu v lineárnej algebre odporúčame zopakovať vedomosti zo základného kurzu algebry napríklad podľa [21], hlbšie uchopenie problematiky a teoretické zázemie nájde čitateľ v [26].

Tvrdenie 2.1.1 *Vektory $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{N-1}$ sú lineárne závislé, ak aspoň jeden z nich sa dá vyjadriť ako lineárna kombinácia ostatných vektorov.*

2.2 Báza a dimenzia vektorového priestoru

Jedným z najdôležitejších pojmov, s ktorými v tejto knihe budeme pracovať je pojem bázy. **Bázou vektorového priestoru** nazveme usporiadanú n -ticu vektorov $\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_{N-1})$ vektorového priestoru \mathbb{V} práve vtedy, ak

- vektory $\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_{N-1}$ sú lineárne nezávislé
- každý z vektorov priestoru \mathbb{V} možno vyjadriť ako lineárnu kombináciu vektorov tejto bázy, teda pre ľubovoľný vektor $\mathbf{x} \in V$ existujú také koeficienty $c_0, c_1, \dots, c_{N-1} \in F$, že platí

$$\mathbf{x} = c_0 \mathbf{b}_0 + c_1 \mathbf{b}_1 + \dots + c_{N-1} \mathbf{b}_{N-1}$$

Poznámka:

Hovoríme, že systém bazových vektorov je **nezávislý** (lineárna nezávislosť vektorov bázy) a **úplný** (všetky vektory sa dajú vyjadriť ako lineárna kombinácia vektorov bázy).

Nezávislosť systému zabezpečí, že dostaneme jednoznačný rozklad (a súčasne najmenší počet vektorov v lineárnej kombinácii). Vďaka úplnosti systému, môžeme do bázy rozložiť ľubovoľný vektor priestoru \mathbb{V} .

Príklad 2.2.1 *Systém vektorov $\mathcal{B} = ((2, 5), (-1, 3))$ je bázou vektorového priestoru \mathbb{V}_2 všetkých dvojzložkových vektorov.*

Riešenie:

Skutočne, ľubovoľný vektor $\mathbf{v} = (v_0, v_1)$ sa dá napísať ako lineárna kombinácia vektorov $(2, 5)$ a $(-1, 3)$

$$(v_0, v_1) = c_0(2, 5) + c_1(-1, 3)$$

Táto vektorová rovnica vedie na lineárnu sústavu rovníc

$$\begin{aligned} v_0 &= c_0 \cdot 2 + c_1 \cdot (-1) \\ v_1 &= c_0 \cdot 5 + c_1 \cdot 3 \end{aligned}$$

Riešením tejto sústavy (s neznámymi c_0, c_1 a parametrami v_0, v_1) sú

$$\begin{aligned} c_0 &= \frac{3v_0 + v_1}{11} \\ c_1 &= \frac{-5v_0 + 2v_1}{11} \end{aligned}$$

Skutočne platí, že

$$\begin{aligned} c_0 \cdot (2, 5) + c_1 \cdot (-1, 3) &= \frac{3v_0 + v_1}{11} \cdot (2, 5) + \frac{-5v_0 + 2v_1}{11} \cdot (-1, 3) = \\ &= \left(\frac{6v_0 + 2v_1}{11}, \frac{15v_0 + 5v_1}{11} \right) + \left(\frac{5v_0 - 2v_1}{11}, \frac{-15v_0 + 6v_1}{11} \right) \\ &= \left(\frac{11v_0}{11}, \frac{11v_1}{11} \right) = (v_0, v_1) \end{aligned}$$

□

Úloha 2.2.1 Zistite, či nasledujúce systémy vektorov tvoria bázu nejakého vektorového priestoru \mathbb{V} :

- $\mathcal{B}_0 = ((1, 0), (1, 1))$
- $\mathcal{B}_1 = ((1, 0, 1), (1, 1, 0))$
- $\mathcal{B}_2 = ((1, 0, 1), (1, 1, 0), (2, 2, 1))$
- $\mathcal{B}_3 = ((1, 0), (1, 1), (2, 3))$
- $\mathcal{B}_4 = ((1, 0, 1), (1, 1, 0), (0, 0, 1), (1, 0, 0))$
- $\mathcal{B}_5 = ((51, 14, 76, 22), (1, 1, 0, 4), (3, 0, 0, 1))$
- $\mathcal{B}_6 = ((34, 42, 50), (61, 19, 30), (0, 0, 0))$

Tvrdenie 2.2.1 Všetky bázy vektorového priestoru \mathbb{V} majú rovnaký počet prvkov.

Predchádzajúce tvrdenie umožňuje zaviesť pojem dimenzie priestoru.

Dimenzia vektorového priestoru \mathbb{V} je počet prvkov bázy tohto priestoru.

Poznámka:

Ak má vektorový priestor dimenziu N , nazýva sa **N -rozmerný vektorový priestor**, značíme $\dim(\mathbb{V}) = N$.

Prvky bázy nazývame **bázové vektory**, alebo **vektory bázy**.

Príklad 2.2.2 Koľko rozmerný je vektorový priestor \mathbb{V} , pozostávajúci zo všetkých trojzložkových vektorov (v_0, v_1, v_2) ?

Riešenie:

Vektory môžeme napísať ako lineárnu kombináciu

$$(v_0, v_1, v_2) = v_0 \cdot (1, 0, 0) + v_1 \cdot (0, 1, 0) + v_2 \cdot (0, 0, 1)$$

Systém troch vektorov $\mathcal{E} = ((1, 0, 0), (0, 1, 0), (0, 0, 1))$ tvorí bázu (presnejšie jednu z mnohých báz) priestoru \mathbb{V} . Ze je to naozaj báza ukážeme tak, že overíme úplnosť a lineárnu nezávislosť systému \mathcal{E} . Vektorový priestor \mathbb{V} , ktorého báza má tri prvky (vektory), je trojrozmerný. □

V príklade 2.2.2 sme ukázali úplne jasné tvrdenie, že trojzložkové vektory tvoria trojrozmerný vektorový priestor. Dôležitejší než výsledok úlohy je však postup, ktorý neskôr použijeme pri určovaní rozmeru v zložitejších priestoroch.

Tvrdenie 2.2.2 Každá množina N lineárne nezávislých vektorov patriacich do N -rozmerného vektorového priestoru \mathbb{V} tvorí bázu vektorového priestoru \mathbb{V} .

Vektorový priestor \mathbb{P} je **vektorovým podpriestorom** vektorového priestoru \mathbb{V} , ak pre každý vektor \mathbf{z} z \mathbb{P} platí:

$$\mathbf{v} \in \mathbb{P} \Rightarrow \mathbf{v} \in \mathbb{V}$$

Príklad 2.2.3 Objekty „priamka“, „rovina“, „priestor“ v geometrickom chápaní môžu byť vektorovými priestormi.

- Objekt „priamka“ je jednorozmerným vektorovým podpriestorom dvojrozmerného vektorového priestoru „rovina“.
- Objekt „rovina“ je dvojrozmerným vektorovým podpriestorom trojrozmerného vektorového priestoru „priestor“.
- Objekt „priamka“ je jednorozmerným vektorovým podpriestorom trojrozmerného vektorového priestoru „priestor“.

Tvrdenie 2.2.3 Množina všetkých lineárnych kombinácií vektorov $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{M-1} \in V_N$ tvorí vektorový podpriestor \mathbb{V}_M priestoru \mathbb{V}_N .

$$\mathbb{V}_M = \{c_0\mathbf{v}_0 + c_1\mathbf{v}_1 + \dots + c_{M-1}\mathbf{v}_{M-1}; c_0, c_1, \dots, c_{M-1} \in F\}$$

Hovoríme, že vektory $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{M-1}$ **generujú** podpriestor \mathbb{V}_M , značíme

$$\mathbb{V}_M = [\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{M-1}]$$

Príklad 2.2.4 Príklady systémov vektorov, ktoré generujú vektorové priestory.

- $\mathcal{S}_1 = ((1, 0, 0), (0, 1, 0), (0, 0, 1))$ generuje trojrozmerný vektorový priestor
- $\mathcal{S}_2 = ((1, 0, 0), (0, 1, 0))$ generuje dvojrozmerný vektorový podpriestor trojrozmerného vektorového priestoru
- $\mathcal{S}_3 = ((1, 1, 0), (2, 2, 0))$ generuje jednorozmerný vektorový podpriestor trojrozmerného vektorového priestoru

Úloha 2.2.2 Zistite, koľkorozmerný vektorový podpriestor generujú vektory

- $\mathcal{B}_1 = ((1, 0, 0, 0), (2, 0, 1, 0), (0, 0, 0, 1))$
- $\mathcal{B}_2 = ((4, 4, 3, 3), (2, 2, 1, 1), (0, 0, 1, 1))$

O podpriestory akého priestoru ide?

Úloha 2.2.3 Ako súvisí rozmer priestoru generovaného systémom vektorov \mathcal{B} s hodnotou matice \mathbf{B} , ktorej riadky sú jednotlivé vektory systému \mathcal{B} ?

Jednotkovou bázou \mathcal{E} vektorového priestoru \mathbb{V}_N , nazveme bázu

$$\mathcal{E} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N)$$

kde $\mathbf{e}_k = (0, \dots, 1, \dots, 0)$ je vektor, ktorý má všetky zložky 0, len na k -tom mieste má 1.

2.3 Súradnice vektora vzhľadom na bázu

Tvrdenie 2.3.1 Každý vektor \mathbf{v} vektorového priestoru \mathbb{V} možno jednoznačne vyjadriť ako lineárnu kombináciu báзовých vektorov

$$\mathbf{v} = \sum_{k=0}^{N-1} c_k \mathbf{b}_k$$

Koeficienty c_0, c_1, \dots, c_{N-1} rozkladu vektora do bázy nazývame **súradnice vektora** \mathbf{v} vzhľadom na bázu \mathcal{B} .

Píšeme

$$\mathbf{v} = (c_0, c_1, \dots, c_{N-1})\mathcal{B}$$

Ak bázu v zápise neoznačíme, budú to súradnice v jednotkovej báze \mathcal{E} :

$$\mathbf{v} = (v_0, v_1, \dots, v_{N-1})$$

$$\mathbf{v} = (v_0, v_1, \dots, v_{N-1})\mathcal{E}$$

Tvrdenie 2.3.2 k -tu súradnicu vektora \mathbf{v} vypočítame ako skalárny súčin vektora $\mathbf{e}_k = (0, \dots, 1, \dots, 0)$ a vektora $\mathbf{v} = (v_0, v_1, \dots, v_{N-1})$:

$$v_k = (\mathbf{v}, \mathbf{e}_k) = ((v_0, v_1, \dots, v_{N-1}), (0, \dots, 1, \dots, 0))$$

Matica patriaca k báze \mathcal{B} , je matica \mathbf{B} , ktorej riadky sú vektory bázy \mathcal{B} .

$$\mathbf{B} = \begin{pmatrix} \mathbf{b}_0 \\ \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_{N-1} \end{pmatrix} = \begin{pmatrix} b_{00} & b_{01} & \dots & b_{0N-1} \\ b_{10} & b_{11} & \dots & b_{1N-1} \\ \vdots & \vdots & \ddots & \vdots \\ b_{N-10} & b_{N-11} & \dots & b_{N-1N-1} \end{pmatrix}$$

Poznámka:

Ak vektor $\mathbf{v} = (v_0, v_1, \dots, v_{N-1})$ má vzhľadom na bázu \mathcal{B} súradnice c_0, c_1, \dots až c_{N-1} , potom platí:

$$\mathbf{v} = (v_0, v_1, \dots, v_{N-1})\mathcal{E}$$

$$\mathbf{v} = (c_0, c_1, \dots, c_{N-1})\mathcal{B} = \sum_{k=0}^{N-1} c_k \mathbf{b}_k = \mathbf{c} \cdot \mathbf{B} \quad (2.3.1)$$

Príklad 2.3.1 Aké sú súradnice vektora $\mathbf{v} = (-2, 11)$ v báze $\mathcal{B} = ((1, 2), (-1, 1))$?

Riešenie:

K báze $\mathcal{B} = ((1, 2), (-1, 1))$ patrí matica \mathbf{B} .

$$\mathbf{B} = \begin{pmatrix} \mathbf{b}_0 \\ \mathbf{b}_1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$$

Vyriešime sústavu rovníc

$$\mathbf{v} = (3, 5)\mathcal{B}$$

a dostaneme, že súradnice vektora $\mathbf{v} = (-2, 11)$ v báze \mathcal{B} sú $c_0 = 3, c_1 = 5$.
Naozaj platí

$$\begin{aligned}\mathbf{v} &= (-2, 11) \\ \mathbf{v} &= (3, 5)\mathcal{B} = c_0\mathbf{b}_0 + c_1\mathbf{b}_1 = 3 \cdot (1, 2) + 5 \cdot (-1, 1)\end{aligned}$$

V maticovom tvare

$$\mathbf{v} = (-2, 11) = (3, 5) \cdot \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} = \mathbf{c} \cdot \mathbf{B}$$

□

Úloha 2.3.1 Ako bude vyzerat' rovnica (2.3.1) v prípade jednorozmernej bázy \mathcal{B} ?

2.4 Zmena bázy

Pri analyzovaní procesu sa na dáta pozeráme rôznym spôsobom. Jednou z metód analýzy je chápať proces ako vektor a zmena pohľadu na dáta je vlastne zmena bázy a v nej vyjadrené súradnice procesu. Zmenou bázy zabezpečíme zmenu pohľadu na dáta ako sme to využili v kapitole 1.

Zmenou bázy \mathcal{B} na bázu \mathcal{B}^* rozumieme prechod od súradníc v báze \mathcal{B} k súradniciam v báze \mathcal{B}^* :

$$\begin{aligned}\mathbf{v} &= (c_0, c_1, \dots, c_{N-1})\mathcal{B} \\ \mathbf{v} &= (c_0^*, c_1^*, \dots, c_{N-1}^*)\mathcal{B}^* \\ \mathbf{v} &= \sum_{k=0}^{N-1} c_k \cdot \mathbf{b}_k = \mathbf{c} \cdot \mathbf{B} \\ \mathbf{v} &= \sum_{k=0}^{N-1} c_k^* \mathbf{b}_k^* = \mathbf{c}^* \cdot \mathbf{B}^*\end{aligned}$$

Z rovnosti

$$\mathbf{v} = \mathbf{c} \cdot \mathbf{B} = \mathbf{c}^* \cdot \mathbf{B}^* \quad (2.4.1)$$

dostávame

$$\mathbf{c}^* = \mathbf{c} \cdot \mathbf{B} \cdot (\mathbf{B}^*)^{-1} \quad (2.4.2)$$

Poznámka:

Pre stĺpcový vektor \mathbf{v} a maticu \mathbf{B} prislúchajúcu báze, tvorenú stĺpcovými vektormi

$$\mathbf{B} = \begin{pmatrix} \mathbf{b}_0 & \mathbf{b}_1 & \dots & \mathbf{b}_{N-1} \end{pmatrix} = \begin{pmatrix} b_{00} & b_{10} & \dots & b_{N-1,0} \\ b_{01} & b_{11} & \dots & b_{N-1,1} \\ \vdots & \vdots & \ddots & \vdots \\ b_{0,N-1} & b_{1,N-1} & \dots & b_{N-1,N-1} \end{pmatrix}$$

vzťahy (2.3.1), (2.4.1) a (2.4.2) majú tvar

$$\begin{aligned}\mathbf{v}_{\mathcal{B}} &= \begin{pmatrix} \mathbf{c}_0 \\ \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_{N-1} \end{pmatrix} \\ \mathbf{v} &= \sum_{k=0}^{N-1} c_k \mathbf{b}_k = \mathbf{B} \cdot \mathbf{c} \\ \mathbf{v} &= \mathbf{B} \cdot \mathbf{c} = \mathbf{B}^* \cdot \mathbf{c}^* \\ \mathbf{c}^* &= (\mathbf{B}^*)^{-1} \cdot \mathbf{B} \cdot \mathbf{c}\end{aligned}$$

□

Súradnice procesu vzhľadom na bázu \mathcal{E} nazývame **namerané hodnoty procesu**. Súradnice procesu vzhľadom na danú bázu \mathcal{B} nazývame **spektrum procesu** vzhľadom na bázu \mathcal{B} .

Príklad 2.4.1 *Nech vektor \mathbf{v} má v báze*

$\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2) = ((-2, 1, -5), (1, 2, 0), (-2, 1, 1))$ *koeficienty $\mathbf{c} = (1, 4, 3)$. Teda $\mathbf{v}_{\mathcal{B}} = (1, 4, 3)$.*

Aké koeficienty bude mať vektor \mathbf{v} v báze

$\mathcal{B}^* = (\mathbf{b}_0^*, \mathbf{b}_1^*, \mathbf{b}_2^*) = ((2, 1, 0), (1, 1, 0), (1, 1, 1))$?

Riešenie:

Označme si dané a hľadané súradnice

$$\begin{aligned}\mathbf{v}_{\mathcal{B}} &= (c_0, c_1, c_2) \\ \mathbf{v}_{\mathcal{B}^*} &= (c_0^*, c_1^*, c_2^*) \\ \mathbf{v} &= (c_0, c_1, c_2) \mathbf{B} = (c_0^*, c_1^*, c_2^*) \mathbf{B}^* \\ \mathbf{v} &= (1, 4, 3) \begin{pmatrix} -2 & 1 & -5 \\ 1 & 2 & 0 \\ -2 & 1 & 1 \end{pmatrix} = (c_0^*, c_1^*, c_2^*) \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \\ (c_0^*, c_1^*, c_2^*) &= (1, 4, 3) \begin{pmatrix} -2 & 1 & -5 \\ 1 & 2 & 0 \\ -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}^{-1} \\ (c_0^*, c_1^*, c_2^*) &= (-4, 12, -2) \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & -1 & 1 \end{pmatrix} = (-16, 30, -2)\end{aligned}$$

Skutočne

$$\begin{aligned}\mathbf{v}_{\mathcal{B}} &= (1, 4, 3) \\ \mathbf{v}_{\mathcal{B}^*} &= (-16, 30, -2)\end{aligned}$$

Lepšie však bude vypočítať najskôr súčin matic \mathbf{B} $(\mathbf{B}^*)^{-1}$, pretože matica

$$\mathbf{B}_{pr} = \mathbf{B} (\mathbf{B}^*)^{-1}$$

je maticou prechodu od bázy \mathcal{B} k báze \mathcal{B}^* .

$$(c_0^*, c_1^*, c_2^*) = (1, 4, 3) \begin{pmatrix} -3 & 9 & -5 \\ -1 & 3 & 0 \\ -3 & 3 & 1 \end{pmatrix} = (-16, 30, -2)$$

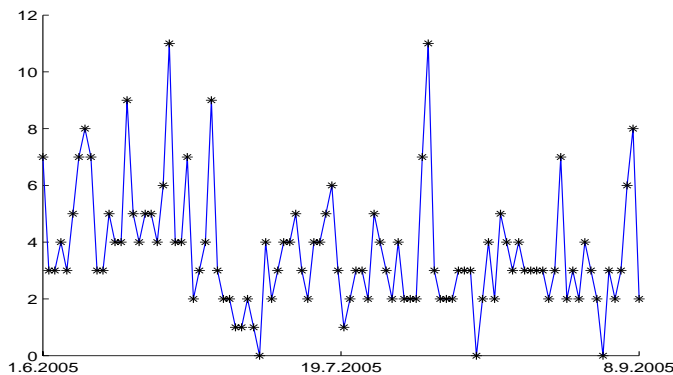
□

Úloha 2.4.1 Bez výpočtu odhadnite súčin vektora $(-16, 30, -2)$ a matice \mathbf{B}^* z príkladu 2.4.1.

Príklad 2.4.2 Zmena škály na súradnicových osiach grafu je vlastne jednoduchá zmena bázy. Napríklad vektor $\mathbf{v}_{\mathcal{E}} = (8, 6)$ v jednotkovej báze \mathcal{E} má v báze $\mathcal{B} = ((2, 0), (0, 2))$ súradnice $\mathbf{v}_{\mathcal{B}} = (4, 3)$.

Úloha 2.4.2 Zistite, kde sa používa logaritmické škálovanie a vysvetlite, prečo je výhodné ho použiť.

Úloha 2.4.3 Vyhľadajte na internete dáta, ktoré tvoria 100 hodnôt nejakého procesu v závislosti od času. Tieto dáta tvoria 100-zložkový vektor. Je to vektor 100-rozmerného vektorového priestoru. Vektor vykreslite a napíšte k nemu stručný popis, teda v akých časových intervaloch sa dáta merali a čo boli namerané hodnoty. Namerané hodnoty sú súradnice procesu vzhľadom na jednotkovú bázu \mathcal{E} . Zmeňte vektory bázy (napríklad desaťnásobne zväčšite ich dĺžku). Ako sa zmenia koeficienty procesu vzhľadom na túto bázu oproti pôvodným nameraným hodnotám?



Obr. 2.9. Počty koncertov ako proces

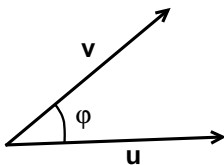
Príklad 2.4.3 100 zložkový vektor môžu tvoriť napríklad počty koncertov, ktoré sa uskutočnili v Bratislave v dobe od 1. júna do 8. októbra 2005, čo je presne 100 dní.

2.5 Skalárny súčin

Spomedzi vektorových priestorov sa v aplikáciách najčastejšie používajú vektorové priestory so skalárnym súčinom. V geometrii bol skalárny súčin definovaný ako skalár, ktorý vznikne ako výsledok binárnej operácie medzi dvoma vektormi. Predpis pre skalárny súčin dvoch vektorov v rovine je

$$\mathbf{v} \odot \mathbf{u} = |\mathbf{v}| \cdot |\mathbf{u}| \cdot \cos \varphi \quad (2.5.1)$$

kde φ je uhol, ktorý zvierajú vektory \mathbf{v} , \mathbf{u} .



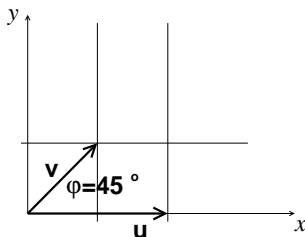
Obr. 2.10. Uhol vektorov \mathbf{v} a \mathbf{u}

Príklad 2.5.1 Vezmime si vektory $\mathbf{v} = (1, 1)$, $\mathbf{u} = (2, 0)$. Pre tieto dva vektory je ľahké určiť, že zvierajú uhol 45° , pozri obr. 2.11. Preto ich skalárny súčin sa vypočíta ako

$$(1, 1) \odot (2, 0) = |(1, 1)| \cdot |(2, 0)| \cdot \cos(45^\circ) = \sqrt{2} \cdot 2 \cdot \frac{\sqrt{2}}{2} = 2$$

Jednoduchší spôsob výpočtu je

$$(1, 1) \odot (2, 0) = 1 \cdot 2 + 1 \cdot 0 = 2$$



Obr. 2.11. Uhol vektorov $\mathbf{v} = (1, 1)$, $\mathbf{u} = (2, 0)$

Úloha 2.5.1 Zistite, prečo oba spôsoby výpočtu v príklade 2.5.1 dávajú rovnaký výsledok.

Úloha 2.5.2 Nakreslite v rovine aspoň tri rôzne dvojice vektorov, ktorých skalárny súčin je rovný nule.

Vo vektorovom priestore \mathbb{V} je teda možné okrem sčítania vektorov definovať ďalšiu operáciu nazývanú skalárny súčin. Zovšeobecnením geometrického skalárneho súčinu je skalárny súčin dvoch vektorov vo vektorovom priestore. Definujeme ho tak, aby v konkrétnom vektorovom priestore z príkladu 2.5.1 platil predpis (2.5.1). Zmena označenia skalárneho súčinu v tomto texte zo spôsobu $\mathbf{v} \odot \mathbf{u}$ na spôsob $\langle \mathbf{u}, \mathbf{v} \rangle$ je zámerná, treba si uvedomiť, že nasledujúca definícia je zovšeobecnením pôvodného pojmu.

Skalárny súčin vo vektorovom priestore $\mathbb{V} = ((V, +), (F, +, \cdot))$ je zobrazenie $\langle \cdot, \cdot \rangle : V \times V \rightarrow F$, pre ktoré platí:

$$\begin{aligned} \langle \mathbf{v}, \mathbf{v} \rangle &\geq 0, \text{ navyše } \langle \mathbf{v}, \mathbf{v} \rangle = 0 \Leftrightarrow \mathbf{v} = 0 && \text{pozitívnosť} \\ \langle \mathbf{v}, \mathbf{u} \rangle &= \langle \mathbf{u}, \mathbf{v} \rangle && \text{symetria} \\ \langle \mathbf{u} + \mathbf{h}, \mathbf{v} \rangle &= \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{h}, \mathbf{v} \rangle && \text{bilinearita} \\ \langle k \mathbf{v}, \mathbf{u} \rangle &= k \langle \mathbf{v}, \mathbf{u} \rangle \end{aligned} \quad (2.5.2)$$

Skalárny súčin nie je súčasťou definície vektorového priestoru. V tejto učebnici však budeme uvažovať iba priestory so skalárnym súčinom. Konečnorozmerný reálny vektorový priestor so skalárnym súčinom sa nazýva **Euklidov vektorový priestor**.

Úloha 2.5.3 *Dokážte, že súčin vektorov N -rozmerného Euklidovho vektorového priestoru \mathbb{V} z príkladu 2.1.1 definovaný ako*

$$\langle \mathbf{v}, \mathbf{u} \rangle = \sum_{k=0}^{N-1} v_k \cdot u_k$$

je naozaj skalárny súčin.

Vektory \mathbf{v}, \mathbf{u} nazveme navzájom **kolmé** (ortogonálne) práve vtedy, ak je ich skalárny súčin rovný 0.

$$\mathbf{v} \perp \mathbf{u} \Leftrightarrow \langle \mathbf{v}, \mathbf{u} \rangle = 0$$

Bázu vektorového priestoru nazveme **ortogonálnou bázou** práve vtedy, ak sú každé dva navzájom rôzne vektory bázy na seba kolmé:

$$\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_{N-1}); \quad \langle \mathbf{b}_i, \mathbf{b}_j \rangle = 0 \quad \text{pre } i \neq j$$

Príklad 2.5.2 *Báza*

- $\mathcal{B}_1 = ((1, 1, 0), (1, -1, 0), (0, 0, 3))$ je ortogonálna
- $\mathcal{B}_2 = ((1, 1, 0), (1, -1, 1), (0, 3, 3))$ nie je ortogonálna

Úloha 2.5.4 *Dokážte, že ortogonálny systém nenulových vektorov je lineárne nezávislý.*

Úloha 2.5.5 *Nájdite príklad použitia neortogonálnej bázy (napr. v deskriptívnej geometrii).*

2.6 Koeficienty vektora v ortogonálnej báze

V odseku 2.4 sme riešili úlohu, ako vypočítať koeficienty v novej báze, ak poznáme koeficienty v pôvodnej. Ilustráciu takéhoto výpočtu koeficientov nájdeme aj v kapitole 1.3. Nasledujúce tvrdenie dáva návod, ako koeficienty vypočítať pomocou skalárneho súčinu. Riešime úlohu, ako z nameraných hodnôt procesu vypočítať koeficienty v ortogonálnej báze \mathcal{B} .

$$\mathbf{v} = (v_0, v_1, \dots, v_{N-1})\mathcal{E}$$

$$\mathbf{v} = (c_0, c_1, \dots, c_{N-1})\mathcal{B}$$

$$c_0 = ?, c_1 = ?, \dots, c_{N-1} = ?$$

Tvrdenie 2.6.1 *Nech je vo vektorovom priestore daná ortogonálna báza \mathcal{B} , potom koeficienty ľubovoľného vektora $\mathbf{v} = (v_0, v_1, \dots, v_{N-1})\mathcal{E}$ sa v báze \mathcal{B} dajú vypočítať podľa vzorca*

$$c_i = \frac{\langle \mathbf{v}, \mathbf{b}_i \rangle}{\langle \mathbf{b}_i, \mathbf{b}_i \rangle} \quad (2.6.1)$$

Dôkaz:

Vypočítajme skalárny súčin s jednotlivými báзовými vektormi:

$$\langle \mathbf{v}, \mathbf{b}_i \rangle = \left\langle \sum_{k=0}^{N-1} c_k \mathbf{b}_k, \mathbf{b}_i \right\rangle = \sum_{k=0}^{N-1} \langle c_k \mathbf{b}_k, \mathbf{b}_i \rangle = \sum_{k=0}^{N-1} c_k \langle \mathbf{b}_k, \mathbf{b}_i \rangle = c_i \langle \mathbf{b}_i, \mathbf{b}_i \rangle$$

odtiaľ

$$c_i = \frac{\langle \mathbf{v}, \mathbf{b}_i \rangle}{\langle \mathbf{b}_i, \mathbf{b}_i \rangle}$$

□

Príklad 2.6.1 *Aké koeficienty má vektor $\mathbf{v}_{\mathcal{E}} = (1, 4, 2)$ v báze $\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2) = ((-2, 1, -5), (1, 2, 0), (-2, 1, 1))$?*

Riešenie:

Najskôr jednoduchým výpočtom overíme, že báza \mathcal{B} je ortogonálna.

Skutočne, skalárny súčin každej dvojice vektorov bázy je 0.

$$\begin{aligned} \langle \mathbf{b}_0, \mathbf{b}_1 \rangle &= (-2, 1, -5) \odot (1, 2, 0) = 0 \\ \langle \mathbf{b}_0, \mathbf{b}_2 \rangle &= (-2, 1, -5) \odot (-2, 1, 1) = 0 \\ \langle \mathbf{b}_1, \mathbf{b}_2 \rangle &= (1, 2, 0) \odot (-2, 1, 1) = 0 \end{aligned}$$

Báza je ortogonálna a preto môžeme použiť vzorec (2.6.1), podľa ktorého

$$\begin{aligned} c_0 &= \frac{\langle \mathbf{v}, \mathbf{b}_0 \rangle}{\langle \mathbf{b}_0, \mathbf{b}_0 \rangle} = -\frac{8}{30} \\ c_1 &= \frac{\langle \mathbf{v}, \mathbf{b}_1 \rangle}{\langle \mathbf{b}_1, \mathbf{b}_1 \rangle} = \frac{9}{5} \\ c_2 &= \frac{\langle \mathbf{v}, \mathbf{b}_2 \rangle}{\langle \mathbf{b}_2, \mathbf{b}_2 \rangle} = \frac{4}{6} \end{aligned}$$

Skutočne

$$c_0 \cdot \mathbf{b}_0 + c_1 \cdot \mathbf{b}_1 + c_2 \cdot \mathbf{b}_2 = -\frac{8}{30} \cdot (-2, 1, -5) + \frac{9}{5} \cdot (1, 2, 0) + \frac{4}{6} \cdot (-2, 1, 1) = (1, 4, 2)$$

□

Úloha 2.6.1 Je možné vypočítať koeficienty vektora pomocou skalárneho súčinu (ako v príklade 2.6.1) aj v prípade, že báza nie je ortogonálna?

Úloha 2.6.2 Navrhnite spôsob, ako pomocou skalárneho súčinu vektor

$$\mathbf{v} = (c_0, c_1, \dots, c_{N-1})\mathcal{B}$$

s koeficientmi v ortogonálnej báze \mathcal{B} vyjadriť ako vektor

$$\mathbf{v} = (c_0^*, c_1^*, \dots, c_{N-1}^*)\mathcal{B}^*$$

v inej ortogonálnej báze \mathcal{B}^* .

2.7 Norma vektora

Teraz už môžeme presne zaviesť pojem veľkosť vektora, ktorý sme zatiaľ zaviedli len intuitívne v geometrickom modeli. Vo vektorovom priestore so skalárnym súčinom môže byť **veľkosť (norma)** vektora definovaná nasledujúco:

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

Poznámka:

V Euklidovom vektorovom priestore môže byť skalárny súčin definovaný ako $\langle \mathbf{v}, \mathbf{u} \rangle = \sum_{k=0}^{N-1} v_k \cdot u_k$. Norma odvodená od tohto skalárneho súčinu má tvar

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{\sum_{k=0}^{N-1} v_k^2}$$

Vo vektorovom priestore môžu byť vzdialenosti definované aj iným spôsobom a normy nemusia byť odvodené od skalárneho súčinu.

Príklad 2.7.1 Aká je norma (veľkosť) vektora $\mathbf{u} = (1, 4, 5, 2, 9, 3, 2)$?

Po dosadení do vzorca dostaneme

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{\sum_{k=0}^{N-1} v_k^2} = \sqrt{1^2 + 4^2 + 5^2 + 2^2 + 9^2 + 3^2 + 2^2} = 11.83$$

Úloha 2.7.1 Zistite, čo znamená slovo ortonormálny. Nájdite dva ortonormálne vektory.

Nasledujúca veta je tvrdenie známe z geometrie, ktorého pravdivosť je možné dokázať aj vo všeobecnom vektorovom priestore. Stačí ak je v tomto priestore definovaný skalárny súčin.

Tvrdenie 2.7.1 (Pytagorova veta)

Vo vektorovom priestore so skalárnym súčinom pre každé dva ortogonálne vektory \mathbf{v} , \mathbf{u} platí:

$$\|\mathbf{v} + \mathbf{u}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{u}\|^2$$

Dôkaz:

Dosadíme súčet vektorov $\mathbf{u} + \mathbf{v}$ do definície normy:

$$\begin{aligned} \|\mathbf{v} + \mathbf{u}\|^2 &= \langle \mathbf{v} + \mathbf{u}, \mathbf{v} + \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{u} \rangle = \\ &= \langle \mathbf{v}, \mathbf{v} \rangle + 0 + 0 + \langle \mathbf{u}, \mathbf{u} \rangle = \|\mathbf{v}\|^2 + \|\mathbf{u}\|^2 \end{aligned}$$

Úloha 2.7.2 Dokážte, že Pytagorova veta platí aj pre rozdiel kolmých vektorov. Vo vektorovom priestore so skalárnym súčinom, pre každé dva ortogonálne vektory \mathbf{v} , \mathbf{u} platí:

$$\|\mathbf{v} - \mathbf{u}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{u}\|^2$$

Úloha 2.7.3 Platí aj nasledujúce tvrdenie?

Vo vektorovom priestore so skalárnym súčinom, pre každé dva ortogonálne vektory \mathbf{v} , \mathbf{u} platí:

$$\|\mathbf{v} + \mathbf{u}\|^2 = \|\mathbf{v} - \mathbf{u}\|^2$$

O aké geometrické tvrdenie ide?

Vo vektorovom priestore so skalárnym súčinom môžeme definovať **vzdialenosť** dvoch vektorov ako veľkosť rozdielového vektora.

$$d(\mathbf{v}, \mathbf{u}) = \|\mathbf{v} - \mathbf{u}\| \quad (2.7.1)$$

Tvrdenie 2.7.2 (Vlastnosti metriky) Dá sa ukázať, že vzdialenosť (metrika) definovaná vzťahom (2.7.1) má nasledujúce vlastnosti:

$$\begin{array}{ll} d(\mathbf{v}, \mathbf{u}) \geq 0, \text{ pričom } d(\mathbf{v}, \mathbf{u}) = 0 \Leftrightarrow \mathbf{v} = \mathbf{u} & \text{pozitívnosť} \\ d(\mathbf{v}, \mathbf{u}) = d(\mathbf{u}, \mathbf{v}) & \text{symetria} \\ d(\mathbf{v}, \mathbf{u}) \leq d(\mathbf{v}, \mathbf{h}) + d(\mathbf{h}, \mathbf{u}) & \text{trojuholníková nerovnosť} \end{array} \quad (2.7.2)$$

Existujú metriky, ktoré nie sú odvodené od skalárneho súčinu. Na to, aby bola binárna operácia (definovaná na dvojiciach vektorov, taká, že jej výsledkom je kladné reálne číslo) **metrika**, stačí aby spĺňala vlastnosti (2.7.2).

Špeciálna metrika, odvodená od skalárneho súčinu $\langle \mathbf{v}, \mathbf{u} \rangle = \sum_{k=0}^{N-1} v_k \cdot u_k$, ktorá má tvar

$$d(\mathbf{v}, \mathbf{u}) = \|\mathbf{v} - \mathbf{u}\| = \sqrt{\langle \mathbf{v} - \mathbf{u}, \mathbf{v} - \mathbf{u} \rangle} = \sqrt{\sum_{k=0}^{N-1} (v_k - u_k)^2} \quad (2.7.3)$$

sa nazýva **Euklidova metrika**.

Príklad 2.7.2 V zmysle Euklidovej metriky (2.7.3) sa problém z úlohy 2.1.5 môže riešiť ako hľadanie vektora s menšou vzdialenosťou.

Riešenie:

Ktorý z vektorov $\mathbf{u} = (1, 4, 5, 2, 9, 3, 2)$, $\mathbf{v} = (4, 2, 1, 3, 5, 2, 3)$ je bližšie k vektoru $\mathbf{a} = (3, 5, 2, 1, 4, 4, 3)$?

K vektoru \mathbf{a} je bližšie vektor \mathbf{v} , pretože

$$\|\mathbf{a} - \mathbf{v}\| = 4.47 < \|\mathbf{a} - \mathbf{u}\| = 6.48$$

□

Euklidova metrika sa používa napríklad pri rozpoznávaní reči. Jednotlivým hláskam sú priradené vektory. Ak sa nameraný zvuk nelíši svojimi hodnotami od vektora niektorej hlásky viac ako je vopred daná, empiricky zistená tolerancia, tak ho môžeme považovať za túto hlásku.

Metrika, ktorá nie je odvodená od skalárneho súčinu je napríklad Hammingova metrika. Používa sa pri kódovaní pomocou vektorov zložených z 0 a 1, kde sa meranie pomocou Euklidovej metriky ukázalo ako nevhodné. **Hammingova vzdialenosť** je počet jednotiek v súčte kódových slov modulo dva:

$$1011101010 + 0011001001 = 1000100011$$

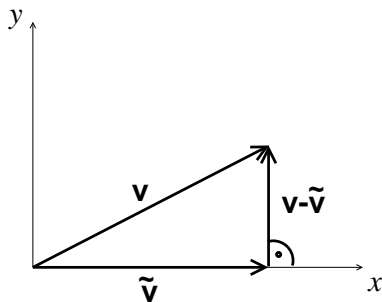
Hammingova metrika sa používa napríklad pri samoopravných kódach.

Úloha 2.7.4 Dokážte, že Hammingova vzdialenosť je skutočne metrika.

Úloha 2.7.5 Zistite ako je definovaná Mahalanobisova metrika a kde sa používa.

2.8 Kolmý priemet vektora do podpriestoru

Riešme v Euklidovom vektorovom priestore úlohu nájsť na x -ovej osi taký vektor $\tilde{\mathbf{v}}$, ktorý dostaneme ako obraz pôvodného vektora \mathbf{v} pri kolmom premietaní. To znamená, že vektory $\mathbf{v} - \tilde{\mathbf{v}}$ a $\tilde{\mathbf{v}}$ budú na seba kolmé. Na obrázku 2.12 je zobrazená geometrická predstava kolmého priemetu vektora \mathbf{v} na os x . Neskôr v tvrdení 3.2.1 ukážeme, že je naozaj pravda to, čo vidno na obrázku, že vektor $\mathbf{v} - \tilde{\mathbf{v}}$ je najkratší zo všetkých vektorov $\mathbf{v} - \tilde{\mathbf{y}}$, kde $\tilde{\mathbf{y}}$ leží na x -ovej osi.



Obr. 2.12. Vektor $\mathbf{v} - \tilde{\mathbf{v}}$ je kolmý na $\tilde{\mathbf{v}}$ (na priemet vektora \mathbf{v} na os x)

Príklad 2.8.1 Ku vektoru $\mathbf{v} = (3, 2)$ nájdite na x -ovej osi taký vektor $\tilde{\mathbf{v}}$, že vektory $\mathbf{v} - \tilde{\mathbf{v}}$ a $\tilde{\mathbf{v}}$ budú na seba kolmé.

Riešenie:

Najprv si povedzme, aké vektory budú ležať na osi x . Budú to všetky násobky vektora $\mathbf{b} = (1, 0)$, ale napríklad aj vektorov $(2, 0)$ alebo $(-3.7, 0)$. Os x je vlastne jednorozmerný vektorový podpriestor dvojrozmerného vektorového priestoru. Vektor ležiaci na osi x bude mať preto tvar

$$\tilde{\mathbf{v}} = c \cdot \mathbf{b} = c \cdot (1, 0) = (c, 0)$$

Pretože $\mathbf{v} - \tilde{\mathbf{v}}$ a $\tilde{\mathbf{v}}$ majú byť na seba kolmé, musí platiť, že ich skalárny súčin bude rovný 0.

$$\langle \mathbf{v} - \tilde{\mathbf{v}}, \tilde{\mathbf{v}} \rangle = 0$$

To nastane práve vtedy, keď

$$\langle \mathbf{v} - \tilde{\mathbf{v}}, \mathbf{b} \rangle = 0$$

Teda riešime rovnicu

$$\langle \mathbf{v} - c \cdot \mathbf{b}, \mathbf{b} \rangle = 0$$

Odtiaľ s využitím vlastností skalárneho súčinu dostaneme

$$\begin{aligned} \langle \mathbf{v}, \mathbf{b} \rangle - c \cdot \langle \mathbf{b}, \mathbf{b} \rangle &= 0 \\ \langle \mathbf{v}, \mathbf{b} \rangle &= c \cdot \langle \mathbf{b}, \mathbf{b} \rangle \end{aligned}$$

$$c = \frac{\langle \mathbf{v}, \mathbf{b} \rangle}{\langle \mathbf{b}, \mathbf{b} \rangle} = \frac{\langle (3, 2), (1, 0) \rangle}{\langle (1, 0), (1, 0) \rangle} = \frac{3}{1} = 3$$

Teda hľadaný priemet vektora $(3, 2)$ na os x bude vektor

$$\tilde{\mathbf{v}} = c \cdot \mathbf{b} = 3 \cdot (1, 0) = (3, 0)$$

□

Nech je daný N -rozmerný vektorový priestor \mathbb{V}_N so skalárnym súčinom a jeho M -rozmerný vektorový podpriestor \mathbb{V}_M , teda $0 < M \leq N$.

Nech $\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_{M-1})$ je báza podpriestoru \mathbb{V}_M .

Priemet vektora $\mathbf{v} \in \mathbb{V}_N$ do podpriestoru \mathbb{V}_M je vektor $\tilde{\mathbf{v}}$, ktorý leží v podpriestore \mathbb{V}_M , pričom vektor $\mathbf{v} - \tilde{\mathbf{v}}$ je kolmý na každý vektor podpriestoru \mathbb{V}_M .

Tvrdenie 2.8.1 *Koeficienty vektora $\tilde{\mathbf{v}} \in \mathbb{V}_M$, ktorý je kolmým priemetom vektora \mathbf{v} do M -rozmerného priestoru \mathbb{V}_M s ortogonálnou bázou*

$\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_{M-1})$, *vypočítame nasledujúco:*

$$c_k = \frac{\langle \mathbf{v}, \mathbf{b}_k \rangle}{\langle \mathbf{b}_k, \mathbf{b}_k \rangle} \quad \text{pre } k = 0, 1, \dots, M-1 \quad (2.8.1)$$

Dôkaz:

Pretože vektor $\mathbf{v} - \tilde{\mathbf{v}}$ musí byť kolmý na podpriestor \mathbb{V}_M , bude kolmý na každý prvok bázy \mathbb{V}_M . Platí

$$\langle \mathbf{v} - \tilde{\mathbf{v}}, \mathbf{b}_k \rangle = 0 \quad \text{pre } k = 0, 1, \dots, M-1$$

$\tilde{\mathbf{v}} \in \mathbb{V}_M$, preto

$$\langle \mathbf{v} - \tilde{\mathbf{v}}, \mathbf{b}_k \rangle = \langle \mathbf{v} - \sum_{i=0}^{M-1} c_i \mathbf{b}_i, \mathbf{b}_k \rangle = 0 \quad \text{pre } k = 0, 1, \dots, M-1$$

$$\langle \mathbf{v}, \mathbf{b}_k \rangle - \sum_{i=0}^{M-1} c_i \langle \mathbf{b}_i, \mathbf{b}_k \rangle = 0$$

$$\langle \mathbf{v}, \mathbf{b}_k \rangle = c_k \langle \mathbf{b}_k, \mathbf{b}_k \rangle \quad \text{pretože } \langle \mathbf{b}_i, \mathbf{b}_k \rangle = 0 \quad \text{pre } i \neq k$$

$$c_k = \frac{\langle \mathbf{v}, \mathbf{b}_k \rangle}{\langle \mathbf{b}_k, \mathbf{b}_k \rangle}$$

Keďže $\mathbf{v} - \tilde{\mathbf{v}}$ je kolmý na báзовé vektory, je kolmý na ich akúkoľvek lineárnu kombináciu, a teda na celý podpriestor \mathbb{V}_M . □

Príklad 2.8.2 *Aký je priemet vektora $\mathbf{v} = (2, 4, 8, 1)$ do vektorového podpriestoru generovaného bázou $\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1)$, kde*

$$\mathbf{b}_0 = (1, 2, 4, 0), \mathbf{b}_1 = (8, 0, -2, 1)?$$

Riešenie:

Priemet vektora $\mathbf{v} = (2, 4, 8, 1)$ do podpriestoru generovaného bázou

$\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1)$, kde $\mathbf{b}_0 = (1, 2, 4, 0)$, $\mathbf{b}_1 = (8, 0, -2, 1)$, je vektor $\tilde{\mathbf{v}}$ so súradnicami

$$\tilde{\mathbf{v}} = (2, 0.02)\mathcal{B}$$

$$\tilde{\mathbf{v}} = (2.16, 4, 7.96, 0.02)\mathcal{E}$$

Pretože báza \mathcal{B} je ortogonálna, koeficienty rozkladu sa vypočítajú zo vzorca (2.8.1).

$$c_0 = \frac{\langle (2, 4, 8, 1), (1, 2, 4, 0) \rangle}{\langle (1, 2, 4, 0), (1, 2, 4, 0) \rangle} = \frac{42}{21} = 2$$

$$c_1 = \frac{\langle (2, 4, 8, 1), (8, 0, -2, 1) \rangle}{\langle (8, 0, -2, 1), (8, 0, -2, 1) \rangle} = \frac{1}{69} = 0.02$$

□

2.9 Gramova-Schmidtova ortogonalizácia

Keď chceme z danej známej bázy $\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_{N-1})$ priestoru \mathbb{V} vyrobiť ortogonálnu bázu $\mathcal{B}' = (\mathbf{b}'_0, \mathbf{b}'_1, \dots, \mathbf{b}'_{N-1})$, môžeme použiť nasledujúci postup:

1. Vyberieme jeden z vektorov bázy \mathcal{B} a označíme ho ako vektor \mathbf{b}'_0 bázy \mathcal{B}' : $\mathbf{b}'_0 = \mathbf{b}_0$.
2. Nájdeme vektor \mathbf{b}'_1 taký, že $\mathbf{b}'_1 \perp \mathbf{b}'_0$ a leží v rovine (podpriestore) určenej vektormi $\mathbf{b}'_0, \mathbf{b}_1$. Túto požiadavku spĺňa vektor $\mathbf{b}'_1 = \mathbf{b}_1 - \tilde{\mathbf{b}}_1$, kde $\tilde{\mathbf{b}}_1$ je kolmý priemet vektora \mathbf{b}_1 na priamku určenú vektorom \mathbf{b}'_0 .
3. Nájdeme vektor \mathbf{b}'_2 taký, že $\mathbf{b}'_2 \perp \mathbf{b}'_0$, $\mathbf{b}'_2 \perp \mathbf{b}'_1$ a leží v podpriestore určenom vektormi $\mathbf{b}'_0, \mathbf{b}'_1, \mathbf{b}_2$. Túto požiadavku spĺňa vektor $\mathbf{b}'_2 = \mathbf{b}_2 - \tilde{\mathbf{b}}_2$, kde $\tilde{\mathbf{b}}_2$ je kolmý priemet vektora \mathbf{b}_2 do podpriestoru určeného vektormi $\mathbf{b}'_0, \mathbf{b}'_1$.
4. Krok 3 opakujeme, až kým v báze \mathcal{B}' nie je N vektorov.

Úloha 2.9.1 *Napište kroky 2 a 3 z algoritmu Gram-Schmidtovej ortogonalizácie ako jeden indukčný krok, v ktorom všeobecne vyjadrite vektor \mathbf{b}'_n pomocou vektorov $\mathbf{b}'_0, \mathbf{b}'_1, \dots, \mathbf{b}'_{n-1}$ a \mathbf{b}_n .*

Úloha 2.9.2 *Čo sa stane, ak systém vektorov \mathcal{B} , z ktorého chceme vyrobiť ortogonálnu bázu \mathcal{B}' pomocou Gram-Schmidtovej ortogonalizácie, nie je ortogonálny? Aký výsledok dá Gram-Schmidtov algoritmus?*

Úloha 2.9.3 *Nájdite ortogonálnu bázu vektorového priestoru určeného bázou*

$$\mathcal{B} = ((1, 2, 3, 4, 5, 6), (3, 4, 2, 1, 3, 4), (5, 4, 1, 0, 1, 2), (1, 3, 4, 5, 6, 1), (2, 3, 4, 5, 1, 2), (1, 2, 3, 4, 2, 1)).$$

Dá sa úloha riešiť bez počítača?

Príklad 2.9.1 *Ortogonalizujte pomocou Gram-Schmidtovej ortogonalizácie bázu $\mathcal{B} = ((1, 1, 1), (0, 1, 2), (0, 1, 4))$.*

Riešenie:

Pôvodná báza je

$$\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2) = ((1, 1, 1), (0, 1, 2), (0, 1, 4))$$

Ako prvý vektor bázy sme vybrali vektor $\mathbf{b}'_0 = (1, 1, 1)$.

Druhý vektor bázy \mathbf{b}'_1 bude rovný kolmému priemetu \mathbf{b}_1 na vektor \mathbf{b}'_0 a vypočítame ho zo vzťahu $\langle \mathbf{b}_1 - c\mathbf{b}'_0, \mathbf{b}'_0 \rangle = 0$, pričom $\mathbf{b}'_1 = \mathbf{b}_1 - c\mathbf{b}'_0$. Po vypočítaní koeficientu c nám vyjde $\mathbf{b}'_1 = (-1, 0, 1)$.

Analogicky bude tretí vektor novej bázy \mathbf{b}'_2 rovný kolmému priemetu \mathbf{b}_2 do roviny určenej vektormi \mathbf{b}'_0 a \mathbf{b}'_1 . Vypočítame ho zo vzťahu $\mathbf{b}'_2 = \mathbf{b}_2 - c_0\mathbf{b}'_0 - c_1\mathbf{b}'_1$ kde koeficienty c_0 a c_1 dostaneme vyriešením sústavy

$$\langle \mathbf{b}_2 - c_0\mathbf{b}'_0 - c_1\mathbf{b}'_1, \mathbf{b}'_0 \rangle = 0 \quad (2.9.1)$$

$$\langle \mathbf{b}_2 - c_0\mathbf{b}'_0 - c_1\mathbf{b}'_1, \mathbf{b}'_1 \rangle = 0 \quad (2.9.2)$$

Dostaneme výsledok $\mathbf{b}'_2 = (\frac{1}{3}, -\frac{2}{3}, \frac{1}{3})$. Ortogonálna báza, ktorú dostaneme Gram-Schmidtovou ortogonalizáciou je báza

$$\mathcal{B}' = (\mathbf{b}'_0, \mathbf{b}'_1, \mathbf{b}'_2) = \left((1, 1, 1), (-1, 0, 1), \left(\frac{1}{3}, -\frac{2}{3}, \frac{1}{3} \right) \right)$$

Neskôr ukážeme užitočnosť takto vytvorenej ortogonálnej bázy \mathcal{B}' . Teraz len poznamenáme, že vektory bázy \mathcal{B}' majú nasledujúce vlastnosti:

- prvý vektor $\mathbf{b}'_0 = (1, 1, 1)$ predstavuje konštantnú zložku procesu,
- vektor $\mathbf{b}'_1 = (-1, 0, 1)$ predstavuje lineárnu zložku procesu,
- vektor $\mathbf{b}'_2 = (\frac{1}{3}, -\frac{2}{3}, \frac{1}{3})$ predstavuje kvadratickú (parabolickú) zložku procesu.

Navyše vektor \mathbf{b}'_1 predstavuje lineárny nárast bez konštantnej zložky a vektor \mathbf{b}'_2 predstavuje kvadratický priebeh bez konštantnej a bez lineárnej zložky. Vektory $\mathbf{b}'_0, \mathbf{b}'_1, \mathbf{b}'_2$ sú vykreslené na obrázku 2.13. \square

Príklad 2.9.2 *Nájdite konštantnú, lineárnu a kvadratickú zložku vektorov $\mathbf{v} = (3, 6, 2)$ a $\mathbf{u} = (0, 2, 7)$.*

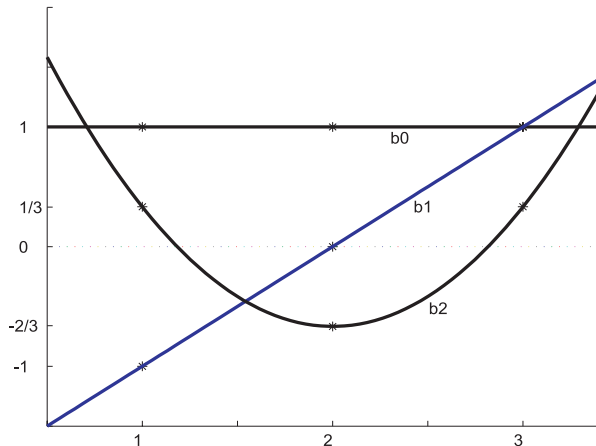
Riešenie:

V trojrozmernom vektorovom priestore bude konštantný priebeh predstavovať vektor

$$\mathbf{b}'_0 = (1, 1, 1)$$

lineárny nárast bude predstavovať vektor

$$\mathbf{b}'_1 = (-1, 0, 1)$$



Obr. 2.13: Ortogonálne vektory $\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2$, ktoré predstavujú očistený konštantný, lineárny a kvadratický priebeh

a kvadratický priebeh bude predstavovať vektor

$$\mathbf{b}'_2 = \left(\frac{1}{3}, -\frac{2}{3}, \frac{1}{3} \right)$$

Tieto vektory sú na seba kolmé, preto môžeme podľa vzorca (2.8.1) vypočítať koeficienty vektora $\mathbf{v} = (3, 6, 2)$:

$$c_0 = \frac{11}{3}, \quad c_1 = -\frac{1}{2}, \quad c_2 = -\frac{7}{2}$$

Vidíme, že platí

$$(3, 6, 2) = \frac{11}{3} \cdot (1, 1, 1) - \frac{1}{2} \cdot (-1, 0, 1) - \frac{7}{2} \cdot \left(\frac{1}{3}, -\frac{2}{3}, \frac{1}{3} \right)$$

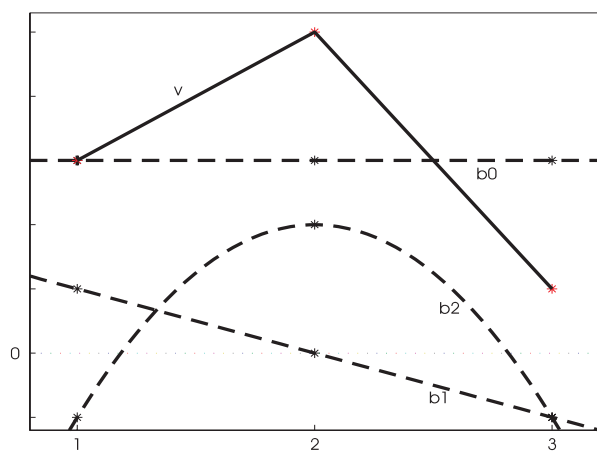
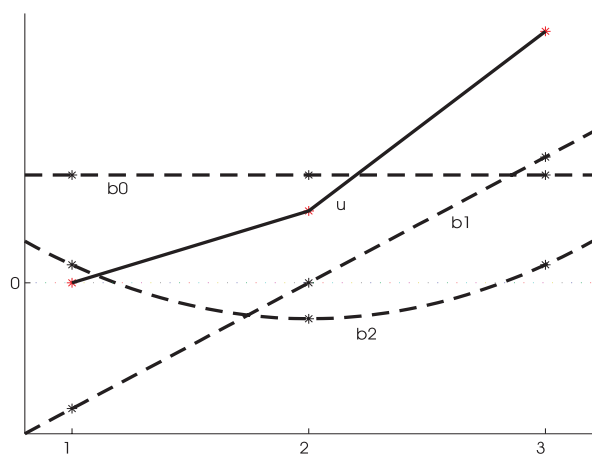
Pre vektor $\mathbf{u} = (0, 2, 7)$ sú koeficienty

$$d_0 = 3, \quad d_1 = 3.5, \quad d_2 = 1.5$$

$$(0, 2, 7) = 3 \cdot (1, 1, 1) + 3.5 \cdot (-1, 0, 1) + 1.5 \cdot \left(\frac{1}{3}, -\frac{2}{3}, \frac{1}{3} \right)$$

Navyše z obrázkov 2.14 a 2.15 vidíme, ako dobre jednotlivé vektory $\mathbf{b}'_0, \mathbf{b}'_1, \mathbf{b}'_2$ aproximujú každý z pôvodných vektorov \mathbf{v} a \mathbf{u} .

Príklad 2.9.3 Nájdite ortogonálnu bázu $\mathcal{B}' = (\mathbf{b}'_0, \mathbf{b}'_1)$ vektorového podpriestoru \mathbb{U} priestoru \mathbb{V} , ak viete, že $\mathbb{U} = [(1, 2, 0, -1), (1, 1, 3, 0)]$.

Obr. 2.14. Vektor $\mathbf{v} = (3, 6, 2)$ a príspevky vektorov $\mathbf{b}'_0, \mathbf{b}'_1, \mathbf{b}'_2$ Obr. 2.15. Vektor $\mathbf{u} = (0, 2, 7)$ a príspevky vektorov $\mathbf{b}'_0, \mathbf{b}'_1, \mathbf{b}'_2$

Riešenie:

Vektory $(1, 2, 0, -1)$, $(1, 1, 3, 0)$ sú lineárne nezávislé, ale nie sú ortogonálne, lebo ich skalárny súčin je 3 a nie 0.

Keby sme úlohu riešili tak, že vyberieme dva kolmé vektory,

napr. $(1, 2, 0, -1)$, $(1, 0, 0, 1)$, dostaneme síce ortogonálnu bázu dvojrozmerného podpriestoru, ale nie podpriestoru \mathbb{U} generovaného vektormi $\mathbf{b}_0 = (1, 2, 0, -1)$, $\mathbf{b}_1 = (1, 1, 3, 0)$.

Napríklad vektor $(1, 1, 3, 0)$ sa nedá napísať ako lineárna kombinácia vektorov $(1, 2, 0, -1)$, $(1, 0, 0, 1)$.

Pri báze, ktorú dostaneme Gram-Schmidtovou ortogonalizáciou však je zaručené, že je bazou toho istého podpriestoru ako báza, z ktorej bola odvodená:

Ako prvý zoberieme do bázy prvý vektor $\mathbf{b}'_0 = (1, 2, 0, -1)$.

Druhý vektor \mathbf{b}'_1 , dostaneme ako rozdiel vektora \mathbf{b}_1 a jeho kolmého priemetu na vektor \mathbf{b}'_0 , ktoré oba ležia v podpriestore \mathbb{U} , teda aj vektor \mathbf{b}'_1 bude ležať v podpriestore \mathbb{U} .

$$\mathbf{b}'_1 = \mathbf{b}_1 - \tilde{\mathbf{b}}_1 = \mathbf{b}_1 - c \cdot \mathbf{b}'_0, \quad \text{kde} \quad c = \frac{\langle \mathbf{b}_1, \mathbf{b}'_0 \rangle}{\langle \mathbf{b}'_0, \mathbf{b}'_0 \rangle}$$

Riešením príkladu je báza $\mathcal{B}' = (\mathbf{b}'_0, \mathbf{b}'_1)$, kde

$$\begin{aligned} \mathbf{b}'_0 &= (1, 2, 0, -1) \\ \mathbf{b}'_1 &= (1, 1, 3, 0) - \tilde{\mathbf{b}}_1 = (1, 1, 3, 0) - 0.5 \cdot (1, 2, 0, -1) = \\ &= (1, 1, 3, 0) - (0.5, 1, 0, -0.5) = (0.5, 0, 3, 0.5) \end{aligned}$$

Z postupu vidíme, že vektor $(0.5, 0, 3, 0.5)$ ortogonálnej bázy, sme dostali ako lineárnu kombináciu vektorov pôvodnej bázy.

Úloha 2.9.4 Nájdite ortogonálnu bázu $\mathcal{B}' = (\mathbf{b}'_0, \mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3)$ vektorového podpriestoru \mathbb{U} priestoru \mathbb{V} , ak viete, že

$$\mathbb{U} = [(1, 2, 3, 4, 5, 6), (3, 4, 2, 1, 3, 4), (5, 4, 1, 0, 1, 2), (1, 3, 4, 5, 6, 1)].$$

3 Modely využívajúce závislosť hodnôt procesu od času

3.1 Vektor ako časová závislosť

V tejto kapitole budeme často pojem vektor zamieňať pojmom proces. Dôvodom je, že teoretické výsledky o vektorových priestoroch, ktoré sme zopakovali v predchádzajúcej kapitole budeme používať pri vyšetrovaní vlastností (analyzovaní) procesov závislých od času. Takéto procesy sa niekedy nazývajú časové rady.

Aby sme si lepšie uvedomili, že objektom nášho skúmania je proces, ktorého hodnoty sú merané postupne, v po sebe nasledujúcich časoch, označíme proces nasledujúco:

$$\mathbf{f} = (f(t_0), f(t_1), \dots, f(t_{N-1}))$$

To znamená, že proces \mathbf{f} chápeme na jednej strane ako vektor, na druhej strane ako funkciu času v okamihoch t_0, t_1, \dots, t_{N-1} . Časové okamihy môžeme tiež považovať za vektor

$$\mathbf{t} = (t_0, t_1, \dots, t_{N-1})$$

Časy môžeme označiť ich poradovým číslom aj keď intervaly medzi nimi nemusia byť rovnaké, tj. ekvidistančné.

Potom vektor \mathbf{f} môžeme zapísať ako

$$\mathbf{f} = (f_0, f_1, \dots, f_{N-1})$$

Základnou úlohou v tejto kapitole bude nájsť v procese nejakú závislosť od času a modelovať proces pomocou tejto časovej funkcie.

Najjednoduchším modelom bude konštantná funkcia.

Príklad 3.1.1 *Nech $\mathbf{f}_0, \mathbf{f}_1, \mathbf{f}_2$ sú vektory tvaru:*

$$\begin{aligned}\mathbf{f}_0 &= (c_0, c_0, \dots, c_0) \\ \mathbf{f}_1 &= (c_1, c_1, \dots, c_1) \\ \mathbf{f}_2 &= (c_2, c_2, \dots, c_2)\end{aligned}$$

Zistite, koľkorozmerný je priestor, ktorý určujú tieto vektory.

Riešenie:

Lubovoľný vektor \mathbf{f}_n spomedzi vektorov $\mathbf{f}_0, \mathbf{f}_1, \mathbf{f}_2$ môžeme vyjadriť ako:

$$\begin{aligned}\mathbf{f}_n &= c_n \cdot \mathbf{b}_0 = c_n \cdot (1, 1, \dots, 1) = (c_n, \dots, c_n) \\ \mathbf{f}_n &= (c_n)\mathcal{B}\end{aligned}$$

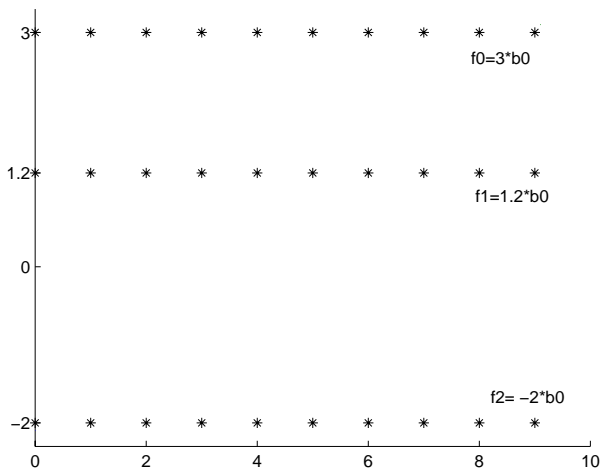
Vidíme, že každý z vektorov \mathbf{f}_n sa dá napísať ako násobok vektora \mathbf{b}_0 , a teda ide o jednorozmerný podpriestor generovaný bázou

$$\mathcal{B} = (\mathbf{b}_0) = ((1, 1, \dots, 1))$$

□

V príklade 3.1.1 boli zadané vektory, z ktorých každý sa dal vyjadriť v tvare $c \cdot (1, 1, \dots, 1)$. To znamená, že všetky tieto vektory ležia v podpriestore generovanom vektorom $(1, 1, \dots, 1)$. Keď si tento podpriestor predstavíme ako priebeh procesu v čase, dostávame konštantný priebeh.

Funkcia času zodpovedajúca tomuto modelu je $f(t_k) = c$. Na obrázku 3.1 sú tri vektory, ktoré ležia v podpriestore \mathbb{V} generovanom vektorom $(1, 1, \dots, 1)$, s hodnotami $c_0 = 3, c_1 = 1.2, c_2 = -2$.



Obr. 3.1. Násobky vektora $(1, 1, \dots, 1)$

Konštanty $c_0 = 3, c_1 = 1.2, c_2 = -2$ a vektor $(1, 1, \dots, 1)$ úplne popisujú procesy $\mathbf{f}_0, \mathbf{f}_1, \mathbf{f}_2$. Teda na popis týchto vektorov stačí omnoho menší počet hodnôt (bitov).

Nie každý vektor sa dá vyjadriť ako násobok vektora $(1, 1, \dots, 1)$. Keď sa budeme snažiť popísať menším počtom hodnôt iné vektory, dopustíme sa chyby. Nevyjadríme presne pôvodný vektor, ale len jeho aproximáciu. V nasledujúcom odseku ukážeme, ako zabezpečiť, aby sme sa dopustili čo najmenšej chyby.

3.2 Aproximácia vektora v podpriestore

V tomto odseku riešime úlohu ako v podpriestore \mathbb{U} nájsť vektor \mathbf{y} , ktorý leží v \mathbb{U} a najlepšie aproximuje zadaný vektor \mathbf{v} .

Hovoríme, že vektor \mathbf{y} ležiaci v podpriestore \mathbb{U} **sa najviac podobá na proces \mathbf{v}** (alebo vektor \mathbf{y} je najbližšie k vektoru \mathbf{v} , vektor \mathbf{y} aproximuje vektor \mathbf{v}), práve vtedy,

ak vzdialenosť medzi týmito vektormi je minimálna:

$$d = \min_{\mathbf{f} \in U} \{d(\mathbf{f}, \mathbf{y})\}$$

Jednou z metód ako úlohu riešiť je použitie diferenciálneho počtu, ako ukazuje nasledujúci príklad.

Príklad 3.2.1 *Nájdite vektor \mathbf{y} ležiaci na osi x , taký, že je najlepšou aproximáciou vektora $\mathbf{v} = (4, 2)$.*

Riešenie:

Vektory na osi x sú násobky vektora $\mathbf{b} = (1, 0)$, teda majú tvar $(c, 0)$, kde c je ľubovoľné reálne číslo.

Najbližšie k vektoru $\mathbf{v} = (4, 2)$ bude ten z vektorov $(c, 0)$, ktorý minimalizuje vzdialenosť medzi nimi

$$d(\mathbf{v}, \mathbf{y}) \rightarrow \min$$

$$d((4, 2), (c, 0)) \rightarrow \min$$

V Euklidovom vektorovom priestore budeme minimalizovať vzdialenosť

$$\sqrt{(4-c)^2 + (2-0)^2}$$

Hľadáme bod c , v ktorom funkcia $\sqrt{(4-c)^2 + (2-0)^2}$ dosiahne minimum.

Stačí ak nájdeme minimum funkcie $g(c) = (4-c)^2 + (2-0)^2$, nakoľko druhá odmocnina je na príslušnej oblasti rastúca funkcia.

$$g'(c) = 2 \cdot (4-c)(-1), \quad g''(c) = 2$$

Odtiaľ dostávame, že $c = 4$ je stacionárny bod a kladná druhá derivácia zabezpečí, že je to minimum.

Teda najbližšie k vektoru $\mathbf{v} = (4, 2)$ je na osi x vektor $\mathbf{y} = (4, 0)$.

Tento výsledok sa dal dosiahnuť intuitívne. Rovnakým spôsobom teraz môžete vyriešiť nasledujúcu úlohu.

Úloha 3.2.1 *Nájdite vektor \mathbf{y} , ležiaci v podpriestore generovanom vektorom $\mathbf{b} = (1, 1, 1, 1, 1)$ taký, že je najlepšou aproximáciou vektora $\mathbf{v} = (4, 2, 3, 2, 3)$.*

Metóda, ktorú sme použili v príklade 3.2.1 sa značne skomplikuje, keď bude podpriestor, v ktorom aproximáciu hľadáme, viacrozmerný. Namiesto extrému funkcie jednej premennej bude potrebné nájsť minimum funkcie viac premenných, čo môže byť náročná úloha.

Príklad 3.2.2 *Nech je proces \mathbf{f} zadaný 100 hodnotami. Ako tento proces popíšeme len tromi hodnotami c_0, c_1, c_2 ?*

Riešenie:

Proces \mathbf{f} vyjadríme vo vektorovom priestore, ktorého báza je $\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2)$ a c_0, c_1, c_2 sú koeficienty v báze \mathcal{B} .

$$\mathbf{f} = (f_0, f_1, \dots, f_{100}) = c_0 \mathbf{b}_0 + c_1 \mathbf{b}_1 + c_2 \mathbf{b}_2$$

Vektor \mathbf{f} je vyjadrený v trojrozmernom vektorovom priestore určenom bázou \mathcal{B} . Niektoré 100 zložkové vektory naozaj možno napísať ako lineárnu kombináciu vektorov z \mathcal{B} . Nie vždy je to však možné. Vtedy sa snažíme nájsť v priestore určenom bázou \mathcal{B} taký vektor $\tilde{\mathbf{f}}$, ktorý je najbližší k vektoru \mathbf{f} . □

Existuje však aj iná metóda, vychádzajúca z geometrického pravidla, že kolmá vzdialenosť je najkratšia, alebo, že odvesna v pravouhlom trojuholníku je vždy kratšia ako prepona (ako vidno na obrázku 3.2). V nasledujúcom tvrdení je dokázané, že v Euklidovom vektorovom priestore je k vektoru \mathbf{f} , spomedzi všetkých vektorov v podpriestore, najbližšie jeho kolmý priemet do tohto podpriestoru.

Tvrdenie 3.2.1 *Nech \mathbb{V} je N -rozmerný vektorový priestor a \mathbb{U} je jeho M -rozmerný vektorový podpriestor.*

Nech $\mathbf{f} \in \mathbb{V}$. Spomedzi vektorov $\mathbf{y} \in \mathbb{U}$ bude najbližší k vektoru \mathbf{f} jeho kolmý priemet do podpriestoru \mathbb{U} , teda vektor $\tilde{\mathbf{f}} \in \mathbb{U}$, taký, že $\mathbf{f} - \tilde{\mathbf{f}}$ je kolmý na všetky vektory \mathbb{U} .

Dôkaz:

Nech \mathbf{y} je ľubovoľný vektor patriaci do \mathbb{U} , potom vektor $\mathbf{y} - \tilde{\mathbf{f}}$ tiež patrí do \mathbb{U} . Vektor $\mathbf{f} - \tilde{\mathbf{f}}$ je kolmý na každý vektor z \mathbb{U} , teda aj na vektor $\mathbf{y} - \tilde{\mathbf{f}}$. Podľa Pytagorovej vety pre kolmé vektory $\mathbf{f} - \tilde{\mathbf{f}}$ a $\mathbf{y} - \tilde{\mathbf{f}}$ platí

$$\|(\mathbf{f} - \tilde{\mathbf{f}}) - (\mathbf{y} - \tilde{\mathbf{f}})\|^2 = \|\mathbf{f} - \tilde{\mathbf{f}}\|^2 + \|\mathbf{y} - \tilde{\mathbf{f}}\|^2$$

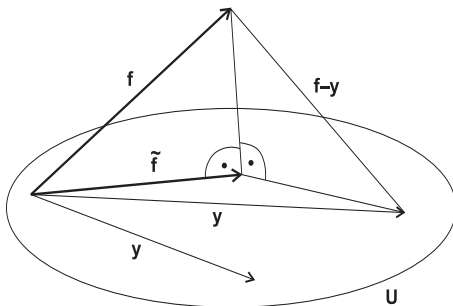
$$\|\mathbf{f} - \mathbf{y}\|^2 = \|\mathbf{f} - \tilde{\mathbf{f}}\|^2 + \|\mathbf{y} - \tilde{\mathbf{f}}\|^2$$

$$d(\mathbf{f}, \mathbf{y}) = d(\mathbf{f}, \tilde{\mathbf{f}}) + d(\mathbf{y}, \tilde{\mathbf{f}})$$

Pretože vzdialenosť je nezáporná, $d(\mathbf{y}, \tilde{\mathbf{f}}) \geq 0$, platí

$$d(\mathbf{f}, \mathbf{y}) \geq d(\mathbf{f}, \tilde{\mathbf{f}})$$

Vektor \mathbf{y} bol ľubovoľný, preto platí, že $d(\mathbf{f}, \tilde{\mathbf{f}})$ je minimálna. □



Obr. 3.2. Najbližší k vektoru \mathbf{f} spomedzi vektorov $\mathbf{y} \in U$ je vektor $\tilde{\mathbf{f}}$

Príklad 3.2.3 *Nájdite v podpriestore $\mathbb{U} = [(1, 0, 2), (-2, 1, 1)]$ vektor, ktorý je najbližšie k vektoru $\mathbf{f} = (1, 2, 3)$.*

Riešenie:

Najbližšie k vektoru $\mathbf{f} = (1, 2, 3)$ bude vektor $\tilde{\mathbf{f}}$, pre ktorý platí

$$\tilde{\mathbf{f}} = c_0 \cdot (1, 0, 2) + c_1 \cdot (-2, 1, 1)$$

$$(\mathbf{f} - \tilde{\mathbf{f}}) \perp (1, 0, 2)$$

$$(\mathbf{f} - \tilde{\mathbf{f}}) \perp (-2, 1, 1)$$

Podmienky môžeme upraviť na tvar

$$\tilde{\mathbf{f}} = c_0 \cdot (1, 0, 2) + c_1 \cdot (-2, 1, 1)$$

$$\langle \mathbf{f} - \tilde{\mathbf{f}}, (1, 0, 2) \rangle = 0$$

$$\langle \mathbf{f} - \tilde{\mathbf{f}}, (-2, 1, 1) \rangle = 0$$

Odtiaľ po dosadení za \mathbf{f}

$$\langle (1, 2, 3), (1, 0, 2) \rangle = c_0 \cdot \langle (1, 0, 2), (1, 0, 2) \rangle + c_1 \cdot \langle (-2, 1, 1), (1, 0, 2) \rangle$$

$$\langle (1, 2, 3), (-2, 1, 1) \rangle = c_0 \cdot \langle (1, 0, 2), (-2, 1, 1) \rangle + c_1 \cdot \langle (-2, 1, 1), (-2, 1, 1) \rangle$$

Teda platí

$$7 = c_0 \cdot 5 + c_1 \cdot 0$$

$$3 = c_0 \cdot 0 + c_1 \cdot 6$$

Nakoniec

$$\tilde{\mathbf{f}} = \frac{7}{5} \cdot (1, 0, 2) + \frac{3}{6} \cdot (-2, 1, 1) = (0.4, 0.5, 3.3)$$

□

Úloha 3.2.2 Nájdite v podpriestore $\mathbb{U} = [(1, 0, 2), (-2, 1, 1)]$ vektor, ktorý je najbližšie k vektoru $\mathbf{f} = (-5, 2, 0)$.

3.3 Metóda najmenších štvorcov

Pomerne známou metódou na výpočet krivky, ktorá čo najlepšie aproximuje dáta, je metóda najmenších štvorcov. Úloha je zvyčajne formulovaná pre priamku a riešená metódami diferenciálneho počtu. Aj v tomto prípade však môžeme úlohu previesť do jazyka vektorových priestorov. Minimalizáciou vzdialenosti medzi vektorom Euklidovského vektorového priestoru \mathbf{f} a jeho priemetom $\tilde{\mathbf{f}}$ do podpriestoru dostaneme metódu najmenších štvorcov. Na obrázku 3.3 je zobrazený priemet do podpriestoru, ktorý určuje priamku $y = a + b \cdot t$.

Nech:

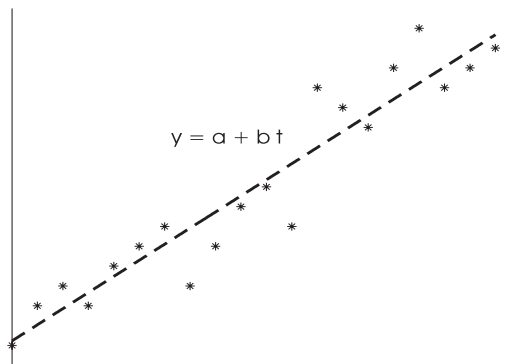
$$\mathbf{f} = (f_0, f_1, \dots, f_{N-1}), \quad \tilde{\mathbf{f}} = (\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_{N-1})$$

Potom

$$d(\mathbf{f}, \tilde{\mathbf{f}}) = \|\mathbf{f} - \tilde{\mathbf{f}}\| = \sqrt{\langle \mathbf{f} - \tilde{\mathbf{f}}, \mathbf{f} - \tilde{\mathbf{f}} \rangle} =$$

$$= \sqrt{(f_0 - \tilde{f}_0)^2 + \dots + (f_{N-1} - \tilde{f}_{N-1})^2} = \sqrt{\sum_{k=0}^{N-1} (f_k - \tilde{f}_k)^2} \rightarrow \min$$

$$d(\mathbf{f}, \tilde{\mathbf{f}})^2 = \sum_{k=0}^{N-1} (f_k - \tilde{f}_k)^2 \rightarrow \min$$



Obr. 3.3. Dáta aproximované priamkou

3.4 Chyba aproximácie

Pri aproximácii procesu nejakým iným procesom je dôležité vedieť, akej chyby sme sa pri odhade dopustili. Pokiaľ poznáme oba procesy, pôvodný aj jeho odhad, vieme presne povedať aká je chyba odhadu.

Nech $\mathbf{f} = (f_0, f_1, \dots, f_{N-1})\mathcal{E}$ je proces vyjadrený v jednotkovej báze N -rozmerného vektorového priestoru. Hodnotu $\|\mathbf{f}\|^2 = (d(\mathbf{f}, \mathbf{0}))^2$ nazveme **energia procesu \mathbf{f}** .

Nech $\tilde{\mathbf{f}}$ je aproximácia procesu \mathbf{f} v podpriestore určenom bázou \mathcal{B} . Vektor $\mathbf{e} = (e_0, e_1, \dots, e_{N-1}) = \mathbf{f} - \tilde{\mathbf{f}}$ nazveme **chybový proces**.

Vzdialenosť

$$d(\mathbf{f}, \tilde{\mathbf{f}}) = \sqrt{\sum_{k=0}^{N-1} (f_k - \tilde{f}_k)^2} = \sqrt{\sum_{k=0}^{N-1} e_k^2} = \|\mathbf{e}\|$$

nazveme **veľkosť chybového procesu**.

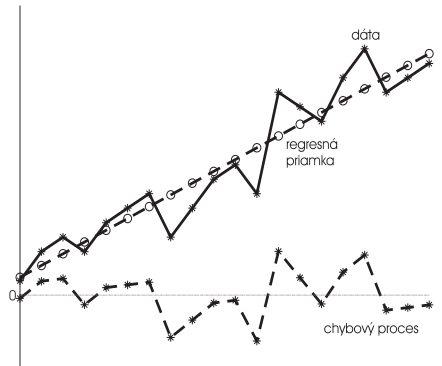
$$\tilde{\mathbf{f}} = (c_0, c_1, \dots, c_{M-1})\mathcal{B}$$

$$\tilde{\mathbf{f}} = (\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_{N-1})\mathcal{E}$$

$$\mathbf{e} = (e_0, e_1, \dots, e_{N-1})\mathcal{E}$$

$$d^2(\mathbf{f}, \tilde{\mathbf{f}}) = \sum_{k=0}^{N-1} e_k^2$$

Hodnota $\|\mathbf{e}\|^2$ je **energia chybového procesu**.



Obr. 3.4: Pôvodný proces, jeho priemet do dvojrozmerného podpriestoru, chybový proces (proces po odstránení lineárneho trendu)

Príklad 3.4.1 *Nech aproximácia časti procesu $\mathbf{y} = (1, 3, 4, 3, 5, 6, 7, 4, 6, 8)$ z obrázku 3.4 je približne proces*

$$\tilde{\mathbf{y}} = (1.23, 2.04, 2.85, 3.67, 4.48, 5.29, 6.11, 6.92, 7.73, 8.54)$$

Aká je veľkosť chyby?

Riešenie:

Chybový proces je proces

$$\mathbf{e} = (-0.23, 0.96, 1.15, -0.67, 0.52, 0.71, 0.89, -2.92, -1.73, -0.54)$$

Veľkosť chybového procesu je $\|\mathbf{e}\| = 4.0154$ a jeho energia je $\|\mathbf{e}\|^2 = 16.1234$. \square

Úloha 3.4.1 *Nájdite dva procesy, ktoré majú rovnakú veľkosť chyby, ale rôzne chybové procesy.*

Tvrdenie 3.4.1 *Ak sú dva procesy na seba kolmé, potom energia súčtového procesu sa rovná súčtu energií jednotlivých procesov.*

Dôkaz:

Upravíme výraz pre energiu súčtu procesov:

$$\begin{aligned} \|\mathbf{f} + \mathbf{g}\|^2 &= \langle \mathbf{f} + \mathbf{g}, \mathbf{f} + \mathbf{g} \rangle = \langle \mathbf{f}, \mathbf{f} + \mathbf{g} \rangle + \langle \mathbf{g}, \mathbf{f} + \mathbf{g} \rangle \\ &= \langle \mathbf{f}, \mathbf{f} \rangle + \langle \mathbf{f}, \mathbf{g} \rangle + \langle \mathbf{g}, \mathbf{f} \rangle + \langle \mathbf{g}, \mathbf{g} \rangle = \langle \mathbf{f}, \mathbf{f} \rangle + 0 + 0 + \langle \mathbf{g}, \mathbf{g} \rangle = \|\mathbf{f}\|^2 + \|\mathbf{g}\|^2 \end{aligned}$$

Toto tvrdenie je vlastne inak formulovaná Pytagorova veta. \square

Predchádzajúce tvrdenie v tvare

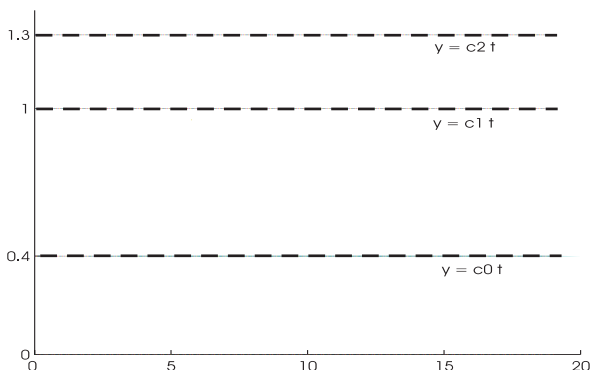
$$\|\mathbf{f}\|^2 = \|\tilde{\mathbf{f}} + \mathbf{e}\|^2 = \|\tilde{\mathbf{f}}\|^2 + \|\mathbf{e}\|^2$$

hovorí o tom, že priemet do podpriestoru má menšiu energiu ako pôvodný proces. Tento fakt sa využíva napríklad pri kompresných metódach.

Zopakujme, že úlohou v tejto kapitole je nájsť v procese nejakú závislosť od času a modelovať proces pomocou tejto časovej funkcie. Modely, ktorými budeme priebeh procesu vyjadrovať, budú lineárne vzhľadom na neznáme koeficienty. Metóda, ktorá k nameraným dátam nájde spomedzi všetkých lineárnych kombinácií nejakých konkrétnych funkcií tú najlepšiu v zmysle Euklidovej metriky, sa nazýva **lineárna regresia**. Nasledujú jednotlivé úlohy lineárnej regresie rozdelené podľa tvaru regresnej krivky.

3.5 Priamka rovnobežná s časovou osou, $y = c$

Je daný proces $\mathbf{f} = (f_0, f_1, \dots, f_{N-1})$ popísaný N hodnotami, ktoré nadobudol v časoch t_0, t_1, \dots, t_{N-1} (tieto časy sú zadané a nemusia byť ekvidištancné). Riešme úlohu popísať tento proces jednou hodnotou, ktorá určí také umiestnenie priamky rovnobežnej s časovou osou, že priamka namerané dáta popisuje najlepšie, viď obrázok 3.5.



Obr. 3.5. Priamky, rovnobežné s časovou osou

Príklad 3.5.1 Chceme teda vyjadriť proces $\mathbf{f} = (f_0, f_1, \dots, f_{N-1})$ ako násobok vektora $\mathbf{b} = (1, 1, \dots, 1)$. Všetky násobky vektora \mathbf{b} tvoria jednorozmerný vektorový podpriestor N -rozmerného vektorového priestoru, z ktorého je vektor \mathbf{f} . Podľa tvrdenia 3.2.1, je najbližšie k vektoru \mathbf{f} spomedzi všetkých vektorov tvaru $c \cdot \mathbf{b}$ taký vektor $\tilde{\mathbf{f}} = c \cdot \mathbf{b}$, ktorý je kolmým priemetom vektora \mathbf{f} do podpriestoru s bázou tvorenou vektorom \mathbf{b} .

$$\mathbf{f} = (f_0, f_1, \dots, f_{N-1})\mathcal{E} \quad \tilde{\mathbf{f}} = (c)\mathcal{B} \quad \tilde{\mathbf{f}} = (c, c, \dots, c)\mathcal{E}$$

$$\tilde{\mathbf{f}} = \mathbf{c} \cdot \mathbf{b}, \quad \text{kde} \quad \mathbf{b} = (1, 1, \dots, 1)$$

$$c = \frac{\langle \mathbf{f}, \mathbf{b} \rangle}{\langle \mathbf{b}, \mathbf{b} \rangle} = \frac{\langle (f_0, f_1, \dots, f_{N-1}), (1, 1, \dots, 1) \rangle}{\langle (1, 1, \dots, 1), (1, 1, \dots, 1) \rangle} = \frac{\sum_{k=0}^{N-1} f_k}{N} = \frac{1}{N} \sum_{k=0}^{N-1} f_k$$

Hodnota, ktorú sme dostali, je aritmetický priemer všetkých hodnôt procesu \mathbf{f} . Uvedeným modelom sú dobre popísané procesy, ktorých hodnoty sú približne rovnaké.

Na internetovej stránke <http://www.adis-computers.sk> sme našli štatistiku návštevnosti stránky za posledných 100 dní. Ide o počet unikátnych IP adries denne. Namerané hodnoty procesu sú

$\mathbf{f} = (20, 34, 20, 35, 24, 16, 34, 29, 28, 27, 25, 20, 22, 28, 34, 33, 30, 32, 27, 20,$
 $36, 31, 33, 39, 30, 24, 33, 28, 43, 33, 30, 27, 22, 25, 29, 26, 28, 29, 24, 8,$
 $14, 30, 28, 21, 20, 31, 9, 28, 31, 18, 29, 33, 30, 20, 21, 26, 28, 33, 27, 28,$
 $15, 29, 32, 36, 33, 43, 24, 21, 25, 33, 40, 30, 31, 31, 17, 27, 24, 33, 26, 24,$
 $23, 22, 14, 26, 38, 27, 27, 30, 13, 17, 25, 31, 30, 31, 26, 14, 20, 26, 32, 30)$

Aproximujte tento proces procesom s konštantným priebehom.

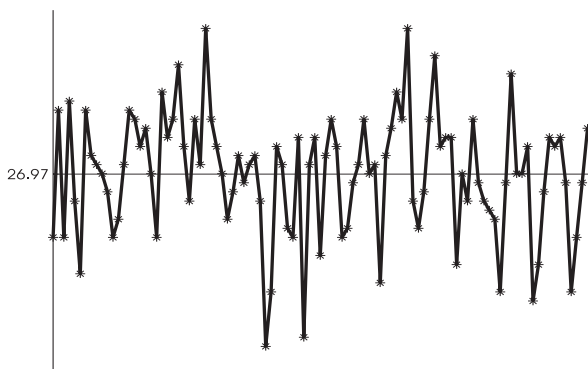
Riešenie:

Konštantnú zložku procesu vyjadríme ako

$$\tilde{\mathbf{f}} = \mathbf{c} \cdot \mathbf{b} = \mathbf{c} \cdot (1, 1, \dots, 1) \quad \text{kde} \quad c = \frac{\langle \mathbf{f}, \mathbf{b} \rangle}{\langle \mathbf{b}, \mathbf{b} \rangle}$$

$$c = \frac{1}{100} \sum_{k=0}^{99} f_k = 26.97$$

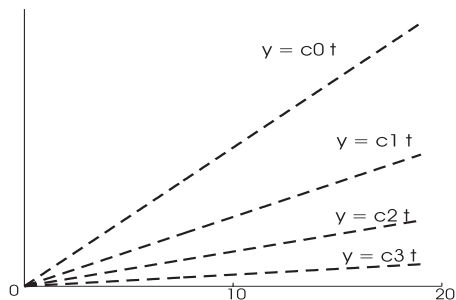
Priebeh procesu a jeho aproximácia konštantným procesom sú zobrazené na obrázku 3.6.



Obr. 3.6. Návštevnosť internetovej stránky a jej aproximácia konštantou

3.6 Priamka prechádzajúca počiatkom, $y = c \cdot t$

Ak hodnoty procesu rastú lineárne a priamka lineárnej závislosti prechádza počiatkom, bude mať proces $\tilde{\mathbf{f}}$ hodnoty $\tilde{f}_k = c \cdot t_k$. Príklady takýchto priamok sú na obrázku 3.7.



Obr. 3.7. Sklon priamky $y = c \cdot t$ závisí od koeficientu c

Teda proces s nameranými hodnotami

$$\mathbf{f} = (f_0, f_1, \dots, f_{N-1})\mathcal{E},$$

ktoré nadobudol v časoch t_0, t_1, \dots, t_{N-1} , popíšeme jednou hodnotou c

$$\tilde{\mathbf{f}} = (\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_{N-1})\mathcal{E} = (c \cdot t_0, c \cdot t_1, \dots, c \cdot t_{N-1}) = c(t_0, t_1, \dots, t_{N-1})$$

$$\tilde{\mathbf{f}} = c \cdot \mathbf{b} = (c \cdot t_0, c \cdot t_1, \dots, c \cdot t_{N-1})$$

Ku procesu \mathbf{f} hľadáme taký proces $\tilde{\mathbf{f}}$, pre ktorý platí

$$d(\mathbf{f}, \tilde{\mathbf{f}}) \rightarrow \min$$

Nahradíme proces $\tilde{\mathbf{f}} = c \cdot \mathbf{b}$, pretože jeho hodnoty ležia na priamke:

$$d(\mathbf{f}, c \cdot \mathbf{b}) \rightarrow \min$$

Na vypočítanie koeficientu c použijeme podmienku $(\mathbf{f} - c \cdot \mathbf{b}) \perp \mathbf{b}$. Podmienka ortogonalnosti vyjadrená pomocou skalárneho súčinu má tvar:

$$\langle \mathbf{f} - c \cdot \mathbf{b}, \mathbf{b} \rangle = 0$$

Využijeme vlastnosti skalárneho súčinu a dostaneme

$$\langle (f_0 - c \cdot t_0, f_1 - c \cdot t_1, \dots, f_{N-1} - c \cdot t_{N-1}), (t_0, t_1, \dots, t_{N-1}) \rangle = 0$$

$$c = \frac{\langle \mathbf{f}, \mathbf{b} \rangle}{\langle \mathbf{b}, \mathbf{b} \rangle} = \frac{\sum_{k=0}^{N-1} f_k \cdot t_k}{\sum_{k=0}^{N-1} t_k^2}$$

Príklad 3.6.1 Na internete sme našli históriu vývoja devízových rezerv Národnej banky Slovenska za obdobie od 29. 1. 1993 do 28. 2. 2001. Sledované hodnoty sa zapisovali a sledovali vždy k poslednému pracovnému dňu v danom mesiaci. Namerané hodnoty procesu sú v miliónoch USD:

$\mathbf{f} = (197, 176, 184, 175, 256, 243, 336, 386, 567, 538, 508, 450, 401, 382, 475, 534, 610, 689, 1125, 1294, 1464, 1565, 1598, 1745, 17411813, 1969, 2022, 2186, 2622, 2630, 2708, 2813, 2873, 3046, 3418, 3307, 3398, 3459, 3407, 3355, 3377, 3504, 3677, 3655, 3602, 3595, 3473, 3434, 3472, 3453, 3347, 2974, 3019, 3010, 3181, 3151, 3411, 3446, 3285, 3161, 3202, 3143, 3349, 3723, 3790, 3770, 3622, 3110, 2987, 2939, 2923, 2860, 2910, 2814, 2732, 2515, 2953, 2864, 2805, 2935, 2952, 2869, 3425, 3444, 3564, 3699, 4198, 4085, 4037, 3988, 4433, 4214, 3994, 4024, 4101, 3973, 3971, 3890, 4074)$

Aproximujte tento proces procesom s lineárnym nárastom s jedným parametrom.

Riešenie:

Lineárny nárast procesu vyjadríme pomocou jedného koeficientu ako

$$\tilde{\mathbf{f}} = c \cdot \mathbf{b} = c \cdot (0, 1, 2, \dots, 99) \quad \text{kde} \quad c = \frac{\langle \mathbf{f}, \mathbf{b} \rangle}{\langle \mathbf{b}, \mathbf{b} \rangle}$$

$$c = \frac{\sum_{k=0}^{99} k \cdot f_k}{\sum_{k=0}^{99} k^2} = 49.0672$$

Priebeh procesu a jeho aproximácia lineárnym procesom s jedným parametrom sú zobrazené na obrázku 3.8.

Úloha 3.6.1 Vyhľadajte na internete stav devízových rezerv NBS k dnešnému dátumu a porovnajte túto hodnotu s odhadom, ktorý urobíte pomocou modelu z príkladu 3.6.1.

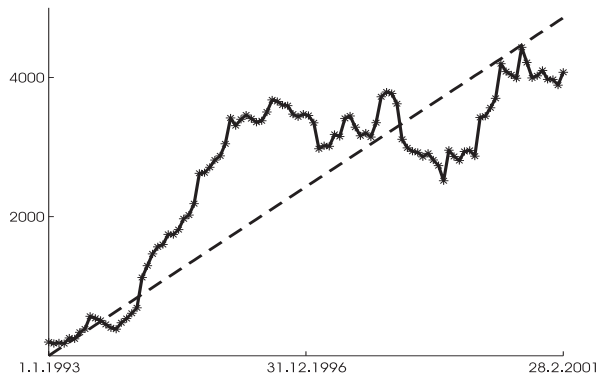
3.7 Priamka neprechádzajúca počiatkom,

$$y = c_0 + c_1 \cdot t$$

Ak hodnoty procesu rastú lineárne a priamka lineárnej závislosti neprechádza počiatkom, bude mať proces $\tilde{\mathbf{f}}$ hodnoty $\tilde{f}_k = c_0 + c_1 \cdot t_k$. Predpokladáme, že proces s nameranými hodnotami $\mathbf{f} = (f_0, f_1, \dots, f_{N-1})$, ktoré nadobudol v časoch t_0, t_1, \dots, t_{N-1} , ležia približne na priamke, ktorá neprechádza počiatkom.

Budeme ich aproximovať hodnotami

$$\tilde{\mathbf{f}} = (\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_{N-1}) = (c_0 + c_1 \cdot t_0, c_0 + c_1 \cdot t_1, \dots, c_0 + c_1 \cdot t_{N-1})$$



Obr. 3.8. Devízové rezervy NBS a ich aproximácia priamkou s jedným parametrom

$$f_k = c_0 + c_1 \cdot t_k$$

$$\tilde{\mathbf{f}} = (c_0, c_1)\mathcal{B}$$

$$\begin{aligned}\tilde{\mathbf{f}} &= c_0(1, 1, \dots, 1) + c_1(t_0, t_1, \dots, t_{N-1})\mathcal{E} = \\ &= (c_0 + c_1 \cdot t_0, c_0 + c_1 \cdot t_1, \dots, c_0 + c_1 \cdot t_{N-1}) = c_0 \cdot \mathbf{b}_0 + c_1 \cdot \mathbf{b}_1\end{aligned}$$

kde báza

$$\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1), \text{ kde } \mathbf{b}_0 = (1, 1, \dots, 1), \mathbf{b}_1 = (t_0, t_1, \dots, t_{N-1})$$

Procesy \mathbf{b}_0 a \mathbf{b}_1 nie sú na seba kolmé. Hľadáme proces $\tilde{\mathbf{f}}$, ktorý leží v podpriestore \mathbb{U} určenom bázou \mathcal{B} a je najbližší k procesu \mathbf{f} . Ukázali sme, že to bude proces $\tilde{\mathbf{f}}$, pre ktorý platí, že $\mathbf{f} - \tilde{\mathbf{f}}$ bude kolmý na \mathbb{U} . To znamená, že bude kolmý na každý proces z \mathbb{U} , teda aj na každý proces bázy $\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1)$. Teda proces $\tilde{\mathbf{f}} = c_0\mathbf{b}_0 + c_1\mathbf{b}_1$ a $\mathbf{f} - \tilde{\mathbf{f}} \perp \mathbb{U}$. Vychádzame zo vzťahov $\mathbf{f} - \tilde{\mathbf{f}} \perp \mathbf{b}_0$ a $\mathbf{f} - \tilde{\mathbf{f}} \perp \mathbf{b}_1$.

Upravíme:

$$\langle \mathbf{f} - \tilde{\mathbf{f}}, \mathbf{b}_0 \rangle = 0; \quad \langle \mathbf{f} - \tilde{\mathbf{f}}, \mathbf{b}_1 \rangle = 0$$

$$\langle \mathbf{f} - c_0 \cdot \mathbf{b}_0 - c_1 \cdot \mathbf{b}_1, \mathbf{b}_0 \rangle = 0$$

$$\langle \mathbf{f} - c_0 \cdot \mathbf{b}_0 - c_1 \cdot \mathbf{b}_1, \mathbf{b}_1 \rangle = 0$$

$$\langle \mathbf{f}, \mathbf{b}_0 \rangle - \langle c_0 \mathbf{b}_0, \mathbf{b}_0 \rangle - \langle c_1 \mathbf{b}_1, \mathbf{b}_0 \rangle = 0$$

$$\langle \mathbf{f}, \mathbf{b}_1 \rangle - \langle c_0 \mathbf{b}_0, \mathbf{b}_1 \rangle - \langle c_1 \mathbf{b}_1, \mathbf{b}_1 \rangle = 0$$

$$\langle \mathbf{f}, \mathbf{b}_0 \rangle - c_0 \langle \mathbf{b}_0, \mathbf{b}_0 \rangle - c_1 \langle \mathbf{b}_1, \mathbf{b}_0 \rangle = 0$$

$$\langle \mathbf{f}, \mathbf{b}_1 \rangle - c_0 \langle \mathbf{b}_0, \mathbf{b}_1 \rangle - c_1 \langle \mathbf{b}_1, \mathbf{b}_1 \rangle = 0$$

$$c_0 \langle \mathbf{b}_0, \mathbf{b}_0 \rangle + c_1 \langle \mathbf{b}_1, \mathbf{b}_0 \rangle = \langle \mathbf{f}, \mathbf{b}_0 \rangle$$

$$c_0 \langle \mathbf{b}_0, \mathbf{b}_1 \rangle + c_1 \langle \mathbf{b}_1, \mathbf{b}_1 \rangle = \langle \mathbf{f}, \mathbf{b}_1 \rangle$$

Dostaneme sústavu dvoch lineárnych rovníc o dvoch neznámych.

$$\begin{aligned}c_0 \cdot N + c_1 \cdot \sum_{k=0}^{N-1} t_k &= \sum_{k=0}^{N-1} f_k \\ c_0 \cdot \sum_{k=0}^{N-1} t_k + c_1 \cdot \sum_{k=0}^{N-1} t_k^2 &= \sum_{k=0}^{N-1} f_k \cdot t_k\end{aligned}$$

Riešením tejto sústavy budú koeficienty procesu $\tilde{\mathbf{f}}$ v báze $\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1)$.

Príklad 3.7.1 *Aproximujte proces z príkladu 3.6.1 procesom s lineárnym nárastom s dvomi parametrami.*

Riešenie:

Lineárny nárast procesu vyjadríme pomocou dvoch koeficientov ako

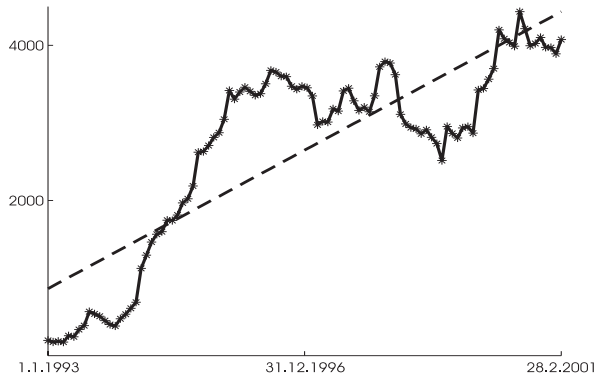
$$\tilde{\mathbf{f}} = c_0 \cdot \mathbf{b}_0 + c_1 \cdot \mathbf{b}_1 = c_0 \cdot (1, 1, \dots, 1) + c_1 \cdot (0, 1, 2, \dots, 99)$$

Koeficienty c_0, c_1 vypočítame ako riešenie sústavy rovníc

$$\begin{aligned} 100 \cdot c_0 + c_1 \cdot \sum_{k=0}^{99} k &= \sum_{k=0}^{99} f_k \\ c_0 \cdot \sum_{k=0}^{99} k + c_1 \cdot \sum_{k=0}^{99} k^2 &= \sum_{k=0}^{99} f_k \cdot k \\ 100 \cdot c_0 + 4950 \cdot c_1 &= 264773 \\ 4950 \cdot c_0 + 328350 \cdot c_1 &= 16111220 \end{aligned}$$

Riešením tejto sústavy sú koeficienty $c_0 = 862.61, c_1 = 36.06$.

Priebeh procesu a jeho aproximácia lineárnym procesom s dvomi parametrami sú zobrazené na obrázku 3.9.



Obr. 3.9. Devízové rezervy NBS a ich aproximácia priamkou s dvomi parametrami

Úloha 3.7.1 *Stav devízových rezerv NBS k dnešnému dátumu porovnajte s odhadom, ktorý urobíte pomocou modelu z príkladu 3.7.1. Chyba tohto odhadu bude zrejme menšia, ako chyba odhadu z úlohy 3.6.1.*

3.8 Parabola, $y = c_0 + c_1 \cdot t + c_2 \cdot t^2$

Bežne sa stáva, že hodnoty procesu nemajú lineárny trend, ale niekedy rastú a inokedy zase klesajú. Niekedy je vhodné takéto hodnoty aproximovať kvadratickou krivkou, nazývanou aj parabola. V takomto prípade bude mať proces $\tilde{\mathbf{f}}$ hodnoty $\tilde{f}_k = c_0 + c_1 \cdot t_k +$

$c_1 \cdot t_k^2$. Predpokladáme teda, že proces s nameranými hodnotami $\mathbf{f} = (f_0, f_1, \dots, f_{N-1})$, ktoré nadobudol v časoch t_0, t_1, \dots, t_{N-1} , ležia približne na parabole. Táto parabola vo všeobecnosti nemusí prechádzať počiatkom.

Budeme ich aproximovať hodnotami

$$\begin{aligned}\tilde{\mathbf{f}} &= (\tilde{f}_0, \tilde{f}_1, \dots, \tilde{f}_{N-1}) \quad \text{kde dosadíme } \tilde{f}_k = c_0 + c_1 \cdot t_k + c_2 \cdot t_k^2 \\ &= (c_0 + c_1 \cdot t_0 + c_2 \cdot t_0^2, c_0 + c_1 \cdot t_1 + c_2 \cdot t_1^2, \dots, c_0 + c_1 \cdot t_{N-1} + c_2 \cdot t_{N-1}^2)\end{aligned}$$

$$\tilde{\mathbf{f}}_{\mathcal{B}} = (c_0, c_1, c_2)$$

$$\begin{aligned}\tilde{\mathbf{f}}_{\mathcal{E}} &= c_0(1, 1, \dots, 1) + c_1(t_0, t_1, \dots, t_{N-1}) + c_2(t_0^2, t_1^2, \dots, t_{N-1}^2) = \\ &= (c_0 + c_1 \cdot t_0 + c_2 \cdot t_0^2, c_0 + c_1 \cdot t_1 + c_2 \cdot t_1^2, \dots, c_0 + c_1 \cdot t_{N-1} + c_2 \cdot t_{N-1}^2) \\ &= c_0 \cdot \mathbf{b}_0 + c_1 \cdot \mathbf{b}_1 + c_2 \cdot \mathbf{b}_2\end{aligned}$$

kde báza $\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2)$, pozostáva z vektorov

$$\begin{aligned}\mathbf{b}_0 &= (1, 1, \dots, 1), \\ \mathbf{b}_1 &= (t_0, t_1, \dots, t_{N-1}), \\ \mathbf{b}_2 &= (t_0^2, t_1^2, \dots, t_{N-1}^2).\end{aligned}$$

Procesy \mathbf{b}_0 , \mathbf{b}_1 a \mathbf{b}_2 nie sú na seba kolmé. Hľadáme proces $\tilde{\mathbf{f}}$, ktorý leží v podpriestore \mathbb{U} určenom bázou \mathcal{B} a je najbližší k procesu \mathbf{f} . Ukázali sme, že to bude proces $\tilde{\mathbf{f}}$, pre ktorý platí, že $\mathbf{f} - \tilde{\mathbf{f}}$ bude kolmý na \mathbb{U} . To znamená, že bude kolmý na každý proces z \mathbb{U} , teda aj na každý proces bázy $\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2)$. Teda proces $\tilde{\mathbf{f}} = c_0 \mathbf{b}_0 + c_1 \mathbf{b}_1 + c_2 \mathbf{b}_2$ a $\mathbf{f} - \tilde{\mathbf{f}} \perp \mathbb{U}$.

Vychádzame zo vzťahov $\mathbf{f} - \tilde{\mathbf{f}} \perp \mathbf{b}_0$, $\mathbf{f} - \tilde{\mathbf{f}} \perp \mathbf{b}_1$ a $\mathbf{f} - \tilde{\mathbf{f}} \perp \mathbf{b}_2$.

Upravíme:

$$\langle \mathbf{f} - \tilde{\mathbf{f}}, \mathbf{b}_0 \rangle = 0; \quad \langle \mathbf{f} - \tilde{\mathbf{f}}, \mathbf{b}_1 \rangle = 0; \quad \langle \mathbf{f} - \tilde{\mathbf{f}}, \mathbf{b}_2 \rangle = 0$$

$$\begin{aligned}\langle \mathbf{f} - c_0 \cdot \mathbf{b}_0 - c_1 \cdot \mathbf{b}_1 - c_2 \cdot \mathbf{b}_2, \mathbf{b}_0 \rangle &= 0 \\ \langle \mathbf{f} - c_0 \cdot \mathbf{b}_0 - c_1 \cdot \mathbf{b}_1 - c_2 \cdot \mathbf{b}_2, \mathbf{b}_1 \rangle &= 0 \\ \langle \mathbf{f} - c_0 \cdot \mathbf{b}_0 - c_1 \cdot \mathbf{b}_1 - c_2 \cdot \mathbf{b}_2, \mathbf{b}_2 \rangle &= 0\end{aligned}$$

$$\begin{aligned}\langle \mathbf{f}, \mathbf{b}_0 \rangle - \langle c_0 \mathbf{b}_0, \mathbf{b}_0 \rangle - \langle c_1 \mathbf{b}_1, \mathbf{b}_0 \rangle - \langle c_2 \mathbf{b}_2, \mathbf{b}_0 \rangle &= 0 \\ \langle \mathbf{f}, \mathbf{b}_1 \rangle - \langle c_0 \mathbf{b}_0, \mathbf{b}_1 \rangle - \langle c_1 \mathbf{b}_1, \mathbf{b}_1 \rangle - \langle c_2 \mathbf{b}_2, \mathbf{b}_1 \rangle &= 0 \\ \langle \mathbf{f}, \mathbf{b}_2 \rangle - \langle c_0 \mathbf{b}_0, \mathbf{b}_2 \rangle - \langle c_1 \mathbf{b}_1, \mathbf{b}_2 \rangle - \langle c_2 \mathbf{b}_2, \mathbf{b}_2 \rangle &= 0\end{aligned}$$

$$\begin{aligned}\langle \mathbf{f}, \mathbf{b}_0 \rangle - c_0 \langle \mathbf{b}_0, \mathbf{b}_0 \rangle - c_1 \langle \mathbf{b}_1, \mathbf{b}_0 \rangle - c_2 \langle \mathbf{b}_2, \mathbf{b}_0 \rangle &= 0 \\ \langle \mathbf{f}, \mathbf{b}_1 \rangle - c_0 \langle \mathbf{b}_0, \mathbf{b}_1 \rangle - c_1 \langle \mathbf{b}_1, \mathbf{b}_1 \rangle - c_2 \langle \mathbf{b}_2, \mathbf{b}_1 \rangle &= 0 \\ \langle \mathbf{f}, \mathbf{b}_2 \rangle - c_0 \langle \mathbf{b}_0, \mathbf{b}_2 \rangle - c_1 \langle \mathbf{b}_1, \mathbf{b}_2 \rangle - c_2 \langle \mathbf{b}_2, \mathbf{b}_2 \rangle &= 0\end{aligned}$$

$$\begin{aligned}c_0 \langle \mathbf{b}_0, \mathbf{b}_0 \rangle + c_1 \langle \mathbf{b}_1, \mathbf{b}_0 \rangle + c_2 \langle \mathbf{b}_2, \mathbf{b}_0 \rangle &= \langle \mathbf{f}, \mathbf{b}_0 \rangle \\ c_0 \langle \mathbf{b}_0, \mathbf{b}_1 \rangle + c_1 \langle \mathbf{b}_1, \mathbf{b}_1 \rangle + c_2 \langle \mathbf{b}_2, \mathbf{b}_1 \rangle &= \langle \mathbf{f}, \mathbf{b}_1 \rangle \\ c_0 \langle \mathbf{b}_0, \mathbf{b}_2 \rangle + c_1 \langle \mathbf{b}_1, \mathbf{b}_2 \rangle + c_2 \langle \mathbf{b}_2, \mathbf{b}_2 \rangle &= \langle \mathbf{f}, \mathbf{b}_2 \rangle\end{aligned}$$

Dostaneme sústavu troch lineárnych rovníc o troch neznámých

$$\begin{aligned} c_0 \cdot N + c_1 \cdot \sum_{k=0}^{N-1} t_k + c_2 \cdot \sum_{k=0}^{N-1} t_k^2 &= \sum_{k=0}^{N-1} f_k \\ c_0 \cdot \sum_{k=0}^{N-1} t_k + c_1 \cdot \sum_{k=0}^{N-1} t_k^2 + c_2 \cdot \sum_{k=0}^{N-1} t_k^3 &= \sum_{k=0}^{N-1} f_k \cdot t_k \\ c_0 \cdot \sum_{k=0}^{N-1} t_k^2 + c_1 \cdot \sum_{k=0}^{N-1} t_k^3 + c_2 \cdot \sum_{k=0}^{N-1} t_k^4 &= \sum_{k=0}^{N-1} f_k \cdot t_k^2 \end{aligned}$$

Riešením tejto sústavy budú koeficienty procesu $\tilde{\mathbf{f}}$ v báze $\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2)$.

Príklad 3.8.1 Na internete sa dá sledovať vývoj hodnôt akciového indexu Dow Jones Industrial Average. Vývoj hodnôt počas turbulentného vývoja po páde americkej banky Lehman Brothers zaznamenal prudký pokles nasledovaný miernejším rastom. Počas obdobia od 1. 7. 2008 do 1. 8. 2009 sme sledovali týždenné hodnoty indexu:

$\mathbf{f} = (11288, 11100, 11496, 11370, 11326, 11734, 11659, 11628, 11543,$
 $11220, 11421, 11388, 11143, 10325, 8451, 8852, 8378, 9325, 8943,$
 $8497, 8046, 8829, 8635, 8629, 8579, 8515, 8515, 9034, 8599, 8281,$
 $8077, 8000, 8280, 7850, 7365, 7062, 6626, 7223, 7278, 7776, 8017,$
 $8083, 8131, 8076, 8212, 8574, 8268, 8277, 8500, 8763, 8799, 8539,$
 $8438, 8280, 8146, 8743, 9093, 9171, 9370)$

Aproximujte tento proces procesom s kvadratickou krivkou s tromi parametrami.

Riešenie:

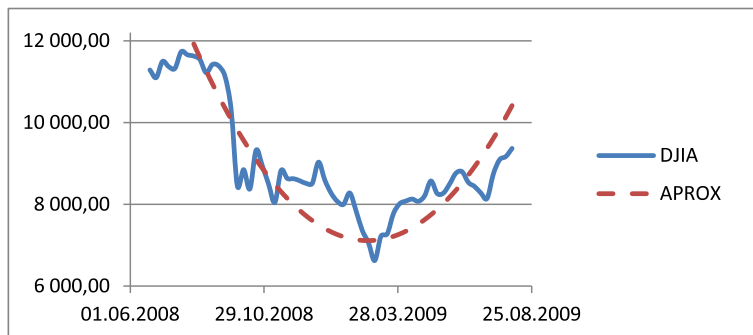
Kvadratický trend tohto procesu sa dá odhadnúť už z grafického znázornenia na obrázku 3.10. Parabolické správanie procesu vyjadríme pomocou troch koeficientov ako

$$\begin{aligned} \tilde{\mathbf{f}} &= c_0 \cdot \mathbf{b}_0 + c_1 \cdot \mathbf{b}_1 + c_2 \cdot \mathbf{b}_2 \\ &= c_0 \cdot (1, 1, \dots, 1) + c_1 \cdot (0, 1, 2, 3, \dots, 58) + c_2 \cdot (0, 1, 4, 9, \dots, 58^2) \end{aligned}$$

Koeficienty c_0, c_1, c_2 vypočítame ako riešenie sústavy rovníc

$$\begin{aligned} c_0 \cdot 59 + c_1 \cdot \sum_{k=0}^{58} k + c_2 \cdot \sum_{k=0}^{58} k^2 &= \sum_{k=0}^{58} f_k \\ c_0 \cdot \sum_{k=0}^{58} k + c_1 \cdot \sum_{k=0}^{58} k^2 + c_2 \cdot \sum_{k=0}^{58} k^3 &= \sum_{k=0}^{58} f_k \cdot k \\ c_0 \cdot \sum_{k=0}^{58} k^2 + c_1 \cdot \sum_{k=0}^{58} k^3 + c_2 \cdot \sum_{k=0}^{58} k^4 &= \sum_{k=0}^{58} f_k \cdot k^2 \\ c_0 \cdot 59 + c_1 \cdot 1722 + c_2 \cdot 66729 &= 533663 \\ c_0 \cdot 1722 + c_1 \cdot 66729 + c_2 \cdot 2927521 &= 14524784 \\ c_0 \cdot 66729 + c_1 \cdot 2927521 + c_2 \cdot 136994637 &= 560982049 \end{aligned}$$

Riešením tejto sústavy sú koeficienty $c_0 = 14657.7, c_1 = -431.942, c_2 = 6.188$. Pribeh procesu a jeho aproximácia kvadratickým procesom s tromi parametrami sú zobrazené na obrázku 3.10.

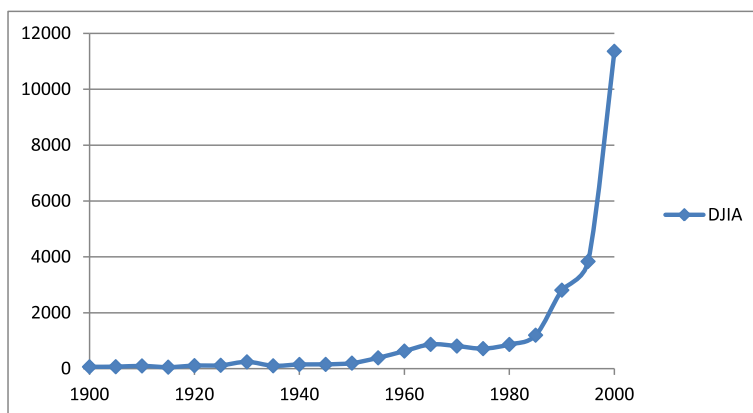


Obr. 3.10: Vývoj akciového indexu Dow Jones Industrial Average v období od 1. 7. 2008 do 1. 8. 2009 a jeho aproximácia parabolou s tromi parametrami

Úloha 3.8.1 *Pokiaľ by sme sledovali vývoj hodnôt indexu DJIA od začiatku minulého storočia, na obrázku 3.11 vidno, že trend by nebol ani lineárny ani kvadratický. V 5-ročných intervaloch by sme namerali tieto hodnoty, začínajúce v roku 1900 a končiace v roku 2000:*

$f = (66.61, 70.39, 98.34, 54.63, 108.76, 121.25, 244.2, 102.76,$
 $151.43, 153.58, 198.89, 388.2, 630.24, 869.78, 809.2, 717.85, 867.15,$
 $1198.87, 2810.15, 3838.24, 11357.51)$

Aproximujte tento proces procesom s exponenciálnou krivkou $y = c_0 e^{c_1 \cdot t}$ s dvoma parametrami c_0 a c_1 .



Obr. 3.11: Vývoj akciového indexu Dow Jones Industrial Average v období od roku 1900 do 2000 v 5-ročných intervaloch

3.9 Lineárna regresia ako priemet do podpriestoru

Nech $\mathbf{f}_{\mathcal{E}} = (f_0, f_1, \dots, f_{N-1})$ je proces vyjadrený v jednotkovej báze N -rozmerného vektorového priestoru. Tento proces aproximujeme procesmi menejrozmerného vektorového podpriestoru \mathbb{U} podobne, ako sme to robili v predchádzajúcich úlohách. Nasledujú príklady priemetu procesu \mathbf{f} v podpriestoroch generovaných bázami:

- Jednorozmerný podpriestor generovaný bázou

$$\mathcal{B}_1 = (\mathbf{b}) = ((1, 1, \dots, 1)).$$

Priemety do tohto podpriestoru majú tvar:

$$\tilde{\mathbf{f}}_{\mathcal{E}} = (c, c, \dots, c)$$

$$\tilde{\mathbf{f}}_{\mathcal{B}_1} = (c)$$

- Dvojrozmerný podpriestor generovaný bázou

$$\mathcal{B}_2 = (\mathbf{b}_0, \mathbf{b}_1) = ((1, 1, \dots, 1), (t_0, t_1, \dots, t_{N-1})).$$

Priemety do tohto podpriestoru majú tvar:

$$\tilde{\mathbf{f}}_{\mathcal{E}} = (c_0 + c_1 \cdot t_0, c_0 + c_1 \cdot t_1, \dots, c_0 + c_1 \cdot t_{N-1})$$

$$\tilde{\mathbf{f}}_{\mathcal{B}_2} = (c_0, c_1)$$

- M -rozmerný podpriestor generovaný bázou

$$\mathcal{B}_3 = ((1, 1, \dots, 1), \dots, (t_0^{M-1}, t_1^{M-1}, \dots, t_{N-1}^{M-1})).$$

Priemety do tohto podpriestoru majú tvar:

$$\begin{aligned} \tilde{\mathbf{f}}_{\mathcal{E}} = & (c_0 + c_1 t_0 + c_2 \cdot t_0^2 + \dots + c_{M-1} \cdot t_0^{M-1}, \\ & c_0 + c_1 t_1 + c_2 \cdot t_1^2 + \dots + c_{M-1} \cdot t_1^{M-1}, \\ & \dots \\ & c_0 + c_1 t_{N-1} + c_2 \cdot t_{N-1}^2 + \dots + c_{M-1} \cdot t_{N-1}^{M-1}) \end{aligned}$$

$$\tilde{\mathbf{f}}_{\mathcal{B}_3} = (c_0, c_1, \dots, c_{M-1})$$

- M -rozmerný podpriestor generovaný bázou

$$\mathcal{B}_4 = ((\varphi_0(t_0), \dots, \varphi_0(t_{N-1})), \dots, (\varphi_{M-1}(t_0), \dots, \varphi_{M-1}(t_{N-1}))).$$

Priemety do tohto podpriestoru majú tvar:

$$\begin{aligned} \tilde{\mathbf{f}}_{\mathcal{E}} = & (c_0 \cdot \varphi_0(t_0) + c_1 \cdot \varphi_1(t_0) + c_2 \cdot \varphi_2(t_0) + \dots + c_{M-1} \cdot \varphi_{M-1}(t_0), \\ & c_0 \cdot \varphi_0(t_1) + c_1 \cdot \varphi_1(t_1) + c_2 \cdot \varphi_2(t_1) + \dots + c_{M-1} \cdot \varphi_{M-1}(t_1), \\ & \dots \\ & c_0 \cdot \varphi_0(t_{N-1}) + c_1 \cdot \varphi_1(t_{N-1}) + \dots + c_{M-1} \cdot \varphi_{M-1}(t_{N-1})) \\ = & \left(\sum_{n=0}^{M-1} c_n \cdot \varphi_n(t_0), \dots, \sum_{n=0}^{M-1} c_n \cdot \varphi_n(t_k), \dots, \sum_{n=0}^{M-1} c_n \cdot \varphi_n(t_{N-1}) \right) \\ \tilde{\mathbf{f}}_{\mathcal{B}_4} = & (c_0, c_1, \dots, c_{M-1}) \end{aligned}$$

Úloha 3.9.1 Nájdite v literatúre alebo na internete návod, ako riešiť úlohu lineárnej regresie pomocou diferenciálneho počtu funkcie viac premenných. Vyriešte pomocou tohto návodu úlohu z príkladu 3.7.1.

Úloha 3.9.2 Porovnajte riešenie rovnakej úlohy v aspoň troch rôznych softvéroch. Zistite, ktorý urobí najmenšiu a ktorý najväčšiu chybu odhadu.

3.10 Všeobecný zápis lineárnej regresie

Všeobecne sa snažíme proces \mathbf{f} nahradiť procesom $\tilde{\mathbf{f}}$ vyjadreným v báze $\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_{M-1}$

$$\tilde{\mathbf{f}} = \sum_{n=0}^{M-1} c_n \mathbf{b}_n$$

pričom

$$\mathbf{b}_0 = (\varphi_0(t_0), \varphi_0(t_1), \dots, \varphi_0(t_k), \dots, \varphi_0(t_{N-1}))$$

$$\mathbf{b}_1 = (\varphi_1(t_0), \varphi_1(t_1), \dots, \varphi_1(t_k), \dots, \varphi_1(t_{N-1}))$$

$$\vdots$$

$$\mathbf{b}_{M-1} = (\varphi_{M-1}(t_0), \varphi_{M-1}(t_1), \dots, \varphi_{M-1}(t_k), \dots, \varphi_{M-1}(t_{N-1}))$$

Teda proces $\tilde{\mathbf{f}}$ je lineárnou kombináciou procesov $\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_{M-1}$ a proces $\mathbf{f} - \tilde{\mathbf{f}}$ je kolmý na všetky báziké procesy. Kolmost' môžeme vyjadriť pomocou skalárneho súčinu:

$$(\mathbf{f} - \tilde{\mathbf{f}}) \perp \mathbf{b}_m \quad \text{pre } m = 0, 1, \dots, M-1$$

$$\langle \mathbf{f} - \tilde{\mathbf{f}}, \mathbf{b}_m \rangle = 0$$

$$\langle \mathbf{f}, \mathbf{b}_m \rangle = \langle \tilde{\mathbf{f}}, \mathbf{b}_m \rangle$$

Dosadíme:

$$\langle \mathbf{f}, \mathbf{b}_m \rangle = \left\langle \sum_{n=0}^{M-1} c_n \cdot \mathbf{b}_n, \mathbf{b}_m \right\rangle$$

Upravíme:

$$\langle \mathbf{f}, \mathbf{b}_m \rangle = \sum_{n=0}^{M-1} c_n \cdot \langle \mathbf{b}_n, \mathbf{b}_m \rangle \quad \text{pre } m = 0, 1, \dots, M-1$$

$$\sum_{k=0}^{N-1} f_k \varphi_m(t_k) = \sum_{n=0}^{M-1} c_n \cdot \sum_{k=0}^{N-1} \varphi_n(t_k) \varphi_m(t_k) \quad \text{pre } m = 0, 1, \dots, M-1$$

Príklad 3.10.1 V časoch 0, 1, 3, 4 nadobudol proces \mathbf{f} hodnoty 7, 3, 5, 7. Zistite, ktorá z nasledujúcich funkcií (modelov) najlepšie vyjadruje priebeh procesu a aká bude približná hodnota procesu v časoch 2 a 5.

a) $\varphi(t) = c$

$$b) \varphi(t) = c \cdot t$$

$$c) \varphi(t) = c \cdot t^2$$

$$d) \varphi(t) = c_0 + c_1 \cdot t$$

$$e) \varphi(t) = c_0 \cdot t + c_1 \cdot t^2$$

$$f) \varphi(t) = c_0 + c_1 \cdot e^t$$

Riešenie:

a) Pre model $\varphi(t) = c$, teda pre konštatný priebeh zvolíme bázu

$$\mathcal{B} = ((1, 1, 1, 1))$$

Proces \mathbf{f} odhadneme procesom $\tilde{\mathbf{f}} = (c)\mathcal{B} = (c, c, c, c)\mathcal{E}$.

Predpokladáme teda, že odhad má tvar $\tilde{\mathbf{f}} = c \cdot \mathbf{b}$ a hodnotu c vypočítame úpravou vzťahu

$$\langle \mathbf{f} - \tilde{\mathbf{f}}, \mathbf{b} \rangle = 0$$

Využitím vlastností skalárneho súčinu dostaneme

$$\langle \mathbf{f}, \mathbf{b} \rangle = \langle \tilde{\mathbf{f}}, \mathbf{b} \rangle$$

$$\langle \mathbf{f}, \mathbf{b} \rangle = c \cdot \langle \mathbf{b}, \mathbf{b} \rangle$$

Preto platí

$$c = \frac{\langle \mathbf{f}, \mathbf{b} \rangle}{\langle \mathbf{b}, \mathbf{b} \rangle} = \frac{22}{4} = 5.5$$

Proces $\mathbf{f} = (7, 3, 5, 7)$ odhadneme (priblížime, aproximujeme) procesom $\tilde{\mathbf{f}} = (5.5, 5.5, 5.5, 5.5)$. Chyba tohto odhadu (chybový vektor) je

$$\mathbf{e} = \mathbf{f} - \tilde{\mathbf{f}} = (1.5, -2.5, -0.5, 1.5)$$

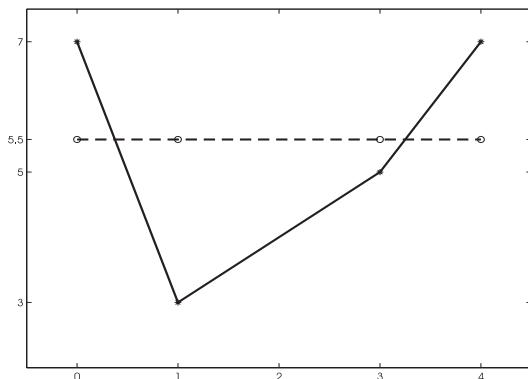
Veľkosť tejto chyby je

$$\|\mathbf{e}\| = \|\mathbf{f} - \tilde{\mathbf{f}}\| = \sqrt{(1.5)^2 + (-2.5)^2 + (-0.5)^2 + (1.5)^2} = 3.3166$$

V čase 2 bude odhadnutá hodnota procesu rovná 5.5 (interpolácia).

V čase 5 bude odhadnutá hodnota procesu rovná 5.5 (extrapolácia).

$$\varphi(t_k) = c, \quad \varphi(2) = 5.5, \quad \varphi(5) = 5.5$$

Obr. 3.12. Aproximácia procesu priamkou $\varphi(t_k) = 5.5$

- b) Pre model $\varphi(t) = c \cdot t$, teda pre lineárny nárast, zvolíme jednorozmernú bázu. Bázový vektor má zložky $c \cdot t_0, c \cdot t_1, c \cdot t_2, c \cdot t_3$, ktoré zodpovedajú hodnotám $\varphi(t_k) = c \cdot t_k$ v časoch t_0, t_1, t_2, t_3 . Pre jednoduchosť zvolíme $c = 1$ a dostaneme bázu

$$\mathcal{B}_1 = ((0, 1, 3, 4))$$

Proces \mathbf{f} odhadneme procesom $\tilde{\mathbf{f}} = (c)\mathcal{B}_1$, $\tilde{\mathbf{f}} = (0, c, 3 \cdot c, 4 \cdot c)\mathcal{E}$.

Hodnotu c vypočítame ako v predchádzajúcej časti zo vzťahu

$$\langle \mathbf{f}, \mathbf{b} \rangle = c \cdot \langle \mathbf{b}, \mathbf{b} \rangle$$

Preto platí

$$c = \frac{\langle \mathbf{f}, \mathbf{b} \rangle}{\langle \mathbf{b}, \mathbf{b} \rangle} = \frac{46}{26} = 1.77$$

Proces $\mathbf{f} = (7, 3, 5, 7)$ odhadneme (priblížime, aproximujeme) procesom $\tilde{\mathbf{f}} = (0, 1.77, 5.31, 7.08)$. Chyba tohto odhadu (chybový vektor) je

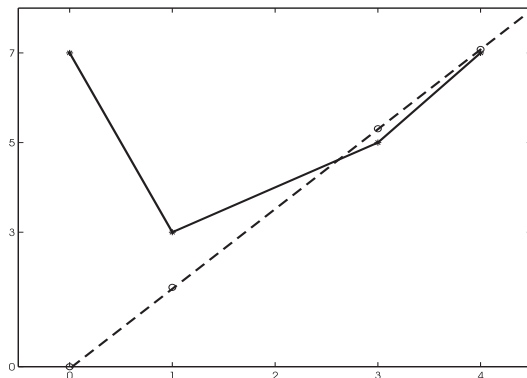
$$\mathbf{e}_1 = \mathbf{f} - \tilde{\mathbf{f}} = (7, 1.23, -0.31, -0.08)$$

Veľkosť tejto chyby je

$$\|\mathbf{e}_1\| = \|\mathbf{f} - \tilde{\mathbf{f}}\| = \sqrt{(7)^2 + (1.23)^2 + (-0.31)^2 + (-0.08)^2} = 7.11$$

Veľkosť chyby ukazuje, že táto aproximácia je horšia ako v predchádzajúcom modeli. Dôvodom, že chyba je veľká je, že v tomto modeli musí regresná priamka $c \cdot t_k$ prechádzať bodom $(0, 0)$, ktorý predstavuje v čase 0 hodnotu procesu 0. Tento bod je ďaleko od skutočnej hodnoty $f(0) = 7$ procesu v bode 0. Príslušné regresné priamky sú na obrázkoch 3.12 a 3.13. V časoch 2 a 5 budú odhadnuté hodnoty procesu rovné 3.54 a 8.85.

$$\varphi(t_k) = c \cdot t_k, \quad \varphi(2) = 1.77 \cdot 2 = 3.54, \quad \varphi(5) = 1.77 \cdot 5 = 8.85$$

Obr. 3.13. Aproximácia procesu priamkou $\varphi(t_k) = 1.77t_k$

- c) Pre model $\varphi(t_k) = c \cdot t_k^2$, teda pre kvadratický priebeh zvolíme bázu, ktorej prvky sú štvorzložkové vektory. Jednotlivé zložky vektorov zodpovedajú modelu v časoch t_0, t_1, t_2, t_3 . Teda zložky vektorov majú postupne hodnoty $c \cdot t_0^2, c \cdot t_1^2, c \cdot t_2^2, c \cdot t_3^2$. Pre jednoduchosť zvolíme $c = 1$ a dostaneme bázu

$$\mathcal{B}_2 = ((0, 1, 9, 16))$$

Proces \mathbf{f} odhadneme procesom $\tilde{\mathbf{f}} = (c)\mathcal{B}_2$, $\tilde{\mathbf{f}} = (0, c, 9 \cdot c, 16 \cdot c)\mathcal{E}$. Hodnotu c vypočítame ako

$$c = \frac{\langle \mathbf{f}, \mathbf{b} \rangle}{\langle \mathbf{b}, \mathbf{b} \rangle} = 0.47$$

Proces $\mathbf{f} = (7, 3, 5, 7)$ odhadneme procesom $\tilde{\mathbf{f}} = (0, 0.47, 4.26, 7.57)$. Chyba tohto odhadu je

$$\mathbf{e}_2 = \mathbf{f} - \tilde{\mathbf{f}} = (7, 2.53, 0.74, -0.57)$$

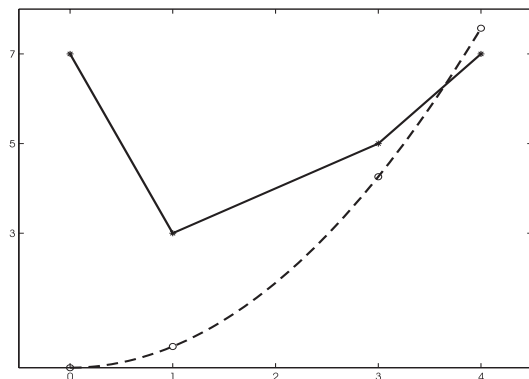
Veľkosť tejto chyby je

$$\|\mathbf{e}_2\| = \|\mathbf{f} - \tilde{\mathbf{f}}\| = \sqrt{(7)^2 + (2.53)^2 + (0.74)^2 + (-0.57)^2} = 7.5007$$

Veľkosť chyby ukazuje, že táto aproximácia je lepšia ako v predchádzajúcom modeli. Chyba je ale stále veľká, pretože v tomto modeli musí regresná krivka $c \cdot t_k^2$ prechádzať bodom $(0, 0)$, ktorý predstavuje v čase 0 hodnotu procesu 0. Tento bod je ďaleko od skutočnej hodnoty (=7) procesu v bode 0. Regresná krivka pre tento model (parabola) je na obrázku 3.14.

V časoch 2 a 5 budú odhadnuté hodnoty procesu rovné 1.89 a 11.83.

$$\varphi(t_k) = c \cdot t_k^2, \quad \varphi(2) = 0.47 \cdot 2^2 = 1.89, \quad \varphi(5) = 0.47 \cdot 5^2 = 11.83$$



Obr. 3.14. Aproximácia procesu regresnou krivkou $\varphi(t_k) = 0.4734 \cdot t_k^2$

- d) Pre model $\varphi(t_k) = c_0 + c_1 \cdot t_k$, teda pre lineárny nárast s konštantnou zložkou zvolíme bázu, ktorej prvky sú štvorzložkové vektory. Jednotlivé zložky vektorov zodpovedajú modelu v časoch t_0, t_1, t_2, t_3 . Teda zložky vektorov majú postupne hodnoty $c_0 + c_1 \cdot t_0, c_0 + c_1 \cdot t_1, c_0 + c_1 \cdot t_2, c_0 + c_1 \cdot t_3$. Pre jednoduchosť zvolíme najprv $c_0 = 1, c_1 = 0$, potom $c_0 = 0, c_1 = 1$ a dostaneme bázu

$$\mathcal{B}_3 = ((1, 1, 1, 1), (0, 1, 3, 4))$$

Proces \mathbf{f} odhadneme procesom

$$\tilde{\mathbf{f}} = (c_0, c_1)\mathcal{B}_3$$

$$\tilde{\mathbf{f}} = (c_0, c_0, c_0, c_0)\mathcal{E} + (0, c_1, 3 \cdot c_1, 4 \cdot c_1)\mathcal{E}$$

Teda $\tilde{\mathbf{f}} = c_0 \cdot \mathbf{b}_0 + c_1 \cdot \mathbf{b}_1$. Hodnoty c_0 a c_1 vypočítame úpravou vzťahov

$$\langle \mathbf{f} - \tilde{\mathbf{f}}, \mathbf{b}_0 \rangle = 0$$

$$\langle \mathbf{f} - \tilde{\mathbf{f}}, \mathbf{b}_1 \rangle = 0$$

$$\langle \mathbf{f}, \mathbf{b}_0 \rangle = (c_0 \cdot \mathbf{b}_0 + c_1 \cdot \mathbf{b}_1, \mathbf{b}_0)$$

$$\langle \mathbf{f}, \mathbf{b}_1 \rangle = (c_0 \cdot \mathbf{b}_0 + c_1 \cdot \mathbf{b}_1, \mathbf{b}_1)$$

Preto sú c_0 a c_1 riešeniami sústavy

$$\langle \mathbf{f}, \mathbf{b}_0 \rangle = c_0 \cdot \langle \mathbf{b}_0, \mathbf{b}_0 \rangle + c_1 \cdot \langle \mathbf{b}_1, \mathbf{b}_0 \rangle$$

$$\langle \mathbf{f}, \mathbf{b}_1 \rangle = c_0 \cdot \langle \mathbf{b}_0, \mathbf{b}_1 \rangle + c_1 \cdot \langle \mathbf{b}_1, \mathbf{b}_1 \rangle$$

$$22 = 4c_0 + 8c_1, \quad 46 = 8c_0 + 26c_1$$

$$c_0 = 5.1, \quad c_1 = 0.2$$

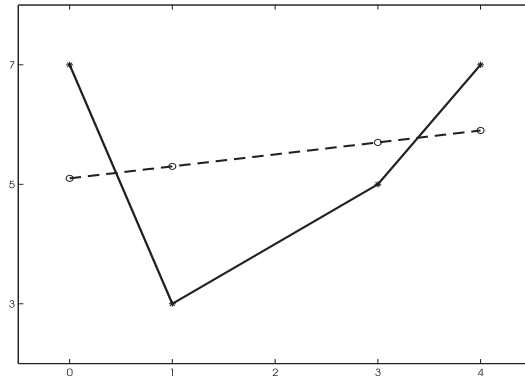
Proces $\mathbf{f} = (7, 3, 5, 7)$ odhadneme (približime, aproximujeme) procesom $\tilde{\mathbf{f}} = (5.1, 5.3, 5.7, 5.9)$. Chyba tohto odhadu (chybový vektor) je

$$\mathbf{e}_3 = \mathbf{f} - \tilde{\mathbf{f}} = (1.9, -2.3, -0.7, 1.1)$$

Veľkosť tejto chyby je

$$\|\mathbf{e}_3\| = \|\mathbf{f} - \tilde{\mathbf{f}}\| = 3.2558$$

Toto je zatiaľ najlepšia aproximácia. Príslušná regresná priamka je na obrázku 3.15.



Obr. 3.15. Aproximácia procesu priamkou $\varphi(t_k) = 5.1 + 0.2 \cdot t_k$

V časoch 2 a 5 budú odhadnuté hodnoty procesu rovné 5.5 a 6.1.

$$\varphi(t_k) = c_0 + c_1 \cdot t_k, \quad \varphi(2) = 5.5, \quad \varphi(5) = 6.1$$

- e) Predtým ako budeme rozoberať model $\varphi(t_k) = c_0 \cdot t_k + c_1 \cdot t_k^2$, zamyslime sa, či vôbec má šancu byť „dobrý“ model, teda či jeho chyba bude nízka. Ak si uvedomíme, že aj tu bude musieť regresná krivka prechádzať bodom $(0, 0)$, tak hneď vidno, že veľkosť chyby bude aspoň 7, teda model bude opäť dosť nepresný.

Pre model $\varphi(t_k) = c_0 \cdot t_k + c_1 \cdot t_k^2$ zvolíme bázu, ktorej prvky sú štvorzložkové vektory. Jednotlivé zložky vektorov zodpovedajú modelu v časoch t_0, t_1, t_2, t_3 . Teda zložky vektorov majú postupne hodnoty $c_0 \cdot t_0 + c_1 \cdot t_0^2, c_0 \cdot t_1 + c_1 \cdot t_1^2, c_0 \cdot t_2 + c_1 \cdot t_2^2, c_0 \cdot t_3 + c_1 \cdot t_3^2$. Pre jednoduchosť zvolíme najprv $c_0 = 1, c_1 = 0$, potom $c_0 = 0, c_1 = 1$ a dostaneme bázu

$$\mathcal{B}_4 = ((0, 1, 3, 4), (0, 1, 9, 16))$$

Proces \mathbf{f} odhadneme procesom

$$\tilde{\mathbf{f}} = (c_0, c_1)\mathcal{B}_4$$

$$\tilde{\mathbf{f}}_{\mathcal{E}} = (0, c_0, 3 \cdot c_0, 4 \cdot c_0) + (0, c_1, 9 \cdot c_1, 16 \cdot c_1)$$

Teda $\tilde{\mathbf{f}} = c_0 \cdot \mathbf{b}_0 + c_1 \cdot \mathbf{b}_1$. Hodnoty c_0 a c_1 vypočítame riešením sústavy

$$46 = 26c_0 + 92c_1, \quad 160 = 92c_0 + 338c_1$$

$$c_0 = 2.56, \quad c_1 = -0.22$$

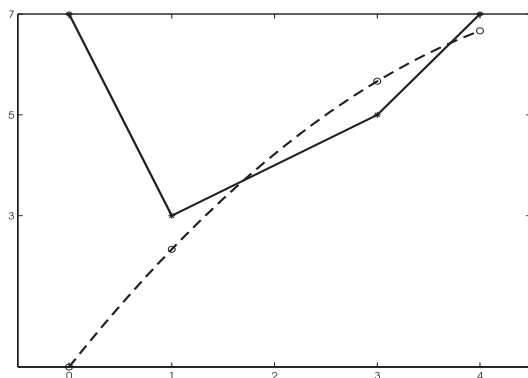
Proces $\mathbf{f} = (7, 3, 5, 7)$ aproximujeme procesom $\tilde{\mathbf{f}} = (0, 2.33, 5.67, 6.67)$.

Chyba tohto odhadu (chybový vektor) je

$$\mathbf{e}_4 = \mathbf{f} - \tilde{\mathbf{f}} = (7, 0.67, -0.67, 0.33)$$

Veľkosť tejto chyby je $\|\mathbf{e}_4\| = \|\mathbf{f} - \tilde{\mathbf{f}}\| = 7.0711$.

Príslušná regresná krivka je na obrázku 3.16.



Obr. 3.16. Aproximácia procesu krivkou $\varphi(t_k) = 2.56 \cdot t_k - 0.22 \cdot t_k^2$

V časoch 2 a 5 budú odhadnuté hodnoty procesu rovné 4.22 a 7.22.

$$\varphi(t_k) = c_0 \cdot t_k + c_1 \cdot t_k^2, \quad \varphi(2) = 2.56 \cdot 2 - 0.22 \cdot 2^2 = 4.22,$$

$$\varphi(5) = 2.56 \cdot 5 - 0.22 \cdot 5^2 = 7.22$$

- f) Pre model $\varphi(t_k) = c_0 + c_1 e^{t_k}$ zvolíme bázu, ktorej prvky sú štvorzložkové vektory. Jednotlivé zložky vektorov zodpovedajú modelu v časoch t_0, t_1, t_2, t_3 . Teda zložky vektorov majú postupne hodnoty $c_0 + c_1 \cdot e^{t_0}, c_0 + c_1 \cdot e^{t_1}, c_0 + c_1 \cdot e^{t_2}, c_0 + c_1 \cdot e^{t_3}$. Pre jednoduchosť zvolíme najprv $c_0 = 1, \quad c_1 = 0$, potom $c_0 = 0, \quad c_1 = 1$ a dostaneme bázu

$$\mathcal{B}_5 = ((1, 1, 1, 1), (1, 2.72, 20.09, 54.6))$$

Proces \mathbf{f} odhadneme procesom

$$\tilde{\mathbf{f}} = (c_0, c_1) \mathcal{B}_5$$

$$\tilde{\mathbf{f}} = (c_0, c_0, c_0, c_0) \mathcal{E} + (e^0 \cdot c_1, e^1 \cdot c_1, e^3 \cdot c_1, e^4 \cdot c_1) \mathcal{E}$$

Teda $\tilde{\mathbf{f}} = c_0 \cdot \mathbf{b}_0 + c_1 \cdot \mathbf{b}_1$. Hodnoty c_0 a c_1 vypočítame riešením sústavy

$$\begin{aligned} 22 &= 4c_0 + 78.4c_1, & 497.77 &= 78.4c_0 + 3392.8c_1 \\ c_0 &= 4.8, & c_1 &= 0.04 \end{aligned}$$

Proces $\mathbf{f} = (7, 3, 5, 7)$ aproximujeme procesom

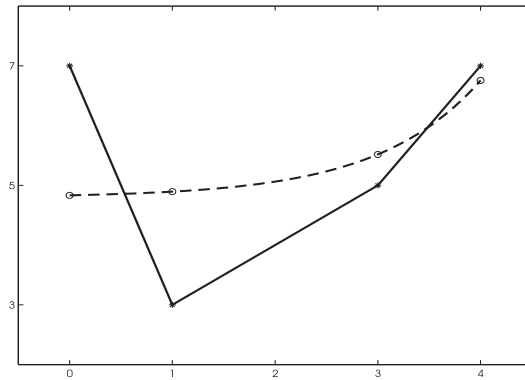
$$\tilde{\mathbf{f}} = (4.83, 4.89, 5.51, 6.75).$$

Chyba tohto odhadu (chybový vektor) je

$$\mathbf{e}_5 = \mathbf{f} - \tilde{\mathbf{f}} = (2.17, -1.9, -0.51, 0.24)$$

Veľkosť tejto chyby je $\|\mathbf{e}_5\| = \|\mathbf{f} - \tilde{\mathbf{f}}\| = 2.94$.

Príslušná regresná krivka je na obrázku 3.17.



Obr. 3.17. Aproximácia procesu krivkou $\varphi(t_k) = 4.8 + 0.04 e^{t_k}$

V časoch 2 a 5 budú odhadnuté hodnoty procesu rovné 5.06 a 10.11.

$$\varphi(2) = 4.8 + 0.04 \cdot e^2 = 5.06, \quad \varphi(5) = 4.8 + 0.04 \cdot e^5 = 10.11$$

□

Najlepší odhad spomedzi uvedených modelov dáva funkcia $\varphi(t_k) = c_0 + c_1 e^{t_k}$. Veľkosť chyby v tomto prípade je 2.94.

Úloha 3.10.1 Navrhnite funkciu s dvomi parametrami c_0 a c_1 , ktorá odhadne proces z príkladu 3.10.1 lepšie, než modely navrhnuté v príklade.

3.11 Koefficienty v neortogónálnej báze

Koefficienty priemetu procesu do bázy sú rôzne v závislosti od bázy. Pre jeden vektor dostávame rôzne jeho vyjadrenia v rôznych bázach.

$$\tilde{\mathbf{f}} = c_0 \mathbf{b}_0 + c_1 \mathbf{b}_1 + \cdots + c_n \mathbf{b}_n = c'_0 \mathbf{b}'_0 + c'_1 \mathbf{b}'_1 + \cdots + c'_n \mathbf{b}'_n$$

Ak bazové vektory predstavujú časové vyjadrenie funkcie, ktorou je proces modelovaný, je možné vypočítané koeficienty použiť na vyjadrenie funkcie, ktorou proces aproximujeme:

$$\varphi(t_k) = c_0\varphi_0(t_k) + c_1\varphi_1(t_k) + \cdots + c_n\varphi_n(t_k) \quad (3.11.1)$$

Výpočet koeficientov c_k vyžaduje riešiť sústavu rovníc. V prípade, že bázu najprv ortogonalizujeme, stačí na výpočet každého z koeficientov jeden vzorec.

$$c'_k = \frac{\langle \mathbf{f}, \mathbf{b}'_k \rangle}{\langle \mathbf{b}'_k, \mathbf{b}'_k \rangle}$$

Preto sa niekedy neortogonálna báza (vyjadrujúca časový priebeh) najskôr ortogonalizuje. Výpočet koeficientov ortogonálnej bázy bude jednoduchší, ale koeficienty c'_k sa nedajú použiť ako koeficienty funkcie $\varphi(t_k)$. V nasledujúcom príklade ukážeme, ako sa zo známych koeficientov ortogonálnej bázy dajú vypočítať koeficienty pôvodnej neortogonálnej bázy.

Príklad 3.11.1 *Aproximujte proces z príkladu 3.6.1, resp. príkladu 3.7.1, procesom s lineárnym nárastom s dvomi parametrami. Bázu najprv ortogonalizujte, potom vypočítajte koeficienty v ortogonálnej báze a nakoniec z týchto koeficientov odvodte koeficienty funkcie*

$$\varphi(t_k) = c_0 + c_1 \cdot t_k$$

Riešenie:

Lineárny nárast procesu vyjadríme pomocou dvoch koeficientov ako

$$\tilde{\mathbf{f}} = c_0 \cdot \mathbf{b}_0 + c_1 \cdot \mathbf{b}_1 = c_0 \cdot (1, 1, \dots, 1) + c_1 \cdot (0, 1, 2, \dots, 99)$$

Zodpovedajúca neortogonálna báza je

$$\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1) = ((1, 1, \dots, 1), (0, 1, 2, \dots, 99))$$

Ortogonálnu bázu \mathcal{B}' dostaneme Gram-Schmidtovou ortogonalizáciou bázy \mathcal{B} .

$$\mathbf{b}'_0 = \mathbf{b}_0$$

$$\mathbf{b}'_1 = \mathbf{b}_1 - \tilde{\mathbf{b}}_1 = \mathbf{b}_1 - k \cdot \mathbf{b}_0 = \mathbf{b}_1 - 49.5 \cdot \mathbf{b}_0$$

Koeficienty c'_0, c'_1 v báze \mathcal{B}' vypočítame zo vzorcov

$$c'_0 = \frac{\langle \mathbf{f}, \mathbf{b}'_0 \rangle}{\langle \mathbf{b}'_0, \mathbf{b}'_0 \rangle} = 2647.7$$

$$c'_1 = \frac{\langle \mathbf{f}, \mathbf{b}'_1 \rangle}{\langle \mathbf{b}'_1, \mathbf{b}'_1 \rangle} = 36.06$$

Koeficienty c_0 a c_1 dostaneme zo vzťahu

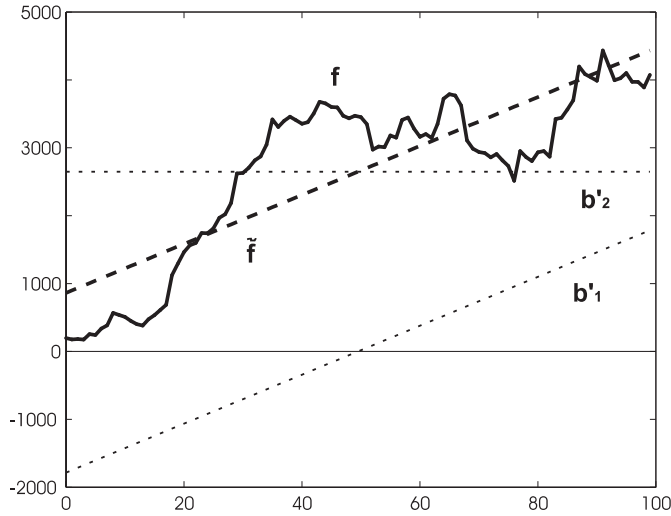
$$\tilde{\mathbf{f}} = c_0 \cdot \mathbf{b}_0 + c_1 \cdot \mathbf{b}_1 = c'_0 \cdot \mathbf{b}'_0 + c'_1 \cdot \mathbf{b}'_1 = 2647.7 \cdot \mathbf{b}_0 + 36.06 \cdot (\mathbf{b}_1 - 49.5 \cdot \mathbf{b}_0) =$$

$$= (2647.7 - 1.7851) \cdot \mathbf{b}_0 + 36.06 \cdot \mathbf{b}_1 = 862.61 \cdot \mathbf{b}_0 + 36.06 \cdot \mathbf{b}_1$$

Odtiaľ dostávame $c_0 = 862.61$ a $c_1 = 36.06$, a teda vyjadrenie procesu ako funkcie času bude

$$\varphi(t_k) = 862.61 + 36.06 \cdot t_k$$

Výslednú aproximáciu vidíme na obrázku 3.18. Rovnaký výsledok sme dostali aj v príklade 3.7.1, keď sme koeficienty c_0 a c_1 hľadali riešením sústavy rovníc.



Obr. 3.18: Aproximácia dát priamkou $\varphi(t_k) = 862.61 + 36.06 \cdot t_k$, vektor \mathbf{b}'_0 reprezentuje konštantný trend a vektor \mathbf{b}'_1 reprezentuje lineárny nárast

Príklad 3.11.2 *Nech je proces \mathbf{f} aproximovaný pomocou časovej funkcie*

$$\varphi(t_k) = c_0\varphi_0(t_k) + c_1\varphi_1(t_k) + c_2\varphi_2(t_k)$$

Jednotlivé funkcie $\varphi_0(t_k), \varphi_1(t_k), \varphi_2(t_k)$ sú vyjadrené procesmi pôvodnej bázy

$$\mathcal{B} = (\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2)$$

Vyjadrenie v pôvodnej báze je

$$\tilde{\mathbf{f}} = c_0\mathbf{b}_0 + c_1\mathbf{b}_1 + c_2\mathbf{b}_2$$

Z bázy \mathcal{B} je Gram-Schmidtovou ortogonalizáciou vytvorená ortogonálna báza:

$$\mathcal{B}' = (\mathbf{b}'_0, \mathbf{b}'_1, \mathbf{b}'_2)$$

Vyjadrite zo známych koeficientov c'_0, c'_1, c'_2 procesu v ortogonálnej báze, koeficienty v pôvodnej, neortogonálnej báze.

Riešenie:

Priemet procesu \mathbf{f} do podpriestoru s ortogonálnou bázou vyjadruje vzťah:

$$\tilde{\mathbf{f}} = c'_0 \mathbf{b}'_0 + c'_1 \mathbf{b}'_1 + c'_2 \mathbf{b}'_2$$

Vyjadríme vektory ortogonálnej bázy pomocou vektorov pôvodnej bázy.

Potom pôvodné bázičné vektory nahradíme jednotlivými funkciami, čím dostaneme funkciu, ktorá aproximuje proces.

Postup ako vyjadriť vektory ortogonálnej bázy pomocou vektorov pôvodnej bázy je možné vyvodiť z postupu ortogonalizácie.

Je zrejmé, že $\mathbf{b}_0 = \mathbf{b}'_0$.

V postupe ortogonalizácie sme použili

$$\mathbf{b}'_1 = \mathbf{b}_1 - k \cdot \mathbf{b}'_0 = \mathbf{b}_1 - k \cdot \mathbf{b}_0, \quad \text{kde} \quad k = \frac{\langle \mathbf{b}_1, \mathbf{b}_0 \rangle}{\langle \mathbf{b}_0, \mathbf{b}_0 \rangle}$$

$$\mathbf{b}'_2 = \mathbf{b}_2 - k_0 \cdot \mathbf{b}_0 - k_1 \mathbf{b}'_1, \quad \text{kde} \quad k_0 = \frac{\langle \mathbf{b}_2, \mathbf{b}_0 \rangle}{\langle \mathbf{b}_0, \mathbf{b}_0 \rangle} \text{ a } k_1 = \frac{\langle \mathbf{b}_2, \mathbf{b}'_1 \rangle}{\langle \mathbf{b}'_1, \mathbf{b}'_1 \rangle}$$

Vo vzťahu pre \mathbf{b}'_2 je síce použité \mathbf{b}'_1 , ktoré je ale vyjadrené vyššie.

Čiže

$$\mathbf{b}'_2 = \mathbf{b}_2 - k_0 \cdot \mathbf{b}_0 - k_1 (\mathbf{b}_1 - k \mathbf{b}_0) = \mathbf{b}_2 - k_1 \mathbf{b}_1 - (k_0 - k_1 \cdot k) \mathbf{b}_0$$

Napríklad funkcia

$$y = a \cdot t_k + b \cdot t_k^2 + c \cdot t_k^3$$

by pomocou koeficientov c'_0, c'_1, c'_2 vypočítaných v ortogonálnej báze vyzerala nasledujúco

$$f(t_k) = c'_0 \cdot t_k + c'_1 \cdot (t_k^2 - kt_k) + c'_2 \cdot (t_k^3 - k_1 \cdot t_k^2 - (k_0 - k_1 k) \cdot t_k)$$

□

Príklad 3.11.3 Je daný proces $\mathbf{f} = (2, 4, 5, 13)$. Vytvorte pre funkciu

$\varphi(t_k) = c_0 + c_1 \cdot t_k + c_2 \ln t_k$ bázu \mathcal{B} zodpovedajúcu časom $\mathbf{t} = (1, 2, 3, 4)$. Ortogonalizujte pomocou Gram-Schmidtovej metódy bázu \mathcal{B} . Vyjadrite koeficienty c'_0, c'_1, c'_2 procesu $\mathbf{f} = (2, 4, 5, 13)$ v ortogonálnej báze \mathcal{B}' . Pomocou koeficientov c'_0, c'_1, c'_2 vyjadrite koeficienty c_0, c_1, c_2 , ktoré je možné použiť na vyjadrenie funkcie $\varphi(t_k) = c_0 + c_1 \cdot t_k + c_2 \ln t_k$.

Riešenie:

$$\mathcal{B} = ((1, 1, 1, 1), (1, 2, 3, 4), (0, 0.69, 1.1, 1.39))$$

$$\mathcal{B}' = (\mathbf{b}'_0, \mathbf{b}'_1, \mathbf{b}'_2), \quad \mathbf{b}'_0 = (1, 1, 1, 1)$$

$$\mathbf{b}'_1 = \mathbf{b}_1 - k \cdot \mathbf{b}_0, \quad \text{kde} \quad k = \frac{\langle \mathbf{b}_1, \mathbf{b}'_0 \rangle}{\langle \mathbf{b}'_0, \mathbf{b}'_0 \rangle} = 2.5$$

$$\mathbf{b}'_1 = (1, 2, 3, 4) - 2.5 \cdot (1, 1, 1, 1) = (-1.5, -0.5, 0.5, 1.5)$$

$$\mathbf{b}'_2 = \mathbf{b}_2 - k_0 \cdot \mathbf{b}'_0 - k_1 \cdot \mathbf{b}'_1, \text{ kde } k_0 = \frac{\langle \mathbf{b}_2, \mathbf{b}'_0 \rangle}{\langle \mathbf{b}'_0, \mathbf{b}'_0 \rangle} = 0.8 \text{ a } k_1 = \frac{\langle \mathbf{b}_2, \mathbf{b}'_1 \rangle}{\langle \mathbf{b}'_1, \mathbf{b}'_1 \rangle} = 0.46$$

$$\begin{aligned} \mathbf{b}'_2 &= (0, 0.69, 1.1, 1.39) - 0.8 \cdot (1, 1, 1, 1) - 0.46 \cdot (-1.5, -0.5, 0.5, 1.5) = \\ &= (-0.11, 0.13, 0.08, -0.09) \end{aligned}$$

Koeficienty odhadu $\tilde{\mathbf{f}} = (2.3, 2.69, 6.74, 12.28)$ procesu $\mathbf{f} = (2, 4, 5, 13)$ v ortogonálnej báze \mathcal{B}' sú

$$c'_0 = 6, c'_1 = 3.4, c'_2 = -12.7$$

Dosaďme do vyjadrenia procesu pôvodné báзовé vektory.

$$\begin{aligned} \tilde{\mathbf{f}} &= c'_0 \cdot \mathbf{b}'_0 + c'_1 \cdot \mathbf{b}'_1 + c'_2 \cdot \mathbf{b}'_2 = c'_0 \cdot \mathbf{b}_0 + c'_1 \cdot (\mathbf{b}_1 - k \cdot \mathbf{b}_0) + c'_2 \cdot (\mathbf{b}_2 - k_0 \cdot \mathbf{b}_0 - k_1 \cdot (\mathbf{b}_1 - k \cdot \mathbf{b}_0)) \\ &= (c'_0 - c'_1 \cdot k - c'_2 \cdot k_0 + c'_2 \cdot k_1 \cdot k) \mathbf{b}_0 + (c'_1 - c'_2 \cdot k_1) \mathbf{b}_1 + c'_2 \cdot \mathbf{b}_2 \end{aligned}$$

Platí zároveň

$$\tilde{\mathbf{f}} = c_0 \cdot \mathbf{b}_0 + c_1 \cdot \mathbf{b}_1 + c_2 \cdot \mathbf{b}_2$$

A odiaľ dostávame

$$c_0 = c'_0 - c'_1 \cdot k - c'_2 \cdot k_0 + c'_2 \cdot k_1 \cdot k = -6.9, c_1 = c'_1 - c'_2 \cdot k_1 = 9.2, c_2 = c'_2 = -12.7$$

Odhad procesu \mathbf{f} je daný funkciou $\varphi(t_k) = c_0 + c_1 \cdot t_k + c_2 \ln t_k$ s koeficientmi $c_0 = -6.9$, $c_1 = 9.2$, $c_2 = -12.7$. Regresná krivka na odhad procesu je na obrázku 3.19. \square

3.12 Linearizácia nelineárnych úloh

Metódou lineárnej regresie nevyriešime napríklad

$$\varphi(t_k) = t_k^{c_0} \ln(c_1 \cdot t_k + c_2)$$

lebo tento vzťah sa nedá vyjadriť ako lineárna kombinácia bázych vektorov s koeficientmi c_0, c_1, c_2 .

Iná je situácia napríklad pre

$$\varphi(t_k) = c_0 \cdot e^{-c_1 \cdot t_k}$$

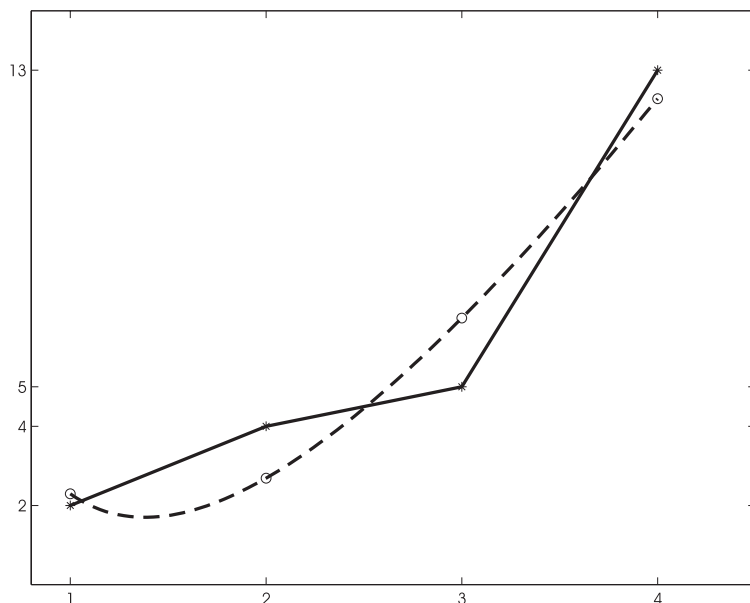
ktorú vieme linearizovať tak, že výraz logaritmujeme

$$\ln \varphi(t_k) = (\ln c_0) - c_1 \cdot t_k$$

Substitúciou $\ln \varphi(t_k) = \varphi^*(t_k)$ dostaneme $\varphi^*(t_k) = -c_1 \cdot t_k + \ln c_0$.

Teda ak sa dá $\varphi(t_k)$ prepísať na tvar

$$\varphi^*(t_k) = \sum_{n=0}^{M-1} c_n \cdot \varphi_n(t_k)$$



Obr. 3.19. Aproximácia procesu krivkou $\varphi(t_k) = -6.9 + 9.2 \cdot t_k - 12.7 \ln t_k$

môžeme pôvodnú úlohu riešiť ako úlohu lineárnej regresie.

Hovoríme, že nelineárnu úlohu môžeme **linearizovať**.

Poznámka:

Na prevod nelineárnej úlohy na úlohu lineárnej regresie sa môže použiť napríklad logaritmus, arcsin, prevrátená hodnota...

Príklad 3.12.1 *Nech sú v časoch $\mathbf{t} = (1, 2, 3, 4)$ namerané hodnoty procesu $\mathbf{f} = (9, 5, 3, 2)$. Nelineárnu úlohu aproximovať proces \mathbf{f} pomocou funkcie*

$$\varphi(t_k) = \frac{c_0 \cdot t_k}{c_1 + t_k}$$

prerobte na úlohu lineárnej regresie.

Riešenie:

Použijeme prevrátenú hodnotu nelineárnej funkcie $\varphi(t_k)$ a dostaneme lineárnu funkciu:

$$\varphi^*(t_k) = \frac{1}{\varphi(t_k)} = \frac{c_1 + t_k}{c_0 \cdot t_k} \quad (3.12.1)$$

Rovnakú úpravu urobíme s hodnotami procesu $\mathbf{f} = (9, 5, 3, 2)$ a dostaneme $\mathbf{f}^* = (0.11, 0.2, 0.33, 0.5)$.

Úpravami vzťahu (3.12.1) dostaneme úlohu lineárnej regresie

$$\varphi^*(t_k) = \frac{c_1}{c_0 \cdot t_k} + \frac{t_k}{c_0 \cdot t_k} = \frac{c_1}{c_0 \cdot t_k} + \frac{1}{c_0}$$

$$\tilde{\mathbf{f}}^* = d_0 \cdot \mathbf{b}_0 + d_1 \cdot \mathbf{b}_1$$

kde

$$d_0 = \frac{c_1}{c_0}, \quad d_1 = \frac{1}{c_0}, \quad \mathbf{b}_0 = \left(\frac{1}{t_0}, \frac{1}{t_1}, \frac{1}{t_2}, \frac{1}{t_3}\right) = \left(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}\right), \quad \mathbf{b}_1 = (1, 1, 1, 1)$$

Koeficienty d_0, d_1 sú riešením sústavy rovníc:

$$\begin{aligned} \langle \mathbf{f}^*, \mathbf{b}_0 \rangle &= d_0 \cdot \langle \mathbf{b}_0, \mathbf{b}_0 \rangle + d_1 \cdot \langle \mathbf{b}_1, \mathbf{b}_0 \rangle \\ \langle \mathbf{f}^*, \mathbf{b}_1 \rangle &= d_0 \cdot \langle \mathbf{b}_0, \mathbf{b}_1 \rangle + d_1 \cdot \langle \mathbf{b}_1, \mathbf{b}_1 \rangle \end{aligned}$$

$$0.45 = d_0 \cdot 1.42 + d_1 \cdot 2.08$$

$$1.14 = d_0 \cdot 2.08 + d_1 \cdot 4$$

$$d_0 = 0.44, \quad d_1 = 0.51$$

$$\tilde{\mathbf{f}}^* = -0.44 \cdot \mathbf{b}_0 + 0.51 \cdot \mathbf{b}_1 = (0.08, 0.3, 0.37, 0.41)$$

Treba si uvedomiť, že lineárnu regresiu sme robili v pokrivenej, transformovanej oblasti. Odhadovaný vektor \mathbf{f}^* aj jeho aproximácia $\tilde{\mathbf{f}}^*$ sa nachádzajú v inom priestore ako pôvodné hodnoty procesu, viď obrázok 3.20.

Odhad pôvodného procesu dostaneme po spätnej transformácii procesu $\tilde{\mathbf{f}}^*$ na proces $\tilde{\mathbf{f}}$, viď obrázok 3.21.

$$\tilde{\mathbf{f}} = \frac{1}{\tilde{\mathbf{f}}^*} = (13.26, 3.39, 2.71, 2.47)$$

Iný postup, ktorý môžeme použiť je vypočítať najskôr hodnoty koeficientov c_0, c_1 využitím vzťahov

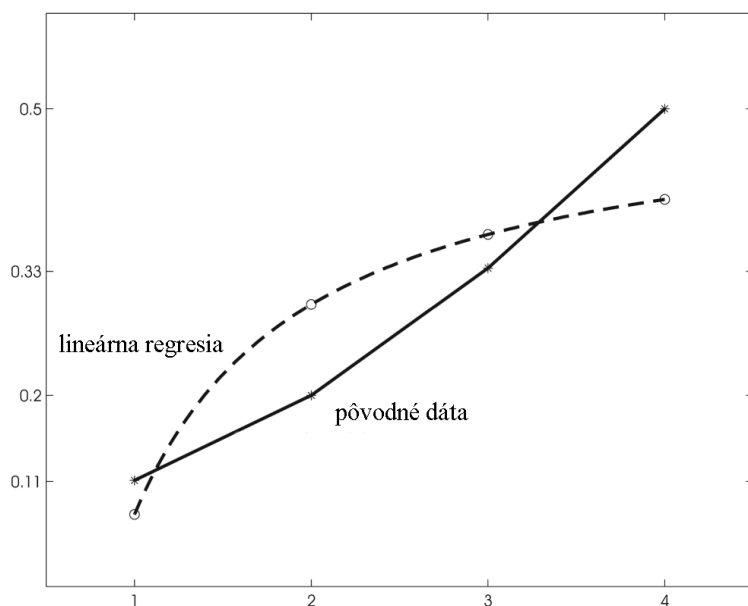
$$d_0 = \frac{c_1}{c_0}, \quad d_1 = \frac{1}{c_0}.$$

$$-0.44 = \frac{c_1}{c_0}, \quad 0.51 = \frac{1}{c_0}$$

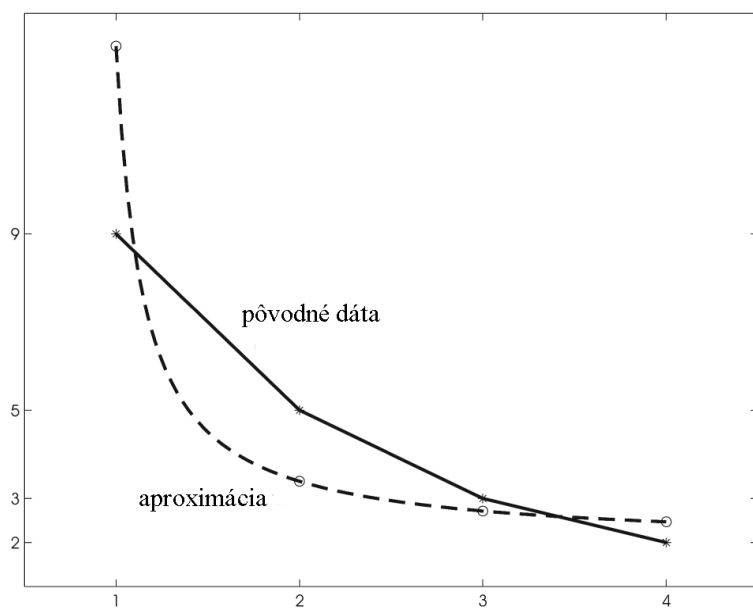
$$c_1 = -0.86, \quad c_0 = 1.95$$

Poznámka:

Je potrebné si uvedomiť, že po linearizácii úlohy už nehľadáme najbližší vektor v metrike pôvodného priestoru. Veľkosť chyby aproximácie bude minimálna iba v zmenenej metrike nového priestoru (po transformácii) a nebude vyjadrená v hodnotách pôvodného priestoru.



Obr. 3.20. Aproximácia procesu regresnou krivkou $\varphi^*(t_k) = \frac{d_0}{t_k} + d_1$



Obr. 3.21. Aproximácia procesu krivkou $\varphi(t_k) = \frac{c_0 \cdot t_k}{c_1 + t_k}$

Príklad 3.12.2 V príklade 3.12.1 je minimalizovaná veľkosť chyby aproximácie $\tilde{\mathbf{f}}^*$ procesu \mathbf{f}^* . Veľkosť chyby v transformovanom priestore je

$$\|\mathbf{e}^*\| = \|\mathbf{f}^* - \tilde{\mathbf{f}}^*\| = 0.14$$

Veľkosť chyby nelineárnej aproximácie $\tilde{\mathbf{f}}$ procesu \mathbf{f} v pôvodnom priestore je

$$\|\mathbf{e}\| = \|\mathbf{f} - \tilde{\mathbf{f}}\| = 4.58$$

Zistite, či veľkosť chyby v pôvodnom priestore je minimálna, alebo či sa dajú nájsť koeficienty c_0 a c_1 také, že chyba aproximácie procesu krivkou

$$\varphi(t_k) = \frac{c_0 \cdot t_k}{c_1 + t_k}$$

bude menšia ako 4.58.

Riešenie:

Hľadanými koeficientmi sú napríklad koeficienty $c_0 = 2.2$ a $c_1 = -0.7$. Pomocou týchto koeficientov dostaneme vektor

$$\tilde{\mathbf{f}} = \varphi(\mathbf{t}) = \frac{c_0 \cdot \mathbf{t}}{c_1 + \mathbf{t}} = (7.33, 3.38, 2.87, 2.67)$$

Veľkosť chyby odhadu je v tomto prípade:

$$\|\mathbf{e}\| = \|\mathbf{f} - \tilde{\mathbf{f}}\| = 2.42$$

Dosiahli sme chybu, ktorá je menšia, než akú sme dosiahli metódou linearizácie nelineárnej úlohy. Naša metóda však spočívala iba v náhodnom hľadaní. Takým hľadaním sa dokonca dajú nájsť koeficienty c_0 a c_1 , pre ktoré je veľkosť chyby menšia než 1.54. Vyskúšajte to! \square

4 Modely využívajúce súvislosti medzi hodnotami procesu

4.1 Modely kĺzavých súčtov

V predchádzajúcej kapitole sme zisťovali, ako sa v procese prejaví časová závislosť. Jednotlivé modely predstavovali explicitné funkcie času, ktoré aproximovali proces. Ako dobre sme proces aproximovali sa dalo zistiť podľa veľkosti chyby, teda veľkosti rozdielu pôvodného procesu a jeho aproximácie. Našou úlohou bolo nájsť tvar a parametre funkcie času tak, aby chyba aproximácie pomocou tejto funkcie bola minimálna.

V tejto kapitole nebudeme sledovať závislosť procesu od času, ale súvislosť hodnôt procesu medzi sebou navzájom. Riešime otázku ako súvisia jednotlivé hodnoty procesu. Súvislosť medzi hodnotami procesu by mohla byť ovplyvnená tým, že hodnoty procesu by boli merané v nepravidelných časoch. Aby sme zamedzili tomuto vplyvu, budeme v tejto kapitole predpokladať, že hodnoty procesu sú merané v ekvidištancných časoch a dĺžka časového intervalu bude 1. V nasledujúcom texte budeme hodnoty $f(t_k)$ značiť iba ako f_k .

Motivačnou úlohou tejto kapitoly je **interpolácia procesu**. Pod interpoláciou rozumieme vyhladenie nameraných hodnôt tak, že ich nahradíme vhodnou krivkou a dopočítame chýbajúce hodnoty pomocou vyjadrenia tejto krivky ako funkcie času. V tejto kapitole sa však nebudeme zaoberať vyjadrením procesu ako explicitnej funkcie času. Našou úlohou bude nájsť súvislosť medzi koeficientmi a vyjadriť chýbajúcu hodnotu pomocou ostatných hodnôt.

Nech sú namerané hodnoty $f_0, f_1, \dots, f_{k-1}, f_{k+1}, \dots, f_{N-1}$ procesu \mathbf{f} , ako to vidíme na obrázku 4.1.

Chceme odhadnúť hodnotu f_k , ktorá nie je nameraná. Predpokladáme, že hodnoty f_{k-1}, f_k a f_{k+1} sa príliš nelíšia. V jednoduchších modeloch, keď nezisťujeme nič o tom ako hodnoty medzi sebou súvisia, sa hodnota f_k vypočíta ako aritmetický priemer susedných hodnôt:

$$\tilde{f}_k = \frac{f_{k-1} + f_{k+1}}{2}$$

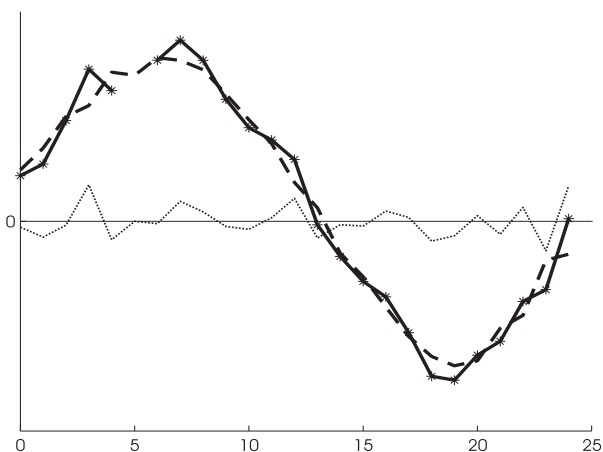
Priebeh procesu s chýbajúcou hodnotou a jeho aproximácia procesom, v ktorom sa použije na výpočet aritmetický priemer dvoch susedných hodnôt, sú zobrazené na obrázku 4.2. Na obrázku je prerušovanou čiarou znázornený aj chybový proces, ktorý vypočítame ako rozdiel pôvodného a odhadnutého procesu.

$$\|\mathbf{e}_1\| = \|\mathbf{f} - \tilde{\mathbf{f}}\| = \sqrt{\sum_{j \neq k} (f_j - \tilde{f}_j)^2}$$



Obr. 4.1. Pribeh procesu s chýbajúcou hodnotou

Veľkosť chyby je v tomto prípade 7.77, odhadnutú hodnotu \tilde{f}_k sme pri výpočte chyby nepoužili.



Obr. 4.2: Pribeh procesu, jeho aproximácia pomocou aritmetického priemeru a chybový vektor

Pokúsime sa teraz úlohu vyriešiť všeobecnejšie. Budeme hľadať také hodnoty c_0, c_1 , pre ktoré proces $\tilde{\mathbf{f}}$, so zložkami

$$\tilde{f}_k = c_0 f_{k-1} + c_1 f_{k+1}$$

bude najlepšie aproximovať pôvodný proces \mathbf{f} .

Vektorový zápis:

$$\tilde{\mathbf{f}} = c_0 \mathbf{b}_0 + c_1 \mathbf{b}_1$$

kde

$$\begin{aligned} \mathbf{f} &= (f_0, f_1, \dots, f_{k-1}, *, f_{k+1}, \dots, f_{N-1}) \\ \mathbf{b}_0 &= (*, f_0, \dots, f_{k-2}, f_{k-1}, *, \dots, f_{N-2}) \\ \mathbf{b}_1 &= (f_1, f_2, \dots, *, f_{k+1}, f_{k+2}, \dots, *) \end{aligned}$$

To nastane vtedy, keď

$$\|\mathbf{f} - \tilde{\mathbf{f}}\| \rightarrow \min$$

odtiaľ

$$(\mathbf{f} - \tilde{\mathbf{f}}) \perp \mathbf{b}_0 \quad \text{a} \quad (\mathbf{f} - \tilde{\mathbf{f}}) \perp \mathbf{b}_1$$

a teda

$$\langle \mathbf{f} - \tilde{\mathbf{f}}, \mathbf{b}_0 \rangle = 0 \quad \text{a} \quad \langle \mathbf{f} - \tilde{\mathbf{f}}, \mathbf{b}_1 \rangle = 0$$

$$\langle \mathbf{f} - c_0 \cdot \mathbf{b}_0 - c_1 \cdot \mathbf{b}_1, \mathbf{b}_0 \rangle = 0$$

$$\langle \mathbf{f} - c_0 \cdot \mathbf{b}_0 - c_1 \cdot \mathbf{b}_1, \mathbf{b}_1 \rangle = 0$$

$$\langle \mathbf{f}, \mathbf{b}_0 \rangle - \langle c_0 \mathbf{b}_0, \mathbf{b}_0 \rangle - \langle c_1 \mathbf{b}_1, \mathbf{b}_0 \rangle = 0$$

$$\langle \mathbf{f}, \mathbf{b}_1 \rangle - \langle c_0 \mathbf{b}_0, \mathbf{b}_1 \rangle - \langle c_1 \mathbf{b}_1, \mathbf{b}_1 \rangle = 0$$

$$\langle \mathbf{f}, \mathbf{b}_0 \rangle - c_0 \langle \mathbf{b}_0, \mathbf{b}_0 \rangle - c_1 \langle \mathbf{b}_1, \mathbf{b}_0 \rangle = 0$$

$$\langle \mathbf{f}, \mathbf{b}_1 \rangle - c_0 \langle \mathbf{b}_0, \mathbf{b}_1 \rangle - c_1 \langle \mathbf{b}_1, \mathbf{b}_1 \rangle = 0$$

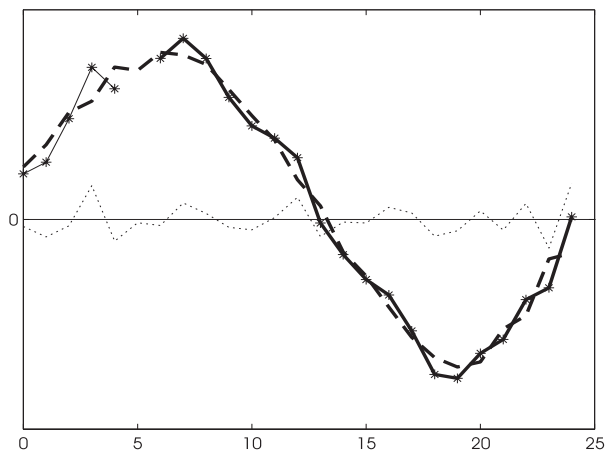
$$c_0 \langle \mathbf{b}_0, \mathbf{b}_0 \rangle + c_1 \langle \mathbf{b}_1, \mathbf{b}_0 \rangle = \langle \mathbf{f}, \mathbf{b}_0 \rangle$$

$$c_0 \langle \mathbf{b}_0, \mathbf{b}_1 \rangle + c_1 \langle \mathbf{b}_1, \mathbf{b}_1 \rangle = \langle \mathbf{f}, \mathbf{b}_1 \rangle$$

Dostaneme sústavu dvoch lineárnych rovníc o dvoch neznámych

$$\begin{aligned} c_0 \cdot \left(\sum_{j=0}^{k-1} f_j^2 + \sum_{j=k+1}^{N-2} f_j^2 \right) + c_1 \cdot \left(\sum_{j=0}^{k-3} f_j \cdot f_{j+2} + f_{k-1} \cdot f_{k+1} + \sum_{j=k+1}^{N-3} f_j \cdot f_{j+2} \right) \\ = \sum_{j=1, j \neq k, j \neq k+1}^{N-1} f_j \cdot f_{j-1} \\ c_0 \cdot \left(\sum_{j=0}^{k-3} f_j \cdot f_{j+2} + f_{k-1} \cdot f_{k+1} + \sum_{j=k+1}^{N-3} f_j \cdot f_{j+2} \right) + c_1 \cdot \left(\sum_{j=1}^{k-1} f_j^2 + \sum_{j=k+1}^{N-1} f_j^2 \right) \\ = \sum_{j=0, j \neq k, j \neq k-1}^{N-2} f_j \cdot f_{j+1} \end{aligned} \quad (4.1.1)$$

Riešením tejto sústavy budú koeficienty procesu $\tilde{\mathbf{f}}$ v báze $\mathcal{B} = \{\mathbf{b}_0, \mathbf{b}_1\}$. Priebeh procesu s chýbajúcou hodnotou a jeho aproximácia procesom, ktorý použije na výpočet priemet do podpriestoru pomocou dvoch susedných hodnôt, sú zobrazené na obrázku 4.3. Na obrázku je prerušovaná čiarou znázornený aj chybový proces, ktorého veľkosť je teraz menej ako v predošlom prípade, $\|\mathbf{e}_2\| = 7.69$.



Obr. 4.3: Priebeh procesu, jeho aproximácia pomocou priemetu do podpriestoru a chybový vektor

Hviezdičky na začiatku a na konci bazových procesov môžeme nahradiť hodnotami podľa povahy procesu, alebo pri výpočte skalárneho súčinu vynechať príslušnú hodnotu vo všetkých vektoroch. Rovnako vynecháme vo všetkých procesoch súradnice zodpovedajúce chýbajúcej hodnote f_k .

Metóda získania koeficientov c_0, c_1 spočíva v hľadaní takého riešenia sústavy (4.1.1), pri ktorom chyba aproximácie pôvodného procesu \mathbf{f} procesom $\tilde{\mathbf{f}}$ bude minimálna. Koeficienty, ktoré minimalizujú veľkosť chyby sa môžu počítať pomocou diferenciálneho počtu rovnako dobre ako metódou priemetu do podpriestoru. Uvedený postup sa nazýva metóda **kľavých súčtov**.

Aritmetický alebo vážený priemer sa používa aj pri známej hodnote f_k . V prípade, že chceme proces vyhladiť, riešime úlohu **filtrácie procesu**. Predpokladajme, že namerané hodnoty sú

$$f_0, f_1, \dots, f_{k-1}, f_k, f_{k+1}, \dots, f_{N-1}$$

ale hodnota f_k je výrazne odlišná a chceme ju prispôbiť ostatným hodnotám. Takáto situácia nastane napríklad keď chceme odstrániť šum na zvukovom zázname. Najjednoduchší postup bude, ak hodnotu f_k nahradíme aritmetickým priemerom susedných hodnôt:

$$\tilde{f}_k = \frac{f_{k-1} + f_{k+1}}{2}$$

To môžeme napísať aj ako

$$\tilde{f}_k = \frac{1}{2}f_{k-1} + \frac{1}{2}f_{k+1}$$

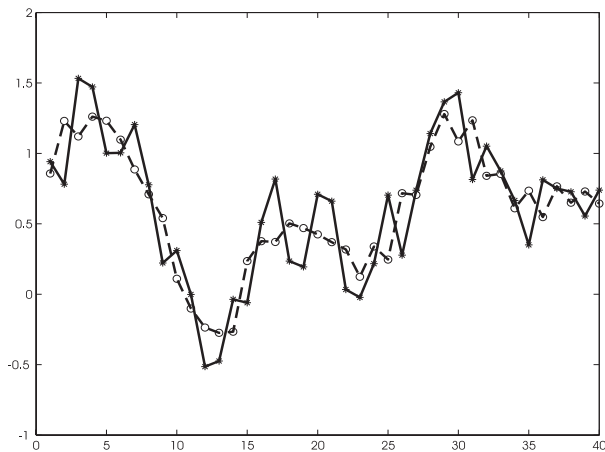
Všeobecnejší zápis a presnejší výsledok dostaneme, ak namiesto koeficientov $\frac{1}{2}$ použijeme vhodné koeficienty c_0 a c_1 .

$$\tilde{f}_k = c_0 \cdot f_{k-1} + c_1 \cdot f_{k+1}$$

Úloha je, ako zvoliť c_0, c_1 , aby sa filtrovaný proces, čo najviac podobal na pôvodný proces \mathbf{f} . Nový proces bude mať súradnice

$$\tilde{\mathbf{f}} = (c_0, c_1)\mathcal{B}$$

v báze \mathcal{B} , ktorú treba vhodne zvoliť. Pôjde teda o priemet do dvojrozmerného vektorového priestoru, určeného bázou \mathcal{B} . Príklad procesu a jeho filtrácie je na obrázku 4.4



Obr. 4.4: Vyhladené dáta pomocou metódy kľzavých súčtov, pôvodné dáta sú označené * a vyhladené sú označené o, použitý je model $f_k = c_0 f_{k-1} + c_1 f_{k+1}$

Príklad 4.1.1 Interpolácia a filtrácia sa používa napríklad aj pri úprave obrázkov. Pri digitálnom prenose čiernobielych obrázkov môžu okrem iných vzniknúť tieto dva druhy zašumenia:

1. Pre niektoré pixely nedostaneme informáciu o hodnote intenzity sivej farby, pričom vieme, pre ktoré pixely táto informácia chýba.
2. To isté čo predtým, len nevieme, ktorým pixelom informácia chýba, prípadne ktoré pixely obsahujú chybnú informáciu.

Na obr. 4.5 vidíme (a) originálny obrázok a (b) obrázok zašumený. Veľkosť obrázka je 256×256 pixelov a pre každý pixel je odtieň sivej reprezentovaný prirodzeným číslom medzi 0 a 255. Použite metódu kľzavých súčtov pri rekonštrukcii obrázka.

Riešenie:

Obrázok si reprezentujeme ako maticu \mathbf{B} . V prípade, že poznáme polohu šumu, tj. súradnice bielych bodov (s_x, s_y) , môžeme pre chýbajúce hodnoty priradiť aritmetický priemer susedných bodov, a teda používame interpoláciu

$$M(s_x, s_y) = \frac{1}{4} (M(s_x - 1, s_y) + M(s_x + 1, s_y) + M(s_x, s_y - 1) + M(s_x, s_y + 1))$$

Takto zrekonštruovaný obrázok je znázornený na obr. 4.5 (c). Vidíme, že obrázok sa podarilo zrekonštruovať veľmi dobre. Horší výsledok sa ale dá očakávať, keď polohu chýbajúcich bodov nepoznáme. Tu už musíme priemerovať všetky pixely, a teda prebehne filtrácia procesu.

Pre lepší výsledok budeme priemerovať 9 pixelov vo štvorci 3×3 okolo jedného bodu

$$M(p, s) = \frac{1}{9} \left(M(p-1, s-1) + M(p-1, s) + M(p-1, s+1) \right. \\ \left. + M(p, s-1) + M(p, s) + M(p, s+1) \right. \\ \left. + M(p+1, s-1) + M(p+1, s) + M(p+1, s+1) \right)$$

pre všetky $p = 0, \dots, 255$ a $s = 0, \dots, 255$.

Výsledok je zobrazený na obr. 4.5 (d). Vidíme, že rekonštrukcia má stále príliš jasné škvrny v miestach pôvodného šumu. Preto upravíme náš filtrovací proces a budeme priemerovať 25 pixelov vo štvorci 5×5 okolo daného bodu.

Výsledok na obr. 4.5 (e) sa mierne zlepšil.

Ďalším zväčšením štvorca na 7×7 získame už príliš rozmazný obrázok, viď obr. 4.5 (f).



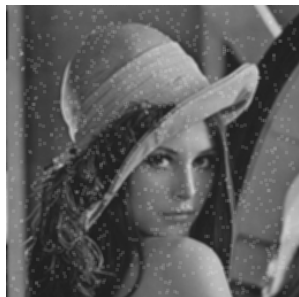
(a) pôvodný obrázok



(b) obrázok so šumom



(c) interpolácia



(d) filtrácia 3×3

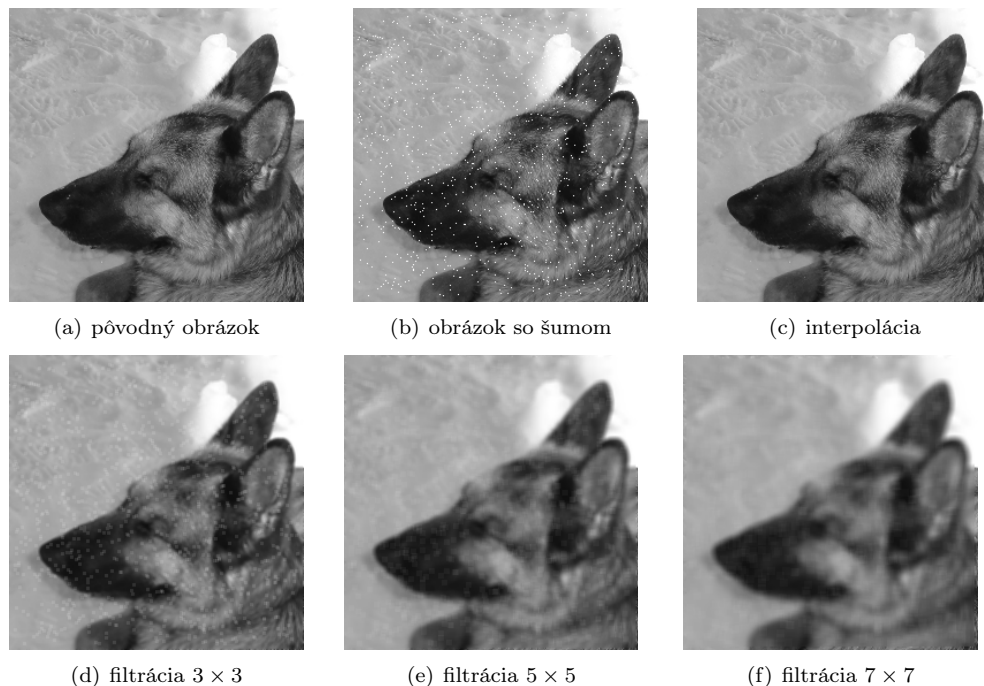


(e) filtrácia 5×5



(f) filtrácia 7×7

Obr. 4.5: Benchmarkový obrázok dievčiny (Lena), ktorý sme filtrovali metódou klzavých súčtov



Obr. 4.6. Bežný obrázok psa (Dena) a jeho filtrácia

Poznámky:

- Pri obrázkoch (d) a (e) vidno, že sú mierne rozmazané. Je to dôsledok toho, že hodnoty boli priemerované. V mieste prechodu medzi dvoma farbami by mal byť ostrý skok, avšak filtráciou sme tento skok zhladili.
- Obrázok (e) je rozmazanejší ako obrázok (d), pretože sme používali väčšiu časť obrázka na priemerovanie.
- Znalosť polohy šumu sa ukázala ako veľká výhoda. Nielenže sme priemerovali len v miestach šumu - a teda sme ostrosť obrázka mimo zašumených miest nestratili - ale zároveň sme chybu neposúvali do vedľajších pixelov. Bohužiaľ informácia o polohe šumu je známa len zriedka.
- Ak porovnáme obr. 4.5 a obr. 4.6 vidíme, že obrázok psa sa ťažšie rekonštruoval. Buď sme mali nedostatočne odstránený šum, alebo bol výsledný obrázok príliš rozmazaný. Je to preto, lebo originálny obrázok psa obsahuje menej súvislých plôch, kde sa farba príliš nemení. Zvieracia srst' obsahuje veľa skokov vo farbe.

Ďalej budeme riešiť úlohu **extrapolácie procesu**. Predpokladajme, že namerané hodnoty procesu f sú f_0, f_1, \dots, f_{k-1} . Chceme odhadnúť (predpovedať) ďalšiu hod-

notu f_k , ktorú odhadujeme pomocou M predchádzajúcich hodnôt:

$$\tilde{f}_k = \sum_{j=1}^M c_j \cdot f_{k-j}$$

Príklad 4.1.2 Ako bude vyzeráť prediktor prvého rádu pre úlohu kľzavých súčtov?

Riešenie:

Ak $M = 1$, dostaneme $\tilde{f}_k = c \cdot f_{k-1}$.

Namerané sú hodnoty pôvodného procesu $\mathbf{f} = (f_0, f_1, \dots, f_k, \dots, f_{N-1})$, nový proces bude mať hodnoty

$$\tilde{\mathbf{f}} = (*, c f_0, c f_1, \dots, c f_k, \dots, c f_{N-2})$$

Báza $\mathcal{B} = \{\mathbf{b}_0\}$ bude obsahovať jediný vektor

$$\mathbf{b}_0 = (*, f_0, f_1, \dots, f_k, \dots, f_{N-2}),$$

pričom prvú zložku vektora \mathbf{b}_0 zanedbáme (hodnotu "*" nahradíme hodnotou 0).

$$c = \frac{\langle \mathbf{f}, \mathbf{b}_0 \rangle}{\langle \mathbf{b}_0, \mathbf{b}_0 \rangle} = \frac{\langle (f_0, f_1, \dots, f_k, \dots, f_{N-1}), (0, f_0, f_1, \dots, f_k, \dots, f_{N-2}) \rangle}{\langle (0, f_0, f_1, \dots, f_k, \dots, f_{N-2}), (0, f_0, f_1, \dots, f_k, \dots, f_{N-2}) \rangle}$$

$$c = \frac{\sum_{k=0}^{N-1} f_k \cdot f_{k-1}}{\sum_{k=0}^{N-1} f_{k-1}^2}$$

To bol prediktor, ktorý používa jednu hodnotu. Nazývame ho tiež prediktor prvého rádu. \square

Úloha 4.1.1 Odvodte vzorec pre prediktor, ktorý používa tri hodnoty, teda prediktor tretieho rádu. Tieto hodnoty môžu byť tri posledné prechádzajúce hodnoty. Ale keď napríklad sledujeme vývoj peňažného kurzu, je lepšie brať do úvahy predošlý deň, predošlý týždeň a predošlý mesiac.

Porovnajete účinnosť navrhovaných modelov na nejakých reálnych dátach.

4.2 Všeobecná formulácia úlohy

Vo všeobecnosti riešime úlohu lineárnej regresie pre model

$$\varphi(t_k) = \sum_{j=1}^M c_j \cdot f_j \tag{4.2.1}$$

V tomto modeli na rozdiel od klasickej úlohy lineárnej regresie v tvare $\varphi(t_k) = \sum_{j=1}^M c_j \cdot \varphi_j(t_k)$ nie sú dopredu určené regresné funkcie $\varphi_j(t_k)$. Namiesto nich sú ako

báza použité známe (namerané) hodnoty procesu f_j v nejakých určených časových okamihoch.

Predpokladáme, že hodnoty sú merané v pravidelných intervaloch, preto pre riešenie úloh kľzavých súčtov je dôležité dodržať podmienku ekvidištancnosti časových okamihov, v ktorých sa merajú hodnoty procesu.

$$t_{j+1} = t_j + \Delta t, \quad j = 1, 2, \dots, N$$

Predchádzajúcu podmienku môžeme normalizovať a dostávame postupnosť časových okamihov:

$$t_1 = 1, t_{j+1} = t_j + 1, \quad j = 1, 2, \dots, N - 1$$

Podľa vzájomného vzťahu hodnôt k a M vo vyjadrení modelu (4.2.1) budeme riešiť úlohy:

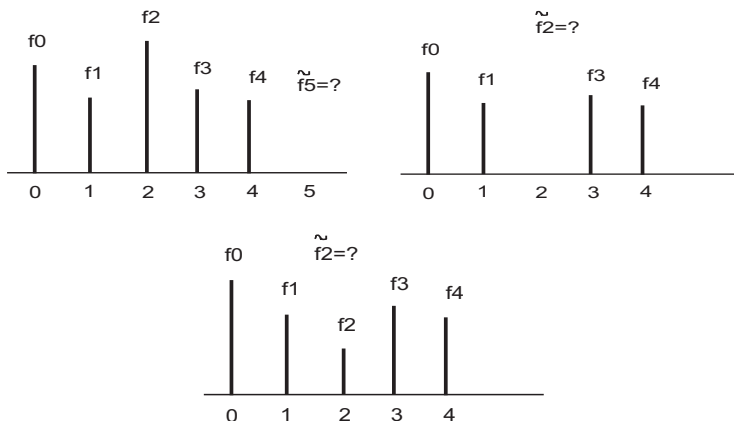
$$\begin{array}{ll} k > M, & \text{extrapolácia} \\ 1 < k < M, & \text{interpolácia} \\ 1 \leq k \leq M, & \text{filtrácia} \end{array}$$

Extrapolácia znamená odhadnúť nasledujúcu (budúcu) hodnotu procesu, ak poznáme jednu, alebo niekoľko predchádzajúcich hodnôt.

Interpolácia znamená odhadnúť chýbajúcu hodnotu ak sú známe hodnoty pred a po tejto neznámej hodnote.

Filtrácia znamená nahradenie nameranej hodnoty procesu inou, takou, ktorá lepšie zodpovedá ostatným hodnotám procesu.

Schématický náčrt týchto postupov je na obrázku 4.7



Obr. 4.7. Extrapolácia, interpolácia a filtrácia procesu

Extrapolácia (predikcia)

Slovo extrapolácia pochádza z latinského extrapolatus a druhá časť slova znamená vyleštiť, vyčistiť, predpona extra tu znamená dopredu, alebo mimo oblasť známych

hodnôt. Rieši sa úloha odhadnúť hodnotu procesu v nasledujúcom časovom okamihu, preto sa táto metóda nazýva aj predikcia. V úlohe sa odhaduje hodnota procesu pomocou vzťahu:

$$\tilde{f}_{k+1} = \sum_{j=0}^{M-1} c_j \cdot f_{k-j} \quad (4.2.2)$$

Uvažujme napríklad $M = 2$.

$$\tilde{f}_{k+1} = c_0 \cdot f_k + c_1 \cdot f_{k-1} \quad (4.2.3)$$

$$\begin{aligned} \mathbf{f} &= (f_0, f_1, f_2, \dots, f_{N-1}) \\ \tilde{\mathbf{f}} &= (0, 0, c_0 \cdot f_1 + c_1 \cdot f_0, c_0 \cdot f_2 + c_1 \cdot f_1, \dots, c_0 \cdot f_{N-1-1} + c_1 \cdot f_{N-1-2}) \end{aligned}$$

$$\tilde{\mathbf{f}} = c_0 \cdot (0, 0, f_1, f_2, \dots, f_{N-1-1}) + c_1 \cdot (0, 0, f_0, f_1, \dots, f_{N-1-2})$$

$$\begin{aligned} \mathbf{b}_0 &= (0, 0, f_1, f_2, \dots, f_{N-1-1}) \\ \mathbf{b}_1 &= (0, 0, f_0, f_1, \dots, f_{N-1-2}) \end{aligned}$$

$$d(\tilde{\mathbf{f}}, \mathbf{f}) \rightarrow \min; \quad d(c_0 \cdot \mathbf{b}_0 + c_1 \cdot \mathbf{b}_1, \mathbf{f}) \rightarrow \min$$

$$\begin{aligned} \mathbf{f} - (c_0 \cdot \mathbf{b}_0 + c_1 \cdot \mathbf{b}_1) &\perp \mathbf{b}_i; & i = 0, 1 \\ \langle \mathbf{f} - c_0 \cdot \mathbf{b}_0 - c_1 \cdot \mathbf{b}_1, \mathbf{b}_i \rangle &= 0; & i = 0, 1 \end{aligned}$$

$$\begin{aligned} i = 0 : \\ \langle (0, 0, f_2 - c_0 f_1 - c_1 f_0, \dots, f_{N-1} - c_0 f_{N-1-1} - c_1 f_{N-1-2}), \\ (0, 0, f_1, \dots, f_{N-1-1}) \rangle &= 0 \\ i = 1 : \\ \langle (0, 0, f_2 - c_0 f_1 - c_1 f_0, \dots, f_{N-1} - c_0 f_{N-1-1} - c_1 f_{N-1-2}), \\ (0, 0, f_0, \dots, f_{N-1-2}) \rangle &= 0 \end{aligned}$$

$$\begin{aligned} i = 0 : \quad \sum_{k=2}^{N-1} (f_k - c_0 f_{k-1} - c_1 f_{k-2}) f_{k-1} &= 0 \\ i = 1 : \quad \sum_{k=2}^{N-1} (f_k - c_0 f_{k-1} - c_1 f_{k-2}) f_{k-2} &= 0 \end{aligned}$$

$$c_0 \sum_{k=2}^{N-1} f_{k-1}^2 + c_1 \sum_{k=2}^{N-1} f_{k-2} f_{k-1} = \sum_{k=2}^{N-1} f_k f_{k-1}$$

$$c_0 \sum_{k=2}^{N-1} f_{k-1} f_{k-2} + c_1 \sum_{k=2}^{N-1} f_{k-2}^2 = \sum_{k=2}^{N-1} f_k f_{k-2}$$

Koeficienty c_0 a c_1 , ktoré sú riešením poslednej sústavy rovníc, použijeme vo vzťahu (4.2.3) na odhad nasledujúcej hodnoty procesu.

Vráťme sa k všeobecnému tvaru (4.2.2) úlohy extrapolácie, v ktorej najprv proces \mathbf{f} odhadneme procesom $\tilde{\mathbf{f}}$ a potom využijeme koeficienty tohto odhadu na výpočet ďalších hodnôt procesu.

$$\tilde{f}_{k+1} = \sum_{j=0}^{M-1} c_j \cdot f_{k-j}$$

$\mathbf{f} =$

$$(f_0, \dots, f_{M-1}, \quad f_M, \quad \dots \quad f_k, \quad \dots \quad f_{N-1})$$

$\tilde{\mathbf{f}} =$

$$(0, \dots, 0, \quad \sum_{j=0}^{M-1} c_j \cdot f_{M-1-j} \quad \dots \quad \sum_{j=0}^{M-1} c_j \cdot f_{k-1-j} \quad \dots \quad \sum_{j=0}^{M-1} c_j \cdot f_{N-2-j})$$

Príslušná báza bude mať tvar

miesto:	0	1	...	$M-1$	M	...	k	...	$N-1$	
\mathbf{b}_0	= (0,	0	...	0,	$f_{M-1},$...	f_{k-1}	...	$f_{N-2})$
\mathbf{b}_1	= (0,	0	...	0,	$f_{M-2},$...	f_{k-2}	...	$f_{N-3})$
\vdots										
\mathbf{b}_n	= (0,	0	...	0,	$f_{M-n-1},$...	f_{k-n-1}	...	$f_{N-n-2})$
\vdots										
\mathbf{b}_{M-1}	= (0,	0	...	0,	$f_0,$...	f_{k-M-2}	...	$f_{N-M-1})$

Vzdialenosť medzi \mathbf{f} a $\tilde{\mathbf{f}}$ je minimálna, ak platí

$$\langle \mathbf{f} - \tilde{\mathbf{f}}, \mathbf{b}_n \rangle = 0, \quad \forall n = 0, 1, \dots, M-1$$

Po dosadení vektorov \mathbf{f} , $\tilde{\mathbf{f}}$ a \mathbf{b}_n dostaneme

$$\left\langle (f_0, f_1, \dots, f_{M-1}, f_M - \sum_{j=0}^{M-1} c_j \cdot f_{M-1-j}, \dots, f_k - \sum_{j=0}^{M-1} c_j \cdot f_{k-1-j}, \dots, f_{N-1} - \sum_{j=0}^{M-1} c_j \cdot f_{N-2-j}), (0, 0, \dots, 0, f_{M-n-1}, \dots, f_{k-n-1}, \dots, f_{N-n-2}) \right\rangle = 0$$

$$\sum_{k=M}^{N-1} \left(f_k - \sum_{j=0}^{M-1} c_j \cdot f_{k-1-j} \right) \cdot f_{k-1-n} = 0$$

Po roznásobení zátvorky a zámene súm dostávame vzorec na výpočet koeficientov pre úlohu extrapolácie.

$$\sum_{j=0}^{M-1} c_j \cdot \sum_{k=M}^{N-1} f_{k-1-j} \cdot f_{k-1-n} = \sum_{k=M}^{N-1} f_k f_{k-1-n} \quad n = 0, 1, \dots, M-1 \quad (4.2.4)$$

Postup, ktorý využíva vzťah (4.2.4) sa nazýva aj lineárne prediktívne kódovanie a príslušné koeficienty sa označujú ako LPC koeficienty.

Interpolácia a filtrácia

Slovo interpolácia pochádza z latinského interpolatus a druhá časť slova znamená vyleštiť, vyčistiť, predpona inter tu znamená dnu, alebo v oblasti známych hodnôt. Rieši sa úloha odhadnúť neznámu hodnotu, alebo nahradiť známu, ale nevyhovujúcu hodnotu procesu v časovom okamihu, ktorý sa nachádza medzi nameranými dátami. V úlohe filtrácie sa odhaduje hodnota procesu pomocou vzťahu:

$$\tilde{f}_k = c_0 \cdot f_k + \sum_{j=1}^M (c_j \cdot f_{k-j} + c_{-j} \cdot f_{k+j}) \quad (4.2.5)$$

V úlohe interpolácie sa odhaduje hodnota procesu pomocou vzťahu:

$$\tilde{f}_k = \sum_{j=1}^M (c_j \cdot f_{k-j} + c_{-j} \cdot f_{k+j}) \quad (4.2.6)$$

Uvažujme napríklad $M = 1$.

$$\tilde{f}_k = c_1 \cdot f_{k-1} + c_{-1} \cdot f_{k+1} \quad (4.2.7)$$

$$\mathbf{f} = (f_0, f_1, f_2, \dots, f_{N-1})$$

$$\tilde{\mathbf{f}} = (0, c_1 \cdot f_0 + c_{-1} \cdot f_2, c_1 \cdot f_1 + c_{-1} \cdot f_3, \dots, c_1 \cdot f_{N-3} + c_{-1} \cdot f_{N-1}, 0)$$

$$\tilde{\mathbf{f}} = c_1 \cdot (0, f_0, f_1, \dots, f_{N-3}, 0) + c_{-1} \cdot (0, f_2, f_3, \dots, f_{N-1}, 0)$$

$$\mathbf{b}_1 = (0, f_0, f_1, \dots, f_{N-3}, 0)$$

$$\mathbf{b}_{-1} = (0, f_2, f_3, \dots, f_{N-1}, 0)$$

$$d(\tilde{\mathbf{f}}, \mathbf{f}) \rightarrow \min; \quad d(c_1 \cdot \mathbf{b}_1 + c_{-1} \cdot \mathbf{b}_{-1}, \mathbf{f}) \rightarrow \min$$

$$\mathbf{f} - (c_1 \cdot \mathbf{b}_1 + c_{-1} \cdot \mathbf{b}_{-1}) \perp \mathbf{b}_i; \quad i = -1, 1$$

$$\langle \mathbf{f} - c_1 \cdot \mathbf{b}_1 - c_{-1} \cdot \mathbf{b}_{-1}, \mathbf{b}_i \rangle = 0; \quad i = -1, 1$$

$$i = 1 :$$

$$\langle (f_0, f_1 - c_1 f_0 - c_{-1} f_2, \dots, f_{N-2} - c_1 f_{N-3} - c_{-1} f_{N-1}, f_{N-1}),$$

$$(0, f_0, \dots, f_{N-3}, 0) \rangle = 0$$

$$i = -1 :$$

$$\langle (f_0, f_1 - c_1 f_0 - c_{-1} f_2, \dots, f_{N-2} - c_1 f_{N-3} - c_{-1} f_{N-1}, f_{N-1}),$$

$$(0, f_2, \dots, f_{N-1}, 0) \rangle = 0$$

$$i = 1 : \quad \sum_{k=1}^{N-2} (f_k - c_1 f_{k-1} + c_{-1} f_{k+1}) f_{k-1} = 0$$

$$i = -1 : \quad \sum_{k=1}^{N-2} (f_k - c_1 f_{k-1} + c_{-1} f_{k+1}) f_{k+1} = 0$$

$$c_1 \sum_{k=1}^{N-2} f_{k-1}^2 + c_{-1} \sum_{k=1}^{N-2} f_{k-1} \cdot f_{k+1} = \sum_{k=1}^{N-2} f_k \cdot f_{k-1}$$

$$c_1 \sum_{k=1}^{N-2} f_{k-1} \cdot f_{k+1} + c_{-1} \sum_{k=1}^{N-2} f_{k+1}^2 = \sum_{k=1}^{N-2} f_k \cdot f_{k+1}$$

Koeficienty c_1 a c_{-1} , ktoré sú riešením poslednej sústavy rovníc, použijeme vo vzťahu (4.2.7) na odhad chýbajúcej, alebo chybnej hodnoty procesu.

Vráťme sa k všeobecnému tvaru (4.2.6) úlohy interpolácie, v ktorej najprv proces \mathbf{f} odhadneme procesom $\tilde{\mathbf{f}}$ a potom využijeme koeficienty tohto odhadu na výpočet chýbajúcich hodnôt procesu.

$$\tilde{f}_k = \sum_{j=1}^M (c_j \cdot f_{k-j} + c_{-j} \cdot f_{k+j}) = \sum_{j=-M; j \neq 0}^M c_j \cdot f_{k-j}$$

V nasledujúcich úvahách budeme uvažovať $c_0 = 0$.

$\mathbf{f} =$

$$(f_0, \dots, f_{M-1}, f_M, \dots, f_k, \dots, f_{N-1-M}, f_{N-1-M+1}, \dots, f_{N-1})$$

$\tilde{\mathbf{f}} =$

$$\underbrace{(0, \dots, 0)}_M, \sum_{j=-M}^M c_j \cdot f_{M-j}, \dots, \sum_{j=-M}^M c_j \cdot f_{k-j}, \dots, \sum_{j=-M}^M c_j \cdot f_{N-1-M-j}, \underbrace{(0, \dots, 0)}_M$$

Príslušná báza bude mať tvar

$$\text{miesto:} \quad 0 \quad \dots \quad M-1 \quad M, \quad \dots \quad N-1-M \quad N-M \quad N-1$$

$$\mathbf{b}_1 = (0, \dots, 0, f_{M-1}, \dots, f_{N-M-2}, 0, \dots, 0)$$

$$\mathbf{b}_{-1} = (0, \dots, 0, f_{M+1}, \dots, f_{N-M}, 0, \dots, 0)$$

\vdots

$$\mathbf{b}_n = (0, \dots, 0, f_{M-n}, \dots, f_{N-1-M-n}, 0, \dots, 0)$$

\vdots

$$\mathbf{b}_M = (0, \dots, 0, f_0, \dots, f_{N-1-2M}, 0, \dots, 0)$$

$$\mathbf{b}_{-M} = (0, \dots, 0, f_{2M}, \dots, f_{N-1}, 0, \dots, 0)$$

Vzdialenosť medzi \mathbf{f} a $\tilde{\mathbf{f}}$ je minimálna ak platí

$$\langle \mathbf{f} - \tilde{\mathbf{f}}, \mathbf{b}_n \rangle = 0, \quad \forall n = -M, \dots, M$$

Po dosadení vektorov \mathbf{f} , $\tilde{\mathbf{f}}$ a \mathbf{b}_n dostaneme

$$\left(\underbrace{(f_0, \dots, f_{M-1}, f_M - \sum_{j=-M}^M c_j \cdot f_{M-j}, \dots, f_{N-1-M} - \sum_{j=-M}^M c_j \cdot f_{N-1-M-j}, f_{N-M}, f_{N-M+1}, \dots, f_{N-1})}_M, \underbrace{(0, \dots, 0, f_{M-n}, \dots, f_{N-1-M-n}, 0, \dots, 0)}_M \right) = 0$$

$$\sum_{k=M}^{N-1-M} \left(f_k - \sum_{j=-M}^M c_j \cdot f_{k-j} \right) f_{k-n} = 0$$

Po roznásobení zátvorky a zámene súm dostávame vzorec na výpočet koeficientov pre úlohu interpolácie.

$$\sum_{j=-M}^M c_j \cdot \sum_{k=M}^{N-1-M} f_{k-j} \cdot f_{k-n} = \sum_{k=M}^{N-1-M} f_k f_{k-n} \quad (4.2.8)$$

kde $n = -M, \dots, M$, $n \neq 0$

Koeficienty vypočítané riešením sústavy (4.2.8) sa využijú vo vzorci (4.2.6) na interpoláciu, alebo filtráciu procesu \mathbf{f} .

Úloha 4.2.1 *Navrhnite postup, ako použiť metódu klzavých súčtov v prípade, že intervaly, v ktorých sú hodnoty procesu namerané, nie sú ekvidištanné.*

5 Periodické vlastnosti procesov

5.1 Periodické modely

V prípade, že sa v modeli pre zovšeobecnenú lineárnu regresiu (3.11.1) použijú ako funkcie $\varphi_n(t_k)$ niektoré z funkcií sínus, alebo kosínus, budú analyzované periodické vlastnosti tohto procesu.

$$\varphi(t_k) = c_0\varphi_0(t_k) + c_1\varphi_1(t_k) + \cdots + c_n\varphi_{N-1}(t_k)$$

Príkladom modelu lineárnej regresie, ktorou je možné aproximovať proces s približne periodickým priebehom, je funkcia:

$$\varphi(t_k) = c_0 \cdot \sin(\omega_0 t_k) + c_1 \cdot \cos(\omega_1 t_k)$$

Hodnoty ω_0 a ω_1 zvolíme vopred podľa periód, ktoré sa v procese vyskytujú. Pomocou postupov z predchádzajúcej kapitoly potom vypočítame koeficienty c_0 a c_1 , ktoré určujú najlepšiu aproximáciu procesu pomocou krivky takéhoto tvaru.

Príklad 5.1.1 *Aproximujte proces $\mathbf{f} = (12, -1, -6, -2, -10, -13, 1, 5, 2, 10)$ regresnou krivkou tvaru $\varphi(t_k) = c_0 \cdot \sin(\omega_0 t_k) + c_1 \cdot \cos(\omega_1 t_k)$.*

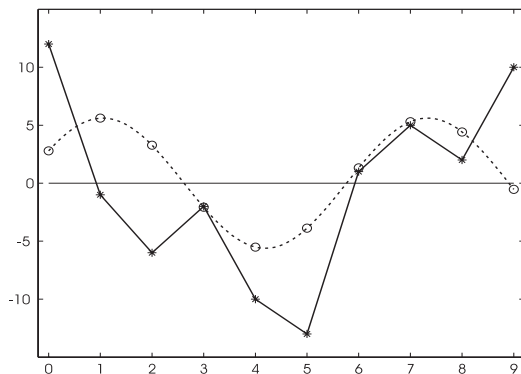
Riešenie:

Metódou lineárnej regresie vypočítame koeficienty c_0 a c_1 , ktoré sú najlepšie pri aproximácii funkciou $\varphi(t_k)$. Hodnoty ω_0 a ω_1 treba zvoliť vopred, pretože sú vo vnútri funkcií sínus a kosínus a nie je možné na ich výpočet použiť lineárne metódy, nie je možné ich exaktne vypočítať. Na obrázkoch 5.1, 5.2 a 5.3 sú vykreslené aproximácie pri intuitívne odhadnutých hodnotách ω_0 a ω_1 . Veľkosti chýb odhadu sú postupne $\|\mathbf{e}_1\| = 20.85$, $\|\mathbf{e}_2\| = 14.89$ a $\|\mathbf{e}_3\| = 10.7$.

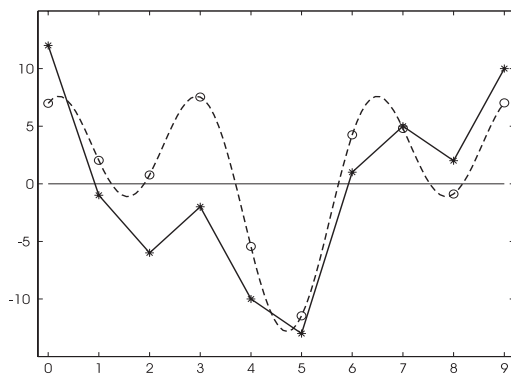
Z obrázkov aj z výpočtu veľkosti chyby vidíme, že koeficienty ω_0 a ω_1 podstatne ovplyvňujú tvar funkcie $\varphi(t_k)$, a teda aj úspešnosť, s akou je proces \mathbf{f} aproximovaný.

Lepšiu aproximáciu by sme mohli dostať, keby sme okrem parametrov ω_0 a ω_1 , ktoré ovplyvňujú dĺžku periód, zaviedli ďalšie dva parametre α_0 a α_1 , ktoré budú ovplyvňovať miesto, kde funkcie sínus a kosínus prechádzajú nulou. Týmito parametrami posunieme obe funkcie po časovej osi. Na obrázku 5.4 je vykreslená aproximácia pri intuitívne odhadnutých hodnotách ω_0 , ω_1 , α_0 a α_1 . Veľkosť chyby odhadu je v tomto prípade $\|\mathbf{e}_4\| = 7.3$.

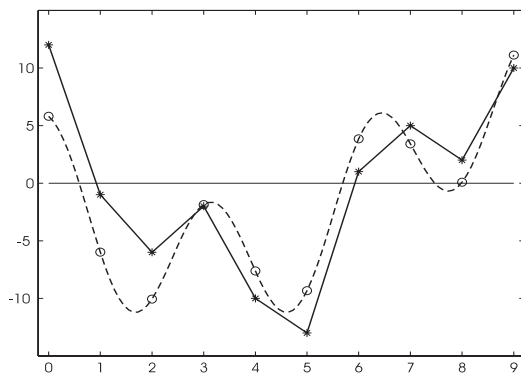
Dá sa dosiahnuť aj omnoho lepšia aproximácia, s veľkosťou chyby $\|\mathbf{e}_5\| = 1$, ale tento výsledok uvedieme až po vybudovaní potrebného matematického aparátu v príklade 5.4.3. \square



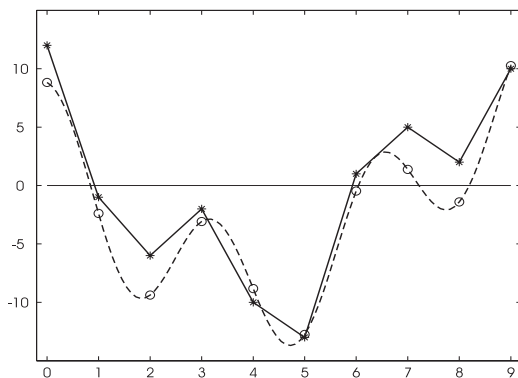
Obr. 5.1. Aproximácia procesu f krivkou $\varphi(t_k) = c_0 \cdot \sin(1 t_k) + c_1 \cdot \cos(1 t_k)$



Obr. 5.2. Aproximácia procesu f krivkou $\varphi(t_k) = c_0 \cdot \sin(1 t_k) + c_1 \cdot \cos(2 t_k)$



Obr. 5.3. Aproximácia procesu f krivkou $\varphi(t_k) = c_0 \cdot \sin(0.5 t_k) + c_1 \cdot \cos(2 t_k)$



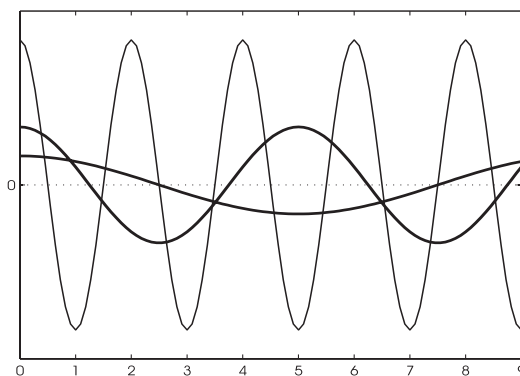
Obr. 5.4: Aproximácia procesu \mathbf{f} krivkou $\varphi(t_k) = c_0 \cdot \sin(0.5 t_k - 0.39) + c_1 \cdot \cos(2 t_k - 0.19)$

Úloha 5.1.1 *S pomocou výpočtovej techniky nájdite také hodnoty parametrov ω_0 , ω_1 , α_0 a α_1 pre funkciu*

$$\varphi(t_k) = c_0 \cdot \sin(\omega_0 t_k + \alpha_0) + c_1 \cdot \cos(\omega_1 t_k + \alpha_1)$$

aby veľkosť chyby odhadu $\|\mathbf{e}\| < 7$.

Problém, ktorý bol riešený v príklade 5.1.1, však môžeme formulovať aj iným spôsobom. Metódou lineárnej regresie vypočítame koeficienty procesu v báze, ktorá pozostáva z veľkého počtu kosínusových priebehov, pričom každý z týchto priebehov má inú periódu. Na obrázku 5.5 sú zobrazené tri takéto funkcie s rôznou periódou aj amplitúdou.



Obr. 5.5. Funkcie kosínus s rôznymi periódami a rôznymi amplitúdami

Podľa toho, aké veľké budú koeficienty pri jednotlivých funkciách sa dá určiť, ktoré sú viac a ktoré menej významné periódy procesu. Báza, ktorá sa na vydeľovanie

periodických trendov z procesu používa, sa nazýva **harmonická báza** a takáto zmena bázy sa nazýva aj **diskrétne Fourierova transformácia**.

Riešime úlohu aproximovať proces ako lineárnu kombináciu funkcií kosínus. Amplitúdy a fázy (= max. výška a posun) týchto funkcií sa vypočítajú ako koeficienty rozkladu procesu do harmonickej bázy.

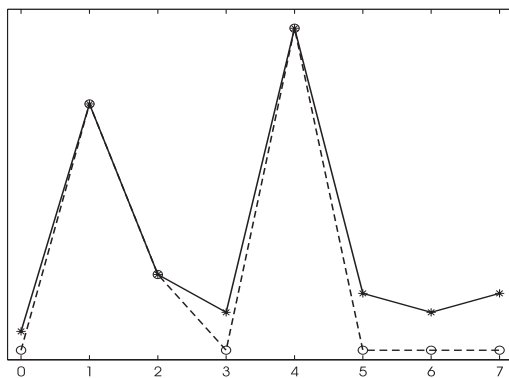
Odhad \mathbf{f} procesu \mathbf{f} vyjadríme v takom podpriestore \mathcal{S} pôvodného priestoru, ktorý bude generovaný báзовými vektormi s najväčšími koeficientmi v rozklade procesu \mathbf{f} . Rozmer podpriestoru \mathcal{S} určíme podľa požiadaviek v zadaní problému. Na ilustráciu tejto myšlienky uvidíme príklad o výbere troch vektorov z jednotkovej bázy vektorového priestoru.

Príklad 5.1.2 Rozložme vektor $\mathbf{f} = (1, 13, 4, 2, 17, 3, 2, 3)$ do jednotkovej bázy. Vyberme z bázy tri báзовé vektory, tak, že v priemet do podpriestoru určeného týmito vektormi bude najlepšie vyjadrovať pôvodný vektor.

Vidíme, že takými vektormi budú vektory

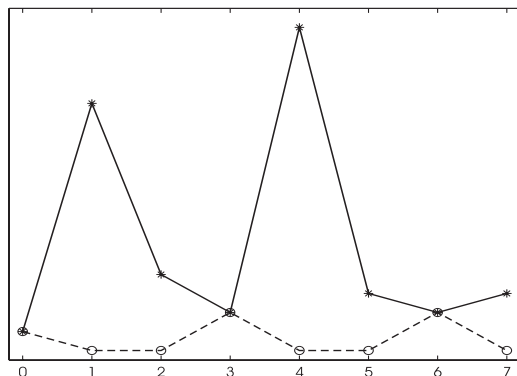
$$\mathbf{e}_4 = (0, 0, 0, 0, 1, 0, 0, 0), \mathbf{e}_1 = (0, 1, 0, 0, 0, 0, 0, 0), \mathbf{e}_2 = (0, 0, 1, 0, 0, 0, 0, 0)$$

Priemet vektora \mathbf{f} do podpriestoru určeného báзовými vektormi \mathbf{e}_4 , \mathbf{e}_1 a \mathbf{e}_2 je vektor $\tilde{\mathbf{f}} = (0, 13, 4, 0, 17, 0, 0, 0)$, ktorý je najlepšou aproximáciou vektora \mathbf{f} spomedzi všetkých trojrozmerných podpriestorov určených trojprvkovou bázou $\mathcal{B}_3 \subset \mathcal{E}$. Teda najlepšie bude zobrať do trojrozmernej bázy vektory s najväčšími koeficientmi $c_1 = 13$, $c_2 = 4$ a $c_4 = 17$. Na obrázku 5.6 je použitá báza $\mathcal{B}_3 = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_4\}$, teda báza s najväčšími koeficientmi. Na obrázku 5.7 je použitá báza $\mathcal{B}_3 = \{\mathbf{e}_0, \mathbf{e}_3, \mathbf{e}_6\}$, teda báza s najmenšími koeficientmi. \square



Obr. 5.6: Aproximácia (dobrá) procesu \mathbf{f} v podpriestore s bázou z vektorov s najväčšími koeficientmi

Úloha 5.1.2 V príklade 5.1.2 je ukázané, že najbližší vektor \mathbf{k} vektoru \mathbf{f} (v podpriestore generovanom trojrozmernou bázou, ktorá je podmnožinou bázy \mathcal{B} celého priestoru), je vektor $\tilde{\mathbf{f}}$, ktorého koeficienty sú tri najväčšie z koeficientov v báze \mathcal{B} . Do



Obr. 5.7: Aproximácia (zlá) procesu \mathbf{f} v podpriestore s bázou z vektorov s najmenšími koeficientmi

bázy \mathcal{B}_3 sú vybrané bázové vektory zodpovedajúce týmto trom najväčším koeficientom. Nájdite takú bázu štvorrozmerného priestoru, v ktorom týmto postupom nenájdeme najlepší trojrozmerný podpriestor. Teda vektor \mathbf{f} s najväčšími koeficientmi nebude najlepšou aproximáciou spomedzi všetkých priemetov do všetkých podpriestorov generovaných bázami $\mathcal{B}_3 \subset \mathcal{B}$.

Úloha 5.1.3 Akú podmienku musí spĺňať báza, aby bolo možné použiť postup z príkladu 5.1.2?

Rovnaký princíp ako v príklade 5.1.2 použijeme aj pri výbere významných koeficientov pri rozklade do harmonickej bázy. Čím väčšia bude absolútna hodnota koeficientu rozkladu vektora \mathbf{f} do bázy, tým lepšie bude daný bázový vektor (vynásobený príslušným koeficientom) aproximovať vektor \mathbf{f} . Že je to naozaj pravda dokážeme tak, že v odseku 5.4 ukážeme, že veľkosť chyby odhadu $\tilde{\mathbf{f}}$ je pri tomto postupe najmenšia.

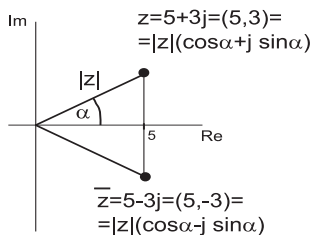
5.1.1 Skalárny súin v komplexnom obore

Aj v tejto kapitole budeme analyzovať procesy, ktorých hodnoty sú reálne čísla. Na ich spracovanie však budeme potrebovať komplexné čísla a komplexné vektory. Preto teraz vysvetlíme matematický aparát potrebný na spracovanie vektorov s komplexnými hodnotami.

Komplexné číslo je číslo, ktoré má okrem reálnej zložky navyše zložku imaginárnu, teda je popísané dvomi hodnotami. Geometricky môžeme čísla znázorniť na priamke (napr. prirodzené, celé, reálne), na znázornenie komplexných čísel je potrebný väčší rozmer, teda rovina. Na obrázku 5.8 je znázornená Gaussova rovina, na ktorej je vyznačený bod zodpovedajúci komplexnému číslu

$$z = 5 + j 3 = 5.83 (\cos(0.54) + j \sin(0.54)) = 5.83 e^{0.54 j}$$

Číslo j vo vyjadrení komplexného čísla sa nazýva komplexná jednotka. Jeho najvýznamnejšia vlastnosť je, že na rozdiel od reálnych čísel, je jeho druhá mocnina číslo

Obr. 5.8. Komplexné číslo $z = 5 + j3$

záporné.

$$j^2 = -1$$

Každé komplexné číslo z môžeme zapísať v algebraickom, geometrickom, goniometrickom alebo exponenciálnom tvare:

- **Algebraický tvar** je súčet

$$z = a + j b$$

kde a je reálna zložka a b je imaginárna zložka komplexného čísla z .

$$\operatorname{Re}\{z\} = a, \quad \operatorname{Im}\{z\} = b$$

- **Geometrický tvar** je usporiadaná dvojica (a, b) , ktorá udáva súradnice komplexného čísla v Gaussovej rovine, pričom prvá zložka udáva hodnotu reálnej časti (os x) a druhá zložka udáva hodnotu imaginárnej časti (os y)

$$z = (a, b)$$

- **Goniometrický tvar** je vyjadrenie komplexného čísla pomocou jeho veľkosti $|z|$ a uhla α . V Gaussovej rovine je hodnota $|z|$ rovná vzdialenosti bodu $z = (a, b)$ od bodu $o = (0, 0)$, a uhol α je uhol, ktorý zvierá úsečka \overline{oz} s kladnou reálnou polosou (kladnou polosou x).

$$z = |z| \cdot (\cos(\alpha) + j \sin(\alpha)); \quad \text{kde } |z| = \sqrt{a^2 + b^2},$$

$$\tan(\alpha) = \frac{b}{a} \Rightarrow \alpha = \arctan\left(\frac{b}{a}\right)$$

$$\cos \alpha = \frac{a}{|z|}, \quad \sin \alpha = \frac{b}{|z|}, \quad \alpha \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

- **Exponenciálny tvar** používa rovnako ako goniometrický tvar veľkosť $|z|$ a uhol α .

$$z = |z| \cdot e^{j\alpha} \quad \text{kde } |z| = \sqrt{a^2 + b^2}, \quad \alpha = \arctan\left(\frac{b}{a}\right)$$

Operáciu sčítania je najjednoduchšie vykonať (s najnižšími nárokmi na pamäť počítača) vtedy, keď sú komplexné čísla v algebraickom tvare.

$$z_1 + z_2 = (a_1 + j b_1) + (a_2 + j b_2) = (a_1 + a_2) + j (b_1 + b_2)$$

Operáciu násobenia je najjednoduchšie vykonať, keď sú komplexné čísla v goniometrickom, alebo exponenciálnom tvare.

$$\begin{aligned} z_1 \cdot z_2 &= |z_1| \cdot (\cos(\alpha_1) + j \sin(\alpha_1)) \cdot |z_2| \cdot (\cos(\alpha_2) + j \sin(\alpha_2)) = \\ &= |z_1||z_2| \cdot (\cos(\alpha_1)\cos(\alpha_2) + j \sin(\alpha_1)\cos(\alpha_2) + j \cos(\alpha_1)\sin(\alpha_2) + \\ &+ j \sin(\alpha_1)j \sin(\alpha_2)) = |z_1||z_2| \cdot (\cos(\alpha_1)\cos(\alpha_2) - \sin(\alpha_1)\sin(\alpha_2)) + \\ &+ j (\sin(\alpha_1)\cos(\alpha_2) + \cos(\alpha_1)\sin(\alpha_2)) = |z_1||z_2| \cdot (\cos(\alpha_1 + \alpha_2) + j \sin(\alpha_1 + \alpha_2)) \end{aligned}$$

respektíve

$$z_1 \cdot z_2 = |z_1| \cdot e^{j \alpha_1} \cdot |z_2| \cdot e^{j \alpha_2} = |z_1| \cdot |z_2| e^{j (\alpha_1 + \alpha_2)}$$

Komplexne združené číslo k číslu $z = a + j b = |z| \cdot e^{j \alpha}$ nazveme číslo

$$\bar{z} = a - j b = |z| \cdot e^{-j \alpha}$$

Definujeme skalárny súčin vektorov, ktorých hodnoty môžu byť aj komplexné čísla. Urobíme to tak, aby zúženie tejto definície na reálne vektory bolo skalárnym súčinom podľa definície v odseku 2.5.

Skalárny súčin je zobrazenie: $\langle \cdot, \cdot \rangle : V \times V \rightarrow F$, pre ktoré:

$$\begin{aligned} \langle \mathbf{f}, \mathbf{f} \rangle &\geq 0, \text{ navyše } \langle \mathbf{f}, \mathbf{f} \rangle = 0 \Leftrightarrow \mathbf{f} = 0 && \text{pozitívnosť} \\ \langle \mathbf{f}, \mathbf{g} \rangle &= \overline{\langle \mathbf{g}, \mathbf{f} \rangle} && \text{symetria} \\ \langle \mathbf{g} + \mathbf{h}, \mathbf{f} \rangle &= \langle \mathbf{g}, \mathbf{f} \rangle + \langle \mathbf{h}, \mathbf{f} \rangle \\ \langle k\mathbf{f}, \mathbf{g} \rangle &= k \cdot \langle \mathbf{f}, \mathbf{g} \rangle && \text{bilinearita} \end{aligned}$$

Predchádzajúca definícia dáva viacero možností ako definovať skalárny súčin v konkrétnom vektorovom priestore. My budeme pracovať s vektorovým priestorom, ktorého prvkami budú usporiadané N -tice komplexných čísel a skalárny súčin pre ľubovoľné dva komplexné vektory \mathbf{f}, \mathbf{g} bude:

$$\mathbf{f} = (f_0, f_1, \dots, f_{N-1}), \quad \mathbf{g} = (g_0, g_1, \dots, g_{N-1}) \quad \text{kde } f_i, g_i \in C$$

Potom skalárny súčin \mathbf{f} a \mathbf{g} je

$$\langle \mathbf{f}, \mathbf{g} \rangle = \sum_{k=0}^{N-1} f_k \overline{g_k}, \quad \text{kde } \overline{g_k} \text{ je číslo komplexne združené ku } g_k$$

Úloha 5.1.4 Dokážte, že takto definovaný súčin, je naozaj skalárny súčin. Teda sú splnené vlastnosti uvedené v definícii skalárneho súčinu.

Úloha 5.1.5 Vynásobte dve komplexné čísla v algebraickom tvare.

Úloha 5.1.6 Sčítajte dve komplexné čísla v goniometrickom, resp. exponenciálnom tvare.

Úloha 5.1.7 Zvoľte si dve komplexné čísla v a w . Do Gaussovej roviny zakreslite súčet aj súčin týchto komplexných čísel.

5.2 Harmonická báza

Na analýzu periodických vlastností v procese budeme používať harmonickú bázu, ktorá bude reprezentovať postupne periódy dĺžky N , $N/2$, $N/3$ až po najkratšiu periódu dĺžky $N/(N-1)$. Všeobecný tvar n -tého vektora **harmonickej bázy** \mathcal{H}_N v N -rozmernom vektorovom priestore je

$$\mathbf{h}_n = \left(\dots, e^{j \frac{2\pi}{N} n \cdot (k-1)}, e^{j \frac{2\pi}{N} n \cdot k}, \dots \right) \quad \text{kde } n = 0, 1, \dots, N-1 \quad (5.2.1)$$

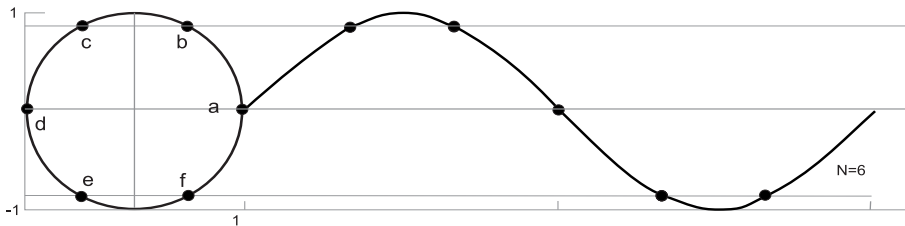
Jednotlivé vektory N -prvkovej harmonickej bázy majú tvar:

$$\begin{aligned} \mathbf{h}_0 &= (1, 1, \dots, 1) \\ \mathbf{h}_1 &= \left(1, e^{j \frac{2\pi}{N}}, e^{j \frac{2\pi}{N} 2}, \dots, e^{j \frac{2\pi}{N} k}, \dots, e^{j \frac{2\pi}{N} (N-1)} \right) \\ \mathbf{h}_2 &= \left(1, e^{j \frac{2\pi}{N} 2}, e^{j \frac{2\pi}{N} 2 \cdot 2}, \dots, e^{j \frac{2\pi}{N} 2 \cdot k}, \dots, e^{j \frac{2\pi}{N} 2 \cdot (N-1)} \right) \\ &\vdots \\ \mathbf{h}_n &= \left(1, e^{j \frac{2\pi}{N} n}, e^{j \frac{2\pi}{N} n \cdot 2}, \dots, e^{j \frac{2\pi}{N} n \cdot k}, \dots, e^{j \frac{2\pi}{N} n \cdot (N-1)} \right) \\ &\vdots \\ \mathbf{h}_{N-1} &= \left(1, e^{j \frac{2\pi}{N} (N-1)}, e^{j \frac{2\pi}{N} (N-1) \cdot 2}, \dots, e^{j \frac{2\pi}{N} (N-1) \cdot k}, \dots, e^{j \frac{2\pi}{N} (N-1) \cdot (N-1)} \right) \end{aligned}$$

Príklad 5.2.1 Nakreslite hodnoty komponentov vektorov harmonickej bázy \mathcal{H}_6 do komplexnej roviny a potom nakreslite imaginárnu zložku každého báзовého vektora ako sínusový priebeh.

Riešenie:

Na obrázku 5.9 sú na jednotkovej kružnici nakreslené jednotlivé hodnoty a, b, c, d, e, f , ktoré sú zložkami vektorov harmonickej bázy \mathcal{H}_6 . Ak vykreslíme proces, ktorého hodnoty budú iba imaginárne zložky báзовých vektorov, dostaneme sínusové priebehy. Na obrázku 5.9 je vykreslený sínusový priebeh zodpovedajúci vektoru \mathbf{h}_1 . Pre hodnoty a, b, c, d, e, f platí, že $b = \bar{f}$ a $c = \bar{e}$.



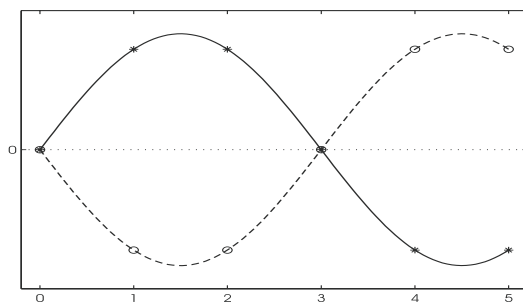
Obr. 5.9. Komplexné čísla a, b, c, d, e, f sú zložky vektora harmonickej bázy \mathcal{H}_6

Navyše uhol α_c zodpovedajúci komplexnému číslu c je dvojnásobný v porovnaní s uhlom α_b , ktorý zodpovedá komplexnému číslu b .

Harmonické bázové vektory 6-rozmerného vektorového priestoru majú teda tvar

$$\begin{aligned}\mathbf{h}_0 &= (1, 1, 1, 1, 1, 1) \\ \mathbf{h}_1 &= (1, b, c, -1, \bar{c}, \bar{b}) \\ \mathbf{h}_2 &= (1, c, \bar{c}, 1, c, \bar{c}) \\ \mathbf{h}_3 &= (1, -1, 1, -1, 1, -1) \\ \mathbf{h}_4 &= (1, \bar{c}, c, 1, \bar{c}, c) \\ \mathbf{h}_5 &= (1, \bar{b}, \bar{c}, -1, c, b)\end{aligned}$$

Vektory \mathbf{h}_1 a \mathbf{h}_5 sú komplexne združené. Rovnako sú komplexne združené vektory \mathbf{h}_2 a \mathbf{h}_4 . Priebiehy, ktoré zodpovedajú komplexne združeným vektorom \mathbf{h}_1 a \mathbf{h}_5 sú na obrázku 5.10. \square



Obr. 5.10: Sínusové priebiehy, zodpovedajúce komplexne združeným bázovým vektorom \mathbf{h}_1 a \mathbf{h}_5 bázy \mathcal{H}_6

Ako vidno na obrázku 5.10 priebiehy, ktoré zodpovedajú komplexne združeným vektorom, môžu byť interpretované ako priebiehy s rôzne dlhými periódami, ale aj ako priebiehy s rovnakou periódou a fázovým posunom o hodnotu π .

Tvrdenie 5.2.1 *Pre vektory harmonickej bázy \mathcal{H}_N platí, že*

$$\mathbf{h}_n = \overline{\mathbf{h}_{N-n}} \quad (5.2.2)$$

Dôkaz:

Vyjadrieme k -tu zložku vektora \mathbf{h}_n :

$$\mathbf{h}_n(k) = e^{j \frac{2\pi}{N} n \cdot k}$$

Vyjadrieme k -tu zložku vektora \mathbf{h}_{N-n} :

$$\mathbf{h}_{N-n}(k) = e^{j \frac{2\pi}{N} (N-n) \cdot k} = e^{j \frac{2\pi}{N} N \cdot k} \cdot e^{j \frac{2\pi}{N} (-n) \cdot k} = \underbrace{e^{j 2\pi \cdot k}}_{=1+j0} \cdot e^{-j \frac{2\pi}{N} n \cdot k} = \overline{\mathbf{h}_n(k)}$$

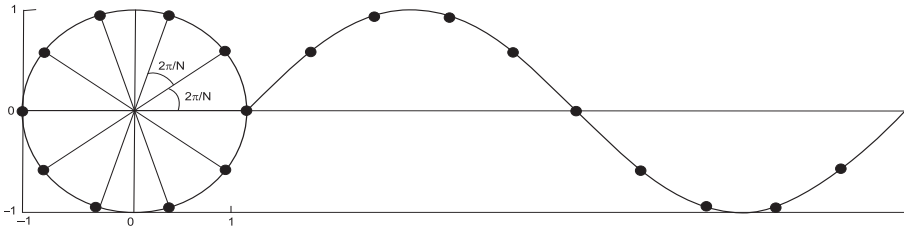
Teda naozaj platí vzťah (5.2.2). \square

Príklad 5.2.2 Aký priebeh má proces, ktorého hodnoty sú imaginárne zložky vektora harmonickej bázy?

Vykreslite hodnoty tohto procesu v časoch $\mathbf{t} = (0, 1, \dots, N - 1)$.

Riešenie:

Harmonická báza \mathcal{H}_N reprezentuje N -zložkové procesy s rôzne veľkými periódami. Harmonická báza N -rozmerného vektorového priestoru má N prvkov. Každý z báзовých vektorov možno interpretovať ako niekoľko obehov po jednotkovej kružnici, respektíve bude reprezentovať niekoľko periód funkcie sínus, ktoré môžeme vykresliť v závislosti od času v intervale $\langle 0, N - 1 \rangle$. Jednotlivé zložky báзовých vektorov sú komplexné čísla, ktoré ležia na jednotkovej kružnici a dve susedné spomedzi týchto komplexných čísel sa líšia o uhol $\frac{2\pi}{N}$ ako vidno na obrázku 5.11, kde je vykreslený sínusový priebeh zodpovedajúci vektoru \mathbf{h}_1 harmonickej bázy \mathcal{H}_N . Vektor \mathbf{h}_0 predstavuje



Obr. 5.11. Zložky vektora \mathbf{h}_1 harmonickej bázy \mathcal{H}_N

konštantnú zložku, respektíve priamku rovnobežnú s osou x .

Vektor \mathbf{h}_1 predstavuje jeden obeh po jednotkovej kružnici, respektíve jednu celú periód funkcie sínus v intervale $\langle 0, N - 1 \rangle$.

Vektor \mathbf{h}_2 predstavuje dva obehy po jednotkovej kružnici, respektíve dve periód funkcie sínus v intervale $\langle 0, N - 1 \rangle$. Pritom „rýchlosť“, akou obehne jednotkovú kružnicu, bude dvojnásobná oproti „rýchlosti“ obehu v prípade vektora \mathbf{h}_1 .

Vektor \mathbf{h}_n predstavuje n obehy po jednotkovej kružnici, respektíve n periód funkcie sínus v intervale $\langle 0, N - 1 \rangle$. \square

Okrem dĺžky periód môže byť periodické chovanie procesu popísané frekvenciou.

Frekvencia je počet obehov (periód), ktoré sa vykonajú za jednotku času, v našom prípade počas celého priebehu procesu. Základnou frekvenciou pre harmonickú bázu je jeden obeh za N časových intervalov, teda

$$\Delta f = \frac{1}{N}$$

Jednotky, v ktorých sa meria frekvencia, môžu byť napr. $[s^{-1}]$, $[rok^{-1}]$.

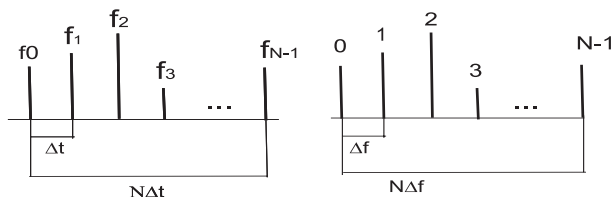
Frekvenciou pre n -tý báзовý vektor \mathbf{h}_n je frekvencia

$$f_n = n \Delta f = \frac{n}{N}$$

a znamená to, že n -tý báзовý vektor je reprezentovaný n periódami počas celého svojho priebehu.

Na obrázku 5.12 je zobrazená súvislosť medzi časovými intervalmi Δt medzi jednotlivými hodnotami procesu a intervalmi Δf medzi frekvenciami procesu.

$$\Delta f = \frac{1}{N\Delta t}$$



Obr. 5.12: Súvislosť medzi časovým priebehom procesu $\mathbf{f} = (f_0, f_1, \dots, f_{N-1})$ a frekvenciami

Vektory harmonickej bázy reprezentujú rôzne dlhé periódy, resp. frekvencie zastúpené v procese. Pre $n = 1$ je to jedna perióda (jeden obeh), pre $n = 2$ sú to 2 periódy (obehy), ale pre $n = N - 1$ je to jedna perióda, pre $n = N - 2$ sú to dve periódy ako to vidno na obrázku 5.10.

Tvrdenie 5.2.2 *Najvyššia frekvencia (resp. najkratšia perióda) zastúpená v harmonickej báze N -rozmerného priestoru je perióda vektora \mathbf{h}_n , kde $n = \frac{N}{2}$ (resp. celá časť $\frac{N}{2}$ ak je N nepárne).*

Tvrdenie 5.2.2 je možné formulovať aj opačne. Ak chceme dobre popísať diskretný proces, treba interval Δt v ktorom meriame hodnoty procesu (vzorkovací interval) zvoliť tak, aby jeho prevrátená hodnota bola aspoň dvojnásobkom maximálnej frekvencie procesu. Teda aby platil vzťah:

$$\frac{1}{\Delta t} \geq 2 \cdot f_{Max}$$

Toto tvrdenie sa nazýva aj Shannon Koteľnikova veta. □

Úloha 5.2.1 *Nájdite na internete, alebo v literatúre presné znenie Shannon Koteľnikovej vety a vysvetlite ho.*

Na začiatku tejto kapitoly sme zaviedli pojem harmonická báza, ale zatiaľ sme nedokázali, že skutočne ide o bázu. Na to, aby množina N vektorov tvorila bázu N -rozmerného vektorového priestoru, musia byť tieto vektory lineárne nezávislé.

Tvrdenie 5.2.3 *Harmonické procesy (vektory) definované vzťahom (5.2.1) tvoria ortogonálnu bázu nejakého N -rozmerného vektorového priestoru.*

Norma každého z vektorov \mathbf{h}_n je rovná N .

Dôkaz:

Najprv dokážeme, že vektory harmonickej bázy \mathbf{h}_n , kde $n = 0, 1, \dots, N - 1$ sú na seba

kolmé a potom, že kolmé vektory sú lineárne nezávislé.

Počítajme skalárny súčin vektorov \mathbf{h}_n a \mathbf{h}_m :

$$\langle \mathbf{h}_n, \mathbf{h}_m \rangle = \begin{cases} 0 & \text{ak je } n \neq m \\ N & \text{ak je } n = m \end{cases}$$

$$n = m$$

$$\begin{aligned} \langle \mathbf{h}_n, \mathbf{h}_m \rangle &= \langle \mathbf{h}_m, \mathbf{h}_m \rangle = \sum_{k=0}^{N-1} b_k \cdot \overline{b_k} = \sum_{k=0}^{N-1} e^{j \frac{2\pi}{N} n \cdot k} \cdot \overline{e^{j \frac{2\pi}{N} n \cdot k}} = \\ &= \sum_{k=0}^{N-1} e^{j \frac{2\pi}{N} n \cdot k} \cdot e^{-j \frac{2\pi}{N} n \cdot k} = \sum_{k=0}^{N-1} 1 = N \end{aligned}$$

$$n \neq m$$

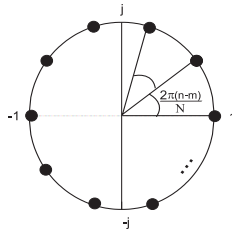
$$\langle \mathbf{h}_n, \mathbf{h}_m \rangle = \sum_{k=0}^{N-1} e^{j \frac{2\pi}{N} n \cdot k} \cdot \overline{e^{j \frac{2\pi}{N} m \cdot k}} = \sum_{k=0}^{N-1} e^{j \frac{2\pi}{N} (n-m) \cdot k}, \quad \text{kde } n - m \neq 0$$

Označme

$$S = \sum_{k=0}^{N-1} e^{j \frac{2\pi}{N} (n-m) \cdot k} \quad \text{pre } n - m \neq 0 \quad (5.2.3)$$

Potrebuje ukázať, že $S = 0$.

Komplexné čísla $e^{j \frac{2\pi}{N} (n-m) \cdot k}$, pre $k = 0, 1, \dots, N-1$ ležia na jednotkovej kružnici ako vidíme na obrázku 5.13.



Obr. 5.13. Komplexné čísla $e^{j \frac{2\pi}{N} (n-m) \cdot k}$, pre $k = 0, 1, \dots, N-1$

Keď vynásobíme rovnicu (5.2.3) číslom $e^{j \frac{2\pi}{N} (n-m)} \neq 0$, dostaneme komplexné čísla, ktorých umiestnenie je na jednotkovej kružnici, len každé číslo je posunuté o uhol $\frac{2\pi}{N} (n-m)$ oproti svojej predchádzajúcej pozícii, preto je na mieste iného čísla (po kružnici vľavo). Umiestnenie čísel je teda rovnaké ako na obrázku 5.13, preto platí

$$S \cdot e^{j \frac{2\pi}{N} (n-m)} = S$$

Pretože sme násobili nenulovým číslom, musí byť $S = 0$, aby posledná rovnosť platila. Teda naozaj pre $n \neq m$ je $\langle \mathbf{h}_n, \mathbf{h}_m \rangle = 0$ a platí, že vektory $\mathbf{h}_n, \mathbf{h}_m$ sú na seba kolmé. Podľa nasledujúcej lemy sú aj lineárne nezávislé.

Lema 5.2.1 Ak sú dva vektory na seba kolmé, potom sú lineárne nezávislé.

Dôkaz:

Predpokladáme

$$\langle \mathbf{f}, \mathbf{g} \rangle = 0$$

hľadáme riešenie rovnice

$$c_1 \cdot \mathbf{f} + c_2 \cdot \mathbf{g} = \mathbf{0}$$

celú rovnicu vynásobíme skalárne vektorom \mathbf{f} . Dostaneme

$$\langle c_1 \cdot \mathbf{f} + c_2 \cdot \mathbf{g}, \mathbf{f} \rangle = \langle \mathbf{0}, \mathbf{f} \rangle$$

po úpravách

$$c_1 \langle \mathbf{f}, \mathbf{f} \rangle + c_2 \langle \mathbf{g}, \mathbf{f} \rangle = 0$$

$$c_1 \langle \mathbf{f}, \mathbf{f} \rangle + 0 = 0$$

posledná rovnosť nastane, len keď $c_1 = 0$. Podobne

$$\langle c_1 \cdot \mathbf{f} + c_2 \cdot \mathbf{g}, \mathbf{g} \rangle = \langle \mathbf{0}, \mathbf{g} \rangle$$

$$c_1 \langle \mathbf{f}, \mathbf{g} \rangle + c_2 \langle \mathbf{g}, \mathbf{g} \rangle = 0$$

$$0 + c_2 \langle \mathbf{g}, \mathbf{g} \rangle = 0$$

a zase rovnosť nastane, len keď $c_2 = 0$.

Teda riešenie rovnice

$$c_1 \cdot \mathbf{f} + c_2 \cdot \mathbf{g} = \mathbf{0}$$

je jediné a to $c_1 = 0$ a $c_2 = 0$, čo je vlastne podmienka nezávislosti, ktorú sme mali dokázať. \square

Ukázali sme, že harmonická báza má niekoľko užitočných vlastností.

- $\mathbf{h}_n \perp \mathbf{h}_m$ pre $n \neq m$
- $\langle \mathbf{h}_m, \mathbf{h}_m \rangle = N$
- $\mathbf{h}_n = \overline{\mathbf{h}_{N-n}}$

5.3 Diskrétna Fourierova transformácia (DFT)

Nech je daný proces

$$\mathbf{f} = (f_0, f_1, \dots, f_{N-1})$$

Tento proces je vyjadrený v jednotkovej báze N -rozmerného vektorového priestoru. Hodnoty procesu v jednotkovej báze \mathcal{E} sú namerané hodnoty procesu.

Vyjadrimo teraz tento proces v tom istom N -rozmernom vektorovom priestore, ktorého báza bude určená harmonickými vektormi tvaru

$$\mathbf{h}_n = \left(\dots, e^{j \frac{2\pi}{N} n \cdot k}, \dots \right) \quad \text{kde } n = 0, 1, \dots, N-1$$

Z priebehu procesu môžeme pomocou rozkladu do harmonickej bázy určiť, ktoré z periód v harmonickej báze sú v danom procese zastúpené výraznejšie, teda ktoré majú v rozklade do harmonickej bázy väčšie koeficienty. Hovoríme, že pomocou harmonickej bázy môžeme analyzovať periodické chovanie sa procesu.

Príklad 5.3.1 *Nájdite na internete príklady konkrétnych procesov, ktoré majú výraznú niektorú periódu.*

Riešenie:

Príkladom je záznam predaja Coca-coly, má periódu približne 1 rok, výskyt slnečných škvŕn, ktorý má periódu cca 13 rokov, ekonomické procesy, ktoré majú periódy rok, mesiac, deň, (sezónne trendy), procesy popisujúce priebeh niektorých samohlások v ľudskej reči majú periódu cca 5 milisekúnd. \square

Harmonická báza je bázou N -rozmerného vektorového priestoru, preto riešime úlohu zmeny bázy (jednotkovej, \mathbf{f} je časový priebeh procesu) v tom istom N -rozmernom priestore. Koeficienty procesu \mathbf{f} v ortogonálnej harmonickej báze vypočítame pomocou vzorca pre výpočet koeficientov v ortogonálnej báze, tak ako sme ich odvodili v kapitole o lineárnej regresii.

$$c_n = \frac{\langle \mathbf{f}, \mathbf{h}_n \rangle}{\langle \mathbf{h}_n, \mathbf{h}_n \rangle} = \frac{1}{N} \sum_{k=0}^{N-1} f_k e^{-j \frac{2\pi}{N} n \cdot k} \quad (5.3.1)$$

Vyjadrenie procesu pomocou týchto koeficientov je

$$\mathbf{f} = \sum_{n=0}^{N-1} c_n \mathbf{h}_n \quad \text{tj.} \quad f_k = \sum_{n=0}^{N-1} c_n \cdot e^{j \frac{2\pi}{N} n \cdot k} \quad (5.3.2)$$

Vzťah (5.3.1) nazývame **spätná diskrétna Fourierova transformácia**.

Vzťah (5.3.2) nazývame **diskrétna Fourierova transformácia** a budeme ju označovať ako DFT.

$$\mathbf{f} = (c_0, c_1, \dots, c_{N-1})_{\mathcal{H}}$$

Vektor koeficientov \mathbf{c} sa nazýva **spektrum** procesu v harmonickej báze, často aj len **spektrum** procesu.

$$\mathbf{c} = (c_0, c_1, \dots, c_{N-1})$$

Vektor absolútnych hodnôt $\mathbf{ca} = (c_0, |c_1|, \dots, |c_{N-1}|)$ spektra procesu nazývame **amplitúdové spektrum** procesu.

Vektor $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_{N-1})$ uhlov α_k spektra procesu, kde $c_k = |c_k| e^{j \alpha_k}$ nazývame **fázové spektrum** procesu.

Poznámka:

Koeficient c_0 má vždy reálnu hodnotu. V amplitúdovom spektre môže ako jediný nadobudnúť aj zápornú hodnotu. Zodpovedajúci uhol fázového spektra α_0 má hodnotu 0. Koeficient c_0 je koeficient bázového vektora $\mathbf{h}_0 = \left(\dots, e^{j \frac{2\pi}{N} 0 \cdot k}, \dots \right) = (1, 1, \dots, 1)$.

Príslušný výpočet je

$$c_0 = \frac{\langle \mathbf{f}, \mathbf{h}_0 \rangle}{\langle \mathbf{h}_0, \mathbf{h}_0 \rangle} = \frac{\sum_{k=1}^{N-1} f_k}{\sum_{k=1}^{N-1} 1} = \frac{1}{N} \sum_{k=1}^{N-1} f_k$$

Posledný vzťah je aritmetický priemer hodnôt procesu \mathbf{f} , preto je koeficient c_0 nazývaný **stredná hodnota procesu**.

Úloha 5.3.1 Ak by sme v amplitúdovom spektre požadovali kladnú hodnotu záporného koeficientu c_0 , zmenila by sa jeho fáza α_0 . Akú hodnotu by mala táto fáza?

Tvrdenie 5.3.1 Pre spektrálne koeficienty procesu podobne ako pre vektory harmonickej bázy platí, že

$$c_n = \overline{c_{N-n}}$$

Dôkaz:

$$\begin{aligned} c_n &= \frac{\langle \mathbf{f}, \mathbf{h}_n \rangle}{\langle \mathbf{h}_n, \mathbf{h}_n \rangle} = \frac{1}{N} \sum_{k=0}^{N-1} f_k e^{-j \frac{2\pi}{N} n k} \\ c_{N-n} &= \frac{\langle \mathbf{f}, \mathbf{h}_{N-n} \rangle}{\langle \mathbf{h}_{N-n}, \mathbf{h}_{N-n} \rangle} = \frac{1}{N} \sum_{k=0}^{N-1} f_k e^{-j \frac{2\pi}{N} (N-n) k} = \\ &= \frac{1}{N} \sum_{k=0}^{N-1} f_k e^{-j \frac{2\pi}{N} N k} e^{-j \frac{2\pi}{N} (-n) k} = \frac{1}{N} \sum_{k=0}^{N-1} f_k e^{j \frac{2\pi}{N} n k} = \overline{\frac{1}{N} \sum_{k=0}^{N-1} f_k e^{-j \frac{2\pi}{N} n k}} \end{aligned}$$

□

Tvrdenie 5.3.2 Pre amplitúdové spektrum procesu platí $|c_n| = |c_{N-n}|$.

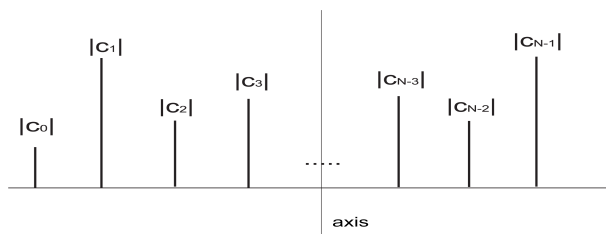
Dôkaz:

Podľa tvrdenia 5.3.1 platí $c_n = \overline{c_{N-n}}$.

$$|c_n| = \sqrt{(Re\{c_n\})^2 + (Im\{c_n\})^2} = \sqrt{(Re\{c_n\})^2 + (-Im\{c_n\})^2} = |\overline{c_n}|$$

Amplitúdové spektrum je symetrické (podľa osi) tak, ako to vidno na obrázku 5.14.

□



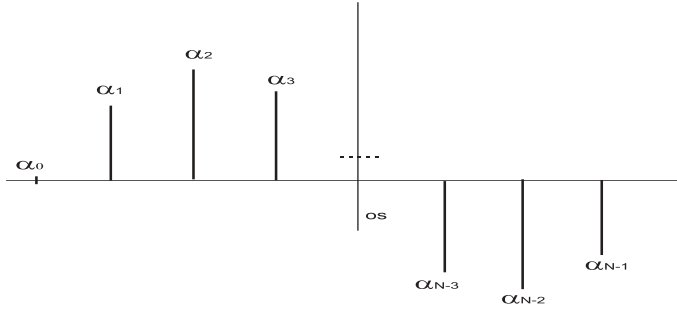
Obr. 5.14. Amplitúdové spektrum je osovo symetrické (okrem hodnoty c_0)

Tvrdenie 5.3.3 Pre fázové spektrum procesu platí $\alpha_n = -\alpha_{N-n}$

Dôkaz:

Podľa tvrdenia 5.3.1 platí $c_n = \overline{c_{N-n}}$. Odtiaľ už priamo vyplýva

$$\alpha_n = -\alpha_{N-n}$$



Obr. 5.15. Fázové spektrum je bodovo symetrické (okrem hodnoty α_0)

Fázové spektrum je symetrické (podľa stredu) tak, ako to vidno na obrázku 5.15. \square Pomocou spektrálnych koeficientov je možné vypočítať veľkosť procesu tak, ako to uvádza nasledujúce tvrdenie.

Tvrdenie 5.3.4 (Veľkosť procesu) *Je daný proces \mathbf{f} s nameranými hodnotami $\mathbf{f} = (f_0, f_1, \dots, f_{N-1})$, nech jeho spektrum v harmonickej báze je $\mathbf{c} = (c_0, c_1, \dots, c_{N-1})$. Potom platí*

$$\|\mathbf{f}\|^2 = N \cdot \sum_{n=0}^{N-1} |c_n|^2$$

Dôkaz:

Dosaďme za \mathbf{f} jeho rozklad do harmonickej bázy.

$$\begin{aligned} \|\mathbf{f}\|^2 = (\mathbf{f}, \mathbf{f}) &= \left(\sum_{n=0}^{N-1} c_n \mathbf{h}_n, \sum_{m=0}^{N-1} c_m \mathbf{h}_m \right) = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} c_n \cdot \overline{c_m} \underbrace{(\mathbf{h}_n, \mathbf{h}_m)}_{=0 \text{ (pre } n \neq m)} = \\ &= \sum_{n=0}^{N-1} |c_n|^2 (\mathbf{h}_n, \mathbf{h}_n) = N \cdot \sum_{n=0}^{N-1} |c_n|^2 \end{aligned}$$

\square

Vlastnosti z tvrdení 5.3.1, 5.3.2 a 5.3.3 znamenajú, že na presné určenie spektra procesu (= koeficientov procesu v harmonickej báze) stačí vypočítať polovicu koeficientov. Ostatné je možné určiť, pretože sú komplexne združené k už vypočítaným koeficientom. V každom z koeficientov sú však uložené dve informácie o báзовom vektore. Jeho amplitúda a posun. V nasledujúcej kapitole uvedieme, ako analyzovať proces pomocou jeho koeficientov v harmonickej báze.

5.4 Priemet do podpriestoru určeného DFT

Úlohy, ktoré môžeme riešiť pomocou harmonickej bázy procesu, sú napríklad:

- odhaľovanie periodických vlastností procesu (periodických trendov)

- filtrácia procesu

Úloha 5.4.1 *Spomedzi všetkých M -rozmerných vektorových podpriestorov, generovaných vektormi harmonickej bázy N -rozmerného vektorového priestoru, nájdite taký, v ktorom bude proces \mathbf{f} najlepšie aproximovaný.*

Označme \mathcal{H}_N harmonickú bázu N -rozmerného vektorového priestoru a \mathcal{H}_M jej M -prvkovú podmnožinu.

$$\mathcal{H}_M \subset \mathcal{H}_N$$

Proces \mathbf{f} môžeme presne vyjadriť v báze \mathcal{H}_N a v báze \mathcal{E} :

$$\mathbf{f} = (f_0, f_1, \dots, f_{N-1})_{\mathcal{E}} = (c_0, c_1, \dots, c_{N-1})_{\mathcal{H}_N}$$

Priemet procesu \mathbf{f} do podpriestoru generovaného bázou \mathcal{H}_M označíme $\tilde{\mathbf{f}}$:

$$\tilde{\mathbf{f}} = (\tilde{c}_0, \tilde{c}_1, \dots, \tilde{c}_{N-1})_{\mathcal{H}_N}$$

pričom M hodnôt spomedzi \tilde{c}_k bude rovných pôvodným koeficientom c_k a $N - M$ koeficientov \tilde{c}_k bude rovných nule. Nule budú rovné tie koeficienty, pre ktoré bude vzdialenosť $d(\mathbf{f}, \tilde{\mathbf{f}})$ minimálna. Chceme, aby $\|\mathbf{f} - \tilde{\mathbf{f}}\| \rightarrow \min$.

To nastane vtedy, keď

$$\|\mathbf{f} - \tilde{\mathbf{f}}\|^2 \rightarrow \min$$

Pre priemet $\tilde{\mathbf{f}}$ platí

$$(\mathbf{f} - \tilde{\mathbf{f}}) \perp \tilde{\mathbf{f}}$$

Podľa Pytagorovej vety

$$\begin{aligned} \|\mathbf{f}\|^2 &= \|\tilde{\mathbf{f}}\|^2 + \|\mathbf{f} - \tilde{\mathbf{f}}\|^2 \\ \|\mathbf{f} - \tilde{\mathbf{f}}\|^2 &= \|\mathbf{f}\|^2 - \|\tilde{\mathbf{f}}\|^2 = (\mathbf{f}, \mathbf{f}) - (\tilde{\mathbf{f}}, \tilde{\mathbf{f}}) = \\ &= N \sum_{n=0}^{N-1} |c_n|^2 - N \sum_{n=0}^{N-1} |\tilde{c}_n|^2 = N \sum_{n=0}^{N-1} (|c_n|^2 - |\tilde{c}_n|^2) \rightarrow \min \end{aligned}$$

Hodnota $N \sum_{n=0}^{N-1} (|c_n|^2 - |\tilde{c}_n|^2)$ bude minimálna, ak do bázy vyberieme tie bázové funkcie \mathbf{h}_n , pre ktoré sú absolútne hodnoty koeficientov $|\tilde{c}_n|$ čo najväčšie.

Hodnota normy rozdielu $\|\mathbf{f} - \tilde{\mathbf{f}}\|^2$ bude potom najmenšia spomedzi všetkých priemetov do podpriestorov pozostávajúcich z rovnakého počtu vektorov harmonickej bázy.

Vektory harmonickej bázy majú komplexné hodnoty, rovnako koeficienty sú komplexné čísla. Vo všeobecnosti môže byť ich lineárna kombinácia vektor s hodnotami, ktoré nie sú reálne čísla. Ukážeme, že môžeme vybrať vektory z harmonickej bázy a im zodpovedajúce koeficienty tak, aby ich lineárna kombinácia bol reálny vektor. Ukázali sme už, že pre vektory a koeficienty harmonickej bázy platí

$$\mathbf{h}_n = \overline{\mathbf{h}_{N-n}} \quad \text{a} \quad c_n = \overline{c_{N-n}}$$

Tvrdenie 5.4.1 *Súčet $c_n \mathbf{h}_n + c_{N-n} \mathbf{h}_{N-n}$ je reálny a navyše platí*

$$c_n \mathbf{h}_n + c_{N-n} \mathbf{h}_{N-n} = 2\operatorname{Im}\{c_n\} \cos\left(\frac{2\pi}{N} n\mathbf{t}\right) - 2\operatorname{Re}\{c_n\} \sin\left(\frac{2\pi}{N} n\mathbf{t}\right) \quad (5.4.1)$$

Dôkaz:

Vypočítajme k -tu zložku vektora $c_n \mathbf{h}_n + c_{N-n} \mathbf{h}_{N-n}$:

$$\begin{aligned}
 (c_n \mathbf{h}_n + c_{N-n} \mathbf{h}_{N-n})_k &= (c_n \mathbf{h}_n + \overline{c_n} \overline{\mathbf{h}_n})_k = \\
 &= (Re\{c_n\} + j Im\{c_n\}) \cdot e^{j \frac{2\pi}{N} n \cdot k} + (Re\{c_n\} - j Im\{c_n\}) \cdot e^{-j \frac{2\pi}{N} n \cdot k} = \\
 &= (Re\{c_n\} + j Im\{c_n\}) \cdot (\cos(\frac{2\pi}{N} n \cdot k) + j \sin(\frac{2\pi}{N} n \cdot k)) + \\
 &+ (Re\{c_n\} - j Im\{c_n\}) \cdot (\cos(-\frac{2\pi}{N} n \cdot k) + j \sin(-\frac{2\pi}{N} n \cdot k)) = \\
 &= 2Re\{c_n\} \cos(\frac{2\pi}{N} n \cdot k) - 2Im\{c_n\} \sin(\frac{2\pi}{N} n \cdot k)
 \end{aligned}$$

Odvodená rovnosť je vlastne k -ta zložka vzťahu (5.4.1). □

Poznámka:

Riešili sme úlohu nájsť priemet procesu do podpriestoru \mathcal{H}_M , ktorý obsahuje vopred daný počet vektorov M . V skutočnosti sa to dá iba v prípade, že M je také číslo (podľa potreby párne, alebo nepárne), že sa do bázy dá spolu s vektorom \mathbf{h}_n zobrať aj vektor \mathbf{h}_{N-n} , ktorý je s ním komplexne združený. Iba vtedy dostaneme ako súčet reálny vektor.

Úloha 5.4.2 V N -rozmernom priestore, kde $N = 8$ chceme urobiť priemet do M -rozmerného podpriestoru generovaného vektormi harmonickej bázy. Určite, aké hodnoty môže mať číslo M .

Príklad 5.4.1 Hodnoty procesu \mathbf{f} v časoch $\mathbf{t} = (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11)$ sú $\mathbf{f} = (1, 2, 4, 2, 5, 6, 3, 5, 7, 4, 6, 8)$. Analyzujte tento proces pomocou Fourierovej transformácie a nájdite dominantnú periódu tohto procesu.

Z procesu odstránime lineárnu časovú závislosť reprezentovanú funkciou

$$\varphi(t_k) = c_0 + c_1 \cdot t_k$$

ako priemet 12-rozmerného vektora \mathbf{f} do dvojrozmerného vektorového priestoru generovaného bázou $\mathcal{B}_2 = \{\mathbf{b}_0, \mathbf{b}_1\}$, kde

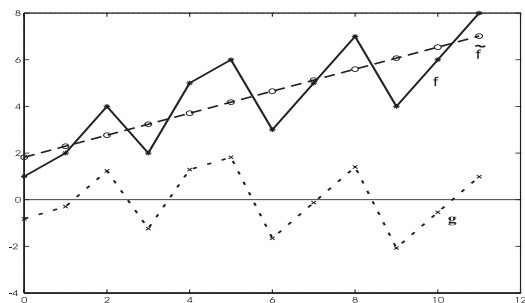
$$\begin{aligned}
 \mathbf{b}_0 &= (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1), \\
 \mathbf{b}_1 &= (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11).
 \end{aligned}$$

Proces, ktorý ostal po odčítaní lineárneho trendu je na obrázku 5.16. Má hodnoty

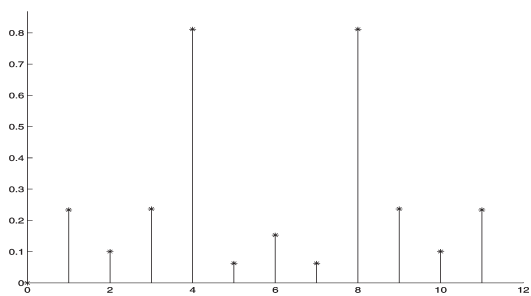
$$\begin{aligned}
 \mathbf{g} &= (-0.82, -0.29, 1.23, -1.23, 1.29, 1.82, \\
 &\quad -1.65, -0.12, 1.40, -2.07, -0.54, 0.99)
 \end{aligned}$$

Spektrum procesu \mathbf{g} v harmonickej báze je

$$\begin{aligned}
 \mathbf{c} &= (0, -0.09 - 0.22j, 0.03 + 0.1j, 0.24 + 0.01j, -0.72 + 0.37j, 0.06 + 0.02j, 0.15, \\
 &\quad 0.06 - 0.02j, -0.72 - 0.37j, 0.24 - 0.01j, 0.03 - 0.1j, -0.09 + 0.22j).
 \end{aligned}$$



Obr. 5.16: Pôvodný proces f , jeho priemet do dvojrozmerného podpriestoru \tilde{f} , proces g je proces f po odstránení lineárneho trendu



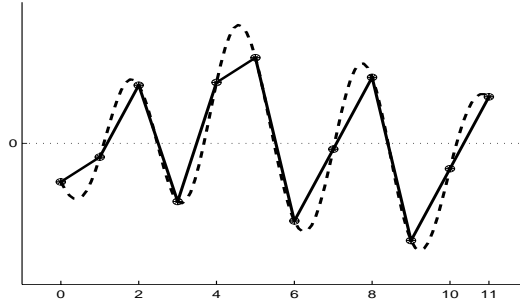
Obr. 5.17. Amplitúdové spektrum procesu g

Amplitúdové spektrum procesu \mathbf{g} je

$$|\mathbf{c}| = (0, 0.23, 0.1, 0.24, 0.81, 0.06, 0.15, 0.06, 0.81, 0.24, 0.1, 0.23).$$

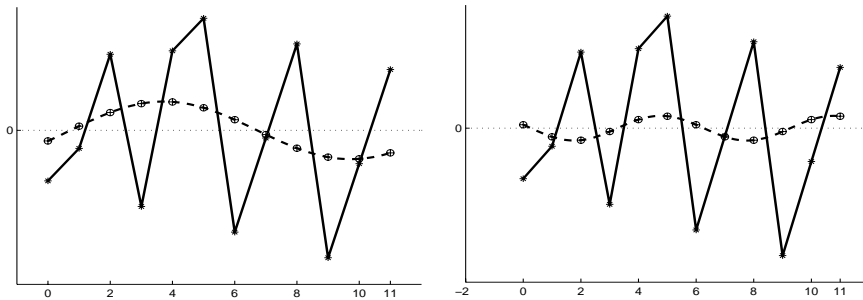
Ak proces \mathbf{g} vyjadríme v harmonickej báze, dostaneme presne hodnoty procesu:

$$\mathbf{g} = c_0 \mathbf{h}_0 + c_1 \mathbf{h}_1 + c_2 \mathbf{h}_2 + c_3 \mathbf{h}_3 + c_4 \mathbf{h}_4 + c_5 \mathbf{h}_5 + c_6 \mathbf{h}_6 + \cdots + c_{10} \mathbf{h}_{10} + c_{11} \mathbf{h}_{11}$$



Obr. 5.18: Hodnoty procesu \mathbf{g} a lineárna kombinácia všetkých vektorov harmonickej bázy, vykreslená ako spojitá funkcia času

Ak proces \mathbf{g} odhadneme pomocou priemetu do dvojrozmerného podpriestoru, dostávame postupne aproximácie zobrazené v obrázkoch 5.19–5.21:

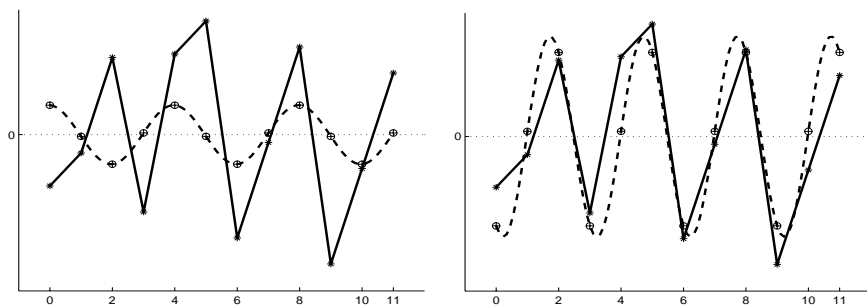


Obr. 5.19: Hodnoty procesu \mathbf{g} a lineárna kombinácia vektorov $\mathbf{h}_1, \mathbf{h}_{11}$ (chyba 4.21); lineárna kombinácia vektorov $\mathbf{h}_2, \mathbf{h}_{10}$ (chyba 4.34)

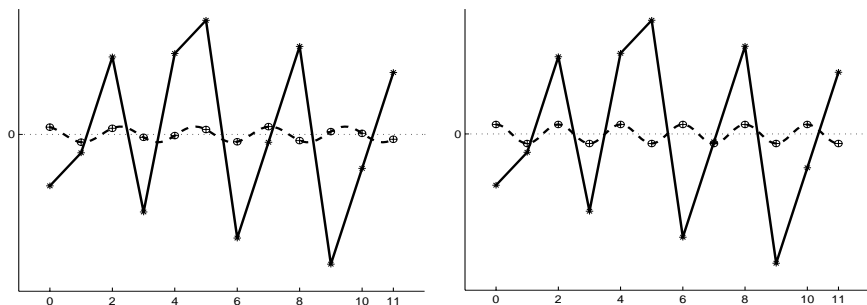
Vidíme, že spomedzi jedno až dvojrozmerných vektorových priestorov je proces \mathbf{g} najlepšie aproximovaný v priestore generovanom bázou $\mathcal{B}_2 = \{\mathbf{h}_4, \mathbf{h}_8\}$. Perióda, ktorú majú báзовé funkcie $\mathbf{h}_4, \mathbf{h}_8$ je teda dominantnou periódou procesu \mathbf{g} .

Dominantná perióda zodpovedá dominantnej frekvencii $\frac{4}{N}$ s uhlom $\omega = \frac{2\pi}{N} 4 = \frac{2\pi}{12} 4 = \frac{2\pi}{3}$. Dĺžka periódy je potom $\frac{2\pi}{\omega}$, čo zodpovedá $\frac{N}{4} = 3$ hodnotám procesu \mathbf{g} .

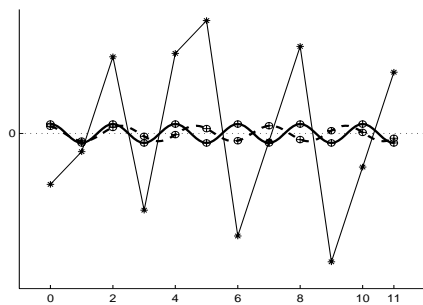
Porovnajme ešte súčet $c_5 \mathbf{h}_5 + c_7 \mathbf{h}_7$ so súčinom $c_6 \mathbf{h}_6$. Na obrázku 5.22 vidíme, že hoci $|c_5| = 0.1527$ je viac ako dvojnásobok hodnoty $|c_6| = 0.0626$, ich príspevky majú takmer rovnakú amplitúdu.



Obr. 5.20: Hodnoty procesu g a lineárna kombinácia vektorov h_3, h_9 (chyba 4.21) a lineárna kombinácia vektorov h_4, h_8 (chyba 1.81)



Obr. 5.21: Hodnoty procesu g a lineárna kombinácia vektorov h_5, h_7 (chyba 4.35); násobok vektora h_6 (chyba 4.33)



Obr. 5.22: Hodnoty procesu g , lineárna kombinácia vektorov b_5, b_7 a násobok vektora b_6

Pri hľadaní najväčšej z hodnôt amplitúdového spektra budeme s ostatnými hodnotami amplitúdového spektra porovnávať polovičnú hodnotu $|c_{\frac{N}{2}}|$, teda budeme hľadať najväčšiu spomedzi hodnôt

$$|c_0|, |c_1|, \dots, |c_{\frac{N}{2}-1}|, \frac{1}{2}|c_{\frac{N}{2}}|, |c_{\frac{N}{2}+1}|, \dots, |c_{N-1}|$$

□

Tvrdenie 5.4.2

Nech má proces $\mathbf{f} = (f_0, f_1, \dots, f_{N-1})$ spektrum $\mathbf{c} = (c_0, c_1, \dots, c_{N-1})$. Potom tento proces môžeme vyjadriť ako lineárnu kombináciu vektorov harmonickej bázy \mathcal{H}_N .

$$\mathbf{f} = \sum_{n=0}^{N-1} c_n \cdot \mathbf{h}_n \quad (5.4.2)$$

Modelom tohto rozkladu je funkcia, ktorá je lineárnou kombináciou funkcií sínus a kosínus. Hodnota $\lfloor \frac{N}{2} \rfloor$ je dolná celá časť podielu $\frac{N}{2}$ a časy $t_k = k$.

$$\varphi(k) = c_0 + \sum_{n=1}^{\lfloor \frac{N}{2} \rfloor} \left(2 \operatorname{Re}\{c_n\} \cos\left(\frac{2\pi n}{N} k\right) - 2 \operatorname{Im}\{c_n\} \sin\left(\frac{2\pi n}{N} k\right) \right) \quad (5.4.3)$$

Funkcia $\varphi(k)$ sa dá vyjadriť aj ako lineárna kombinácia funkcií kosínus, posunutých o fázu určenú fázovým spektrom.

$$\varphi(k) = c_0 + \sum_{n=1}^{\lfloor \frac{N}{2} \rfloor} 2|c_n| \cos\left(\frac{2\pi n}{N} k + \alpha_n\right) \quad (5.4.4)$$

Dôkaz:

Vzťah (5.4.2) je vlastne spätná Fourierova transformácia.

Vzťah (5.4.3) vznikne dosadením rovnosti (5.4.1) pre $f_k = \varphi(k)$ do vzťahu (5.4.2).

Úpravou vzťahu (5.4.3) dostaneme vzťah (5.4.4).

$$\begin{aligned} \varphi(k) &= c_0 + \sum_{n=1}^{\lfloor \frac{N}{2} \rfloor} \left(2 \operatorname{Re}\{c_n\} \cos\left(\frac{2\pi n}{N} k\right) - 2 \operatorname{Im}\{c_n\} \sin\left(\frac{2\pi n}{N} k\right) \right) = \\ &= c_0 + \sum_{n=1}^{\lfloor \frac{N}{2} \rfloor} \left(2 \operatorname{Re}\{c_n\} \frac{|c_n|}{\operatorname{Re}\{c_n\}} \overbrace{\frac{\operatorname{Re}\{c_n\}}{|c_n|}}^{\cos \alpha_n} \cos\left(\frac{2\pi n}{N} k\right) - \right. \\ &\quad \left. - 2 \operatorname{Im}\{c_n\} \frac{|c_n|}{\operatorname{Im}\{c_n\}} \underbrace{\frac{\operatorname{Im}\{c_n\}}{|c_n|}}_{\sin \alpha_n} \sin\left(\frac{2\pi n}{N} k\right) \right) = \end{aligned}$$

$$\begin{aligned}
&= c_0 + \sum_{n=1}^{\lfloor \frac{N}{2} \rfloor} 2|c_n| \left(\cos \alpha_n \cos \left(\frac{2\pi n}{N} k \right) - \sin \alpha_n \sin \left(\frac{2\pi n}{N} k \right) \right) = \\
&= c_0 + \sum_{n=1}^{\lfloor \frac{N}{2} \rfloor} 2|c_n| \cos \left(\frac{2\pi n}{N} k + \alpha_n \right)
\end{aligned}$$

V poslednom kroku sme použili goniometrický súčtový vzorec:

$$\cos \alpha \cdot \cos \beta - \sin \alpha \cdot \sin \beta = \cos(\alpha + \beta)$$

□

Úloha 5.4.3 Navrhните tvar vzorcov pre diskrétnu Fourierovu transformáciu pre proces $\mathbf{f} = (f_0, f_1, \dots, f_{N-1})$, ktorého hodnoty sú namerané v ekvidistančných časových intervaloch

$$\mathbf{t} = (t_0, t_1, \dots, t_{N-1}) = (t_0, t_0 + \Delta t, \dots, t_0 + (N-1)\Delta t)$$

Úloha 5.4.4 Navrhните tvar vzorcov pre diskrétnu Fourierovu transformáciu pre proces $\mathbf{f} = (f_0, f_1, \dots, f_{N-1})$, ktorého hodnoty sú namerané v neekvidistančných časových intervaloch

$$\mathbf{t} = (t_0, t_1, \dots, t_{N-1})$$

Príklad 5.4.2 Proces $\mathbf{f} = (1, -2, 4, 2, -4, 0, 1, -1, 1, 2)$ aproximujte postupne v zväčšujúcich sa podpriestoroch \mathcal{S}_n určených podmnožinami harmonickej bázy 10-rozmerného vektorového priestoru.

Riešenie:

Vypočítajme spektrum procesu \mathbf{f} v harmonickej báze \mathcal{H}_{10} .

$$\begin{aligned}
\mathbf{c} &= (0.4, 0.47 - 0.04j, -0.48 - 0.1j, -0.32 + 1.21j, 0.53 - 0.06j, \\
&0.2, 0.53 + 0.06j, -0.32 - 1.21j, -0.48 + 0.1j, 0.47 + 0.04j)
\end{aligned}$$

Amplitúdové spektrum procesu \mathbf{f} je

$$\mathbf{ca} = (|c_0|, |c_1|, \dots, |c_9|) = (0.4, 0.47, 0.49, 1.25, 0.53, 0.2, 0.53, 1.25, 0.49, 0.47)$$

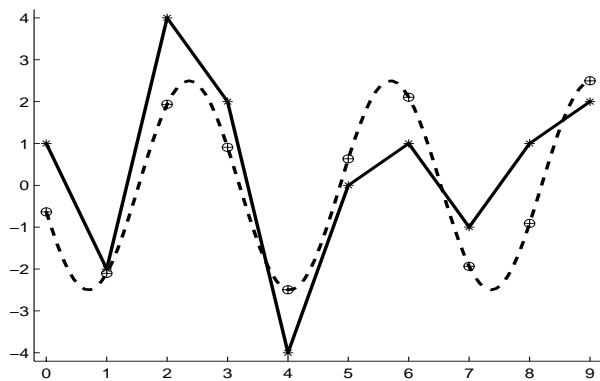
Fázové spektrum procesu \mathbf{f} je

$$\alpha = (\alpha_0, \alpha_1, \dots, \alpha_9) = (0, -0.09, -2.95, 1.83, -0.11, 0, 0.11, -1.83, 2.95, 0.09)$$

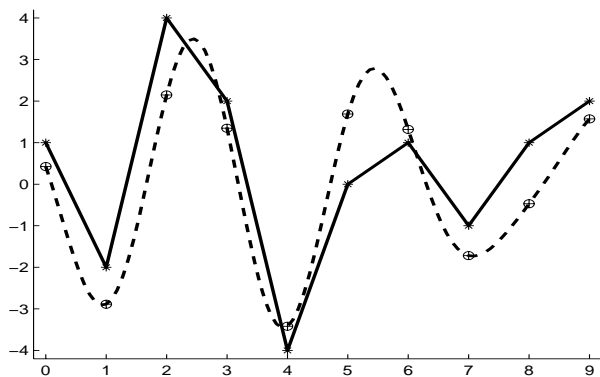
Do bázy vyberieme vektory ktoré zodpovedajú najväčším hodnotám amplitúdového spektra, teda postupne hodnotám $|c_3|$, $|c_4|$, $|c_2|$, $|c_1|$, $|c_0|$ a $|c_5|$.

Na obrázku 5.23 je aproximácia procesu \mathbf{f} procesom $\tilde{\mathbf{f}}_1 \in \mathcal{S}_1$, ktorý je generovaný bázou $\mathcal{B}_1 = \{\mathbf{h}_3, \mathbf{h}_7\}$. Veľkosť chyby odhadu je $\mathbf{e}_1 = \|\mathbf{f} - \tilde{\mathbf{f}}_1\| = 4.1$.

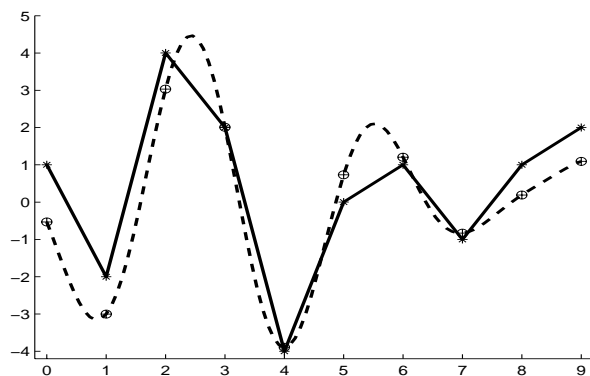
Na obrázku 5.24 je aproximácia procesu \mathbf{f} procesom $\tilde{\mathbf{f}}_2 \in \mathcal{S}_2$, ktorý je generovaný bázou $\mathcal{B}_2 = \{\mathbf{h}_3, \mathbf{h}_7, \mathbf{h}_4, \mathbf{h}_6\}$. Veľkosť chyby odhadu je $\mathbf{e}_2 = \|\mathbf{f} - \tilde{\mathbf{f}}_2\| = 3.34$.



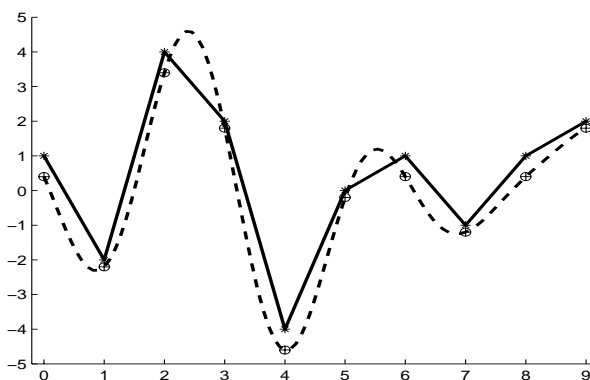
Obr. 5.23. Aproximácia procesu f krivkou $\varphi_1(t_k) = 2.5 \cos(1.89 t_k + 1.83)$



Obr. 5.24: Aproximácia procesu f krivkou $\varphi_2(t_k) = 2.5 \cos(1.89 t_k + 1.83) + 1.06 \cos(2.51 t_k - 0.11)$



Obr. 5.25: Aproximácia procesu \mathbf{f} krivkou $\varphi_3(t_k) = 2.5 \cos(1.89t_k + 1.83) + 1.06 \cos(2.51 \cdot t_k - 0.11) + 0.98 \cos(1.26 t_k - 2.95)$

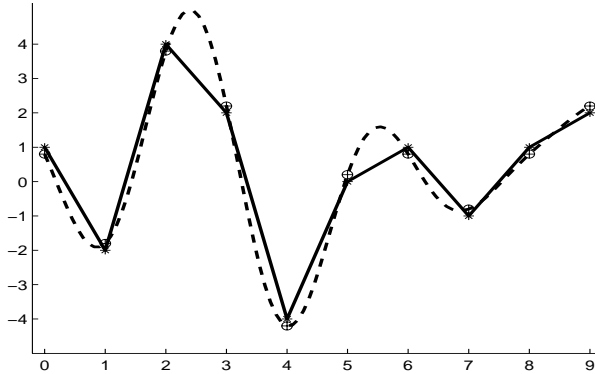


Obr. 5.26: Aproximácia procesu \mathbf{f} krivkou $\varphi_4(t_k) = 2.5 \cos(1.89t_k + 1.83) + 1.06 \cos(2.51 \cdot t_k - 0.11) + 0.98 \cos(1.26 t_k - 2.95) + 0.93 \cos(0.63 t_k - 0.09)$

Na obrázku 5.25 je aproximácia procesu \mathbf{f} procesom $\tilde{\mathbf{f}}_3 \in \mathcal{S}_3$, ktorý je generovaný bázou $\mathcal{B}_3 = \{\mathbf{h}_3, \mathbf{h}_7, \mathbf{h}_4, \mathbf{h}_6, \mathbf{h}_2, \mathbf{h}_8\}$. Veľkosť chyby odhadu je $\mathbf{e}_3 = \|\mathbf{f} - \tilde{\mathbf{f}}_3\| = 2.53$.

Na obrázku 5.26 je aproximácia procesu \mathbf{f} procesom $\tilde{\mathbf{f}}_4 \in \mathcal{S}_4$, ktorý je generovaný bázou $\mathcal{B}_4 = \{\mathbf{h}_3, \mathbf{h}_7, \mathbf{h}_4, \mathbf{h}_6, \mathbf{h}_2, \mathbf{h}_8, \mathbf{h}_1, \mathbf{h}_9\}$. Veľkosť chyby odhadu je $\mathbf{e}_4 = \|\mathbf{f} - \tilde{\mathbf{f}}_4\| = 1.41$.

Na obrázku 5.27 je aproximácia procesu \mathbf{f} procesom $\tilde{\mathbf{f}}_5 \in \mathcal{S}_5$, ktorý je generovaný bázou $\mathcal{B}_5 = \{\mathbf{h}_3, \mathbf{h}_7, \mathbf{h}_4, \mathbf{h}_6, \mathbf{h}_2, \mathbf{h}_8, \mathbf{h}_1, \mathbf{h}_9, \mathbf{h}_0\}$. Veľkosť chyby odhadu je $\mathbf{e}_5 = \|\mathbf{f} - \tilde{\mathbf{f}}_5\| = 0.63$.



Obr. 5.27: Aproximácia procesu \mathbf{f} krivkou $\varphi_5(t_k) = 2.5 \cos(1.89t_k + 1.83) + 1.06 \cos(2.51 \cdot t_k - 0.11) + 0.98 \cos(1.26 t_k - 2.95) + 0.93 \cos(0.63 t_k - 0.09) + 0.4$

Na obrázku 5.28 je aproximácia procesu \mathbf{f} procesom $\tilde{\mathbf{f}}_6 \in \mathcal{S}_6$, ktorý je generovaný bázou $\mathcal{B}_6 = \{\mathbf{h}_3, \mathbf{h}_7, \mathbf{h}_4, \mathbf{h}_6, \mathbf{h}_2, \mathbf{h}_8, \mathbf{h}_1, \mathbf{h}_9, \mathbf{h}_0, \mathbf{h}_5\}$. Veľkosť chyby odhadu je $\mathbf{e}_6 = \|\mathbf{f} - \tilde{\mathbf{f}}_6\| = 0$.

Posledná aproximácia $\tilde{\mathbf{f}}_6 \in \mathcal{S}_6$ je vektorom pôvodného 10-rozmerného priestoru, preto $\tilde{\mathbf{f}}_6 = \mathbf{f}$ a chyba aproximácie je 0. \square

Príklad 5.4.3 Aproximujte proces $\mathbf{f} = (12, -1, -6, -2, -10, -13, 1, 5, 2, 10)$ z príkladu 5.1.1 regresnou krivkou tvaru $\varphi(t_k) = c_0 t_k + \alpha_0 + c_1 \cos(\omega_1 t_k + \alpha_1)$. Pri riešení použite riešenie použite diskretnú Fourierovu transformáciu.

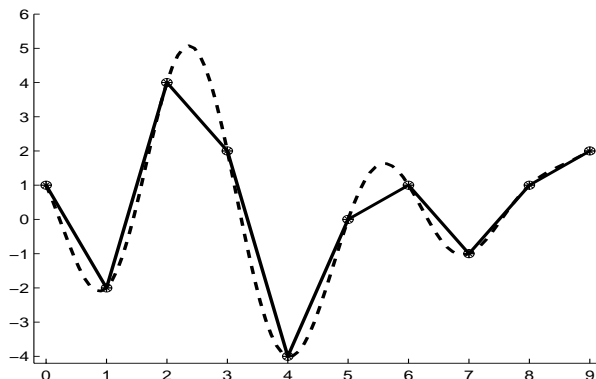
Riešenie:

Vypočítajme spektrum procesu \mathbf{f} v harmonickej báze \mathcal{H}_{10} .

$$\mathbf{c} = (c_0, c_1, \dots, c_9) = (-0.2, 3.74 + 2.72j, -0.02 + 0.06j, 2.51 + 1.21j, \\ -0.13 - 0.1j, 0, -0.13 + 0.1j, 2.51 - 1.21j, -0.02 - 0.06j, 3.74 - 2.72j)$$

Amplitúdové spektrum procesu \mathbf{f} je

$$\mathbf{ca} = (|c_0|, |c_1|, \dots, |c_9|) = (-0.2, 4.62, 0.06, 2.79, 0.16, 0, 0.16, 2.79, 0.06, 4.62)$$



Obr. 5.28: Aproximácia procesu \mathbf{f} krivkou $\varphi_6(t_k) = 2.5 \cos(1.89t_k + 1.83) + 1.06 \cos(2.51 \cdot t_k - 0.11) + 0.98 \cos(1.26 t_k - 2.95) + 0.93 \cos(0.63 t_k - 0.09) + 0.4 + 0.2 \cos(3.14 t_k)$

Fázové spektrum procesu \mathbf{f} je

$$\alpha = (\alpha_0, \alpha_1, \dots, \alpha_9) = (0, 0.63, 1.88, 0.45, -2.51, 0, 2.51, -0.45, -1.88, -0.63)$$

Do bázy vyberieme vektory s najväčšími hodnotami amplitúdového spektra, teda \mathbf{h}_1 , \mathbf{h}_9 , \mathbf{h}_3 a \mathbf{h}_7 .

$$\tilde{\mathbf{f}} = c_1 \cdot \mathbf{h}_1 + c_9 \cdot \mathbf{h}_9 + c_3 \cdot \mathbf{h}_3 + c_7 \cdot \mathbf{h}_7 = (12.5, -1, -5.5, -2, -10, -12.5, 1, 5.5, 2, 10)$$

Chyba odhadu je

$$\|\mathbf{f} - \tilde{\mathbf{f}}\| = 1$$

Zodpovedajúca funkcia času je

$$\begin{aligned} \varphi(t_k) &= 2 \cdot |c_1| \cos\left(\frac{2\pi}{10} t_k + \alpha_1\right) + 2 \cdot |c_3| \cos\left(\frac{2\pi}{10} t_k + \alpha_3\right) = \\ &= 9.24 \cos(0.63 t_k + 0.63) + 5.58 \cos(1.89 t_k + 0.45) \end{aligned}$$

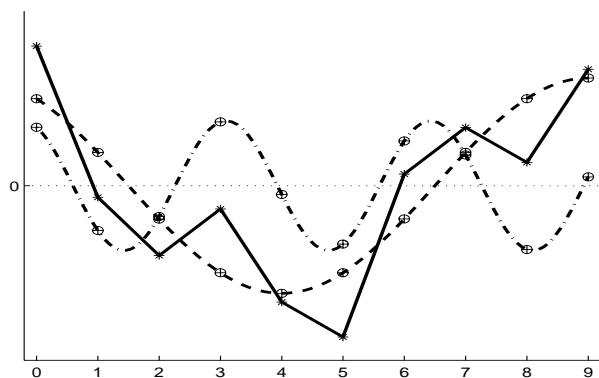
Odhad je súčtom dvoch funkcií kosínus, ktoré sú zobrazené na obrázku 5.29. Odhad je zobrazený na nasledujúcom obrázku 5.30.

Jednu z funkcií kosínus prerobíme na funkciu sínus posunutím o fázu $\frac{\pi}{2}$. Výsledná funkcia má tvar

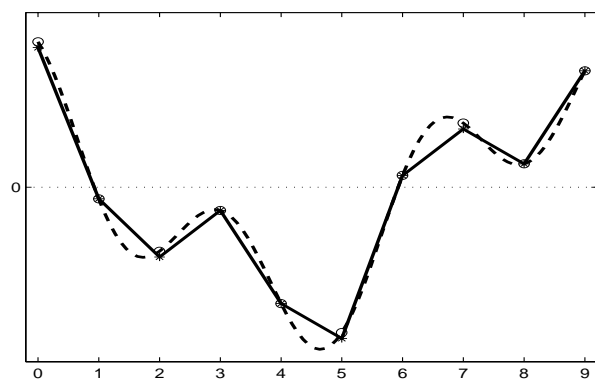
$$\begin{aligned} \varphi(t_k) &= 9.24 \sin\left(0.63 \cdot t_k + 0.63 + \frac{\pi}{2}\right) + 5.58 \cos(1.89 t_k + 0.45) = \\ &= 9.24 \sin(0.63 t_k + 0.22) + 5.58 \cos(1.89 t_k + 0.45) \end{aligned}$$

Chyba odhadu je $\|\mathbf{e}_5\| = 1$.

Úloha 5.4.5 Popíšte čo znamená slovo „diskrétna“ v pojme *diskrétna Fourierova transformácia*. (Okrem *diskrétnej Fourierovej transformácie* je používaná ešte *spojitá Fourierova transformácia*.)



Obr. 5.29: Aproximácia procesu f krivkou $\varphi_1(t_k) = 9.24 \cos(0.63 t_k + 0.22)$ a $\varphi_2(t_k) = 5.58 \cos(1.89 t_k + 0.45)$



Obr. 5.30: Aproximácia procesu f krivkou $\varphi(t_k) = 9.24 \sin(0.63 t_k + 0.22) + 5.58 \cos(1.89 t_k + 0.45)$

5.5 Walsh-Hadamardova báza

Fourierova báza nie je jediná báza, ktorou sa dajú popísať periodické vlastnosti procesu. Inou diskretnou bázou, ktorá popisuje periodické chovanie procesu je napríklad **Walsh-Hadamardova báza**. Táto báza je zložená len z vektorov, ktorých hodnoty sú 1 alebo -1.

Návod na generovanie Walsh-Hadamardovej bázy je nasledujúci. Vytvoríme n -tú Hadamardovu maticu podľa predpisu:

$$H_0 = \quad (1)$$

$$\mathbf{H}_{2^n} = \begin{pmatrix} \mathbf{H}_{2^{n-1}} & \mathbf{H}_{2^{n-1}} \\ \mathbf{H}_{2^{n-1}} & -\mathbf{H}_{2^{n-1}} \end{pmatrix} \quad (5.5.1)$$

Riadky Hadamardovej matice \mathbf{H}_{2^n} sú vektory Walsh-Hadamardovej bázy 2^n -rozmerného priestoru.

Príklad 5.5.1 Vyjadrite Hadamardove matice \mathbf{H}_0 , \mathbf{H}_2 , \mathbf{H}_4 a \mathbf{H}_8 .

Riešenie:

Podľa predpisu (5.5.1) budú Hadamardove matice postupne

$$\mathbf{H}_0 = (1) \quad \mathbf{H}_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \mathbf{H}_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

$$\mathbf{H}_8 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{pmatrix}$$

□

Úloha 5.5.1 *Napíšte desiaty riadok Hadamardovej matice \mathbf{H}_{16} .*

Úloha 5.5.2 *Dokážte, že Walsh-Hadamardova báza 8-rozmerného vektorového priestoru je ortogonálna a každý jej vektor má veľkosť $\sqrt{8}$.*

Vektor koeficientov \mathbf{c} sa nazýva **spektrum** procesu vo Walsh-Hadamardovej báze.

$$\mathbf{c} = (c_0, c_1, \dots, c_{N-1})$$

Vektor absolútnych hodnôt $\mathbf{ca} = (|c_0|, |c_1|, \dots, |c_{N-1}|)$ spektra procesu sa nazýva **amplitúdové spektrum** procesu.

Úloha 5.5.3 *Spomedzi všetkých M -rozmerných podpriestorov, generovaných vektormi Walsh-Hadamardovej bázy, najlepšia aproximácia pôvodného procesu bude v podpriestore, v ktorom sa nachádzajú vektory s najväčšími hodnotami amplitúdového spektra pôvodného procesu. Dokážte toto tvrdenie.*

Príklad 5.5.2 *Napíšte Walsh-Hadamardovu bázu 8-rozmerného vektorového priestoru a vyjadrite v tejto báze koeficienty procesu*

$$\mathbf{f} = (12, -1, -6, -2, -10, -13, 1, 5)$$

Čo najlepšie aproximujte proces \mathbf{f} v trojrozmernom podpriestore 8-rozmerného euklidovského priestoru generovaného vektormi Walsh-Hadamardovej bázy.

Riešenie:

$$\begin{aligned} \mathcal{W}_8 = \{ & (1, 1, 1, 1, 1, 1, 1, 1), (1, -1, 1, -1, 1, -1, 1, -1), \\ & (1, 1, -1, -1, 1, 1, -1, -1), (1, -1, -1, 1, 1, -1, -1, 1), \\ & (1, 1, 1, 1, -1, -1, -1, -1), (1, -1, 1, -1, -1, 1, -1, 1), \\ & (1, 1, -1, -1, -1, 1, 1, 1), (1, -1, -1, 1, -1, 1, 1, -1) \} \end{aligned}$$

Spektrum procesu \mathbf{f} v báze \mathcal{W}_8 je

$$\mathbf{c} = (-1.75, 1, -1.25, 3, 2.5, 1.25, 6, 1.25)$$

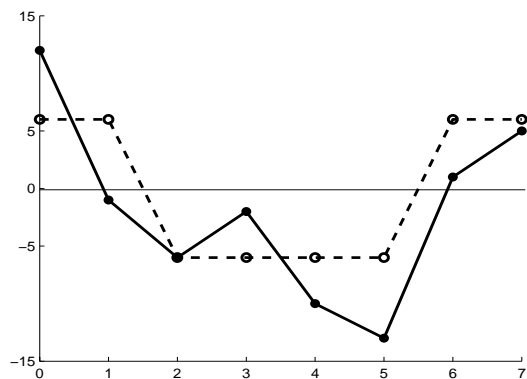
Najlepšou aproximáciou procesu \mathbf{f} v jednorozmernom podpriestore generovanom vektorom Walsh-Hadamardovej bázy je vektor

$$c_6 \mathbf{wh}_6 = 6(1, 1, -1, -1, -1, -1, 1, 1) = (6, 6, -6, -6, -6, -6, 6, 6)$$

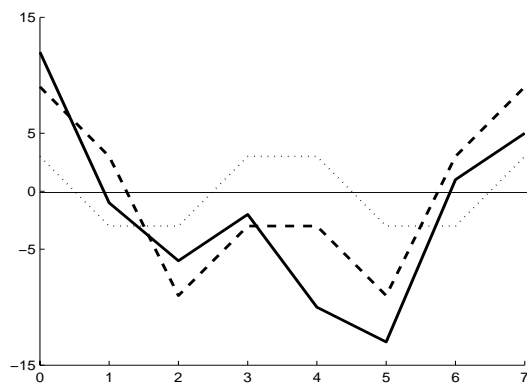
Veľkosť chyby tejto aproximácie je $\|\mathbf{e}_6\| = 13.86$ a pôvodný vektor a jeho aproximácia sú znázornené na obrázku 5.31

Najlepšou aproximáciou procesu \mathbf{f} v dvojrozmernom podpriestore generovanom dvomi vektormi Walsh-Hadamardovej bázy je vektor

$$\begin{aligned} c_3 \mathbf{wh}_3 + c_6 \mathbf{wh}_6 &= 3(1, -1, -1, 1, 1, -1, -1, 1) + 6(1, 1, -1, -1, -1, -1, 1, 1) = \\ &= (9, 3, -9, -3, -3, -9, 3, 9) \end{aligned}$$



Obr. 5.31: Aproximácia procesu \mathbf{f} v jednorozmernom podpriestore, procesom $c_6 \mathbf{wh}_6$



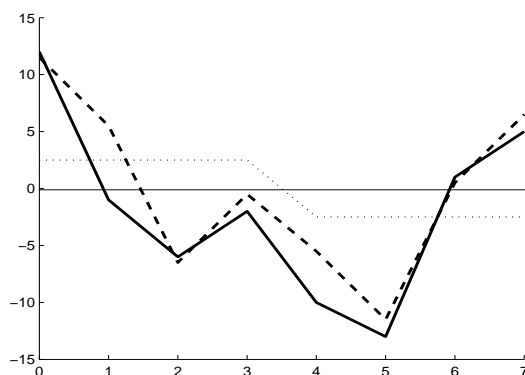
Obr. 5.32: Aproximácia procesu \mathbf{f} v dvojrozmernom podpriestore procesom $c_6 \mathbf{wh}_6 + c_3 \mathbf{wh}_3$ a vektor $c_3 \mathbf{wh}_3$

Veľkosť chyby tejto aproximácie je $\|\mathbf{e}_{36}\| = 10.96$ a pôvodný vektor a jeho aproximácia sú znázornené na obrázku 5.32

Najlepšou aproximáciou procesu \mathbf{f} v trojrozmernom podpriestore generovanom tromi vektormi Walsh-Hadamardovej bázy je vektor

$$\begin{aligned} c_3 \mathbf{wh}_3 + c_4 \mathbf{wh}_4 + c_6 \mathbf{wh}_6 &= 3(1, -1, -1, 1, 1, -1, -1, 1) + \\ &+ 2.5(1, 1, 1, 1, -1, -1, -1, -1) + \\ &+ 6(1, 1, -1, -1, -1, -1, 1, 1) + \\ &= (11.5, 5.5, -6.5, -0.5, -5.5, -11.5, 0.5, 6.5) \end{aligned}$$

Veľkosť chyby tejto aproximácie je $\|\mathbf{e}_{346}\| = 8.37$ a pôvodný vektor a jeho aproximácia sú znázornené na obrázku 5.33 □



Obr. 5.33: Aproximácia procesu \mathbf{f} v trojrozmernom podpriestore procesom $c_6 \mathbf{wh}_6 + c_3 \mathbf{wh}_3 + c_4 \mathbf{wh}_4$ a vektor $c_4 \mathbf{wh}_4$

Poznámka:

Metódou priemetu vektora do podpriestoru sme riešili tri úlohy: lineárnu regresiu, kľzavý priemer a diskretnú Fourierovu transformáciu. Ku základnej myšlienke priemetu sa v rôznych metódach menil spôsob výberu bázy vhodného podpriestoru.

- Lineárna regresia:
Časová zákonitosť; prekladanie procesu analytickou funkciou.
Hodnota v čase súvisí s hodnotami v iných časoch (je splnená podmienka linearity). Báza priestoru, do ktorého robíme projekciu, je tvorená vektormi, ktorých zložky sú funkciami času.
- Kľzavý priemer:
Hodnotu procesu v čase vypočítame z hodnôt tohto procesu v „susedných časoch“. Báza priestoru, do ktorého robíme projekciu, je tvorená vektormi, ktoré dostaneme posunutím hodnôt pôvodného vektora.
- Diskretná Fourierova transformácia alebo Walsh-Hadamardova transformácia:
Harmonická aj Walsh-Hadamardova báza sú dopredu dané a podpriestor vyberieme až po spočítaní koeficientov v tejto báze. Vyberieme tie báзовé vektory, ktorých koeficienty majú najväčšiu absolútnu hodnotu.

5.6 Použitie Fourierovej transformácie v praxi

Diskretná Fourierova transformácia našla široké použitie v praxi. Pri spracovaní obrazu sa používa jej variant, diskretná kosínusová transformácia, pri kompresii formátu JPEG. V spracovaní zvuku je Fourierova transformácia štandardným nástrojom pri odhade spektra zvuku, čo je nevyhnutný krok pri jeho digitálnom spracovaní ako aj napríklad pri rozpoznávaní reči.

Dôležitým problémom pri praktickom použití Fourierovej transformácie je jej rýchly výpočet. Fourierova transformácia spočíva v podstate vo vynásobení vektora fixnou štvorcovou maticou $n \times n$. Takýto výpočet vo všeobecnosti nemožno spraviť podstatne rýchlejšie než s pomocou n^2 operácií.

Cooley a Tukey [5] prišli na to, ako rýchlo počítať Fourierovu transformáciu v prípade, že n je presná mocnina 2 (napríklad 256). V tomto prípade na výpočet stačí len $\sim n \log n$ operácií. Tento algoritmus sa nazýva Fast Fourier Transform (FFT), a bol známy už aj Gaussovi.

V súčasnosti poznáme veľa druhov Fourierových transformácií [24], [7], a sú známe aj rýchle algoritmy, keď n je prvočíslo [14]. Najnovšie prebieha výskum algoritmov, ktoré počítajú Fourierovu transformáciu signálov s malým počtom nenulových spektrálnych koeficientov [9].

6 Analýza viacrozmerných objektov

V tejto kapitole si ukážeme metódy, ktoré sa používajú na analýzu štruktúry mnoho-rozmerných dát. Ako jednoduchý príklad použitia môžeme uviesť charakterové kvízy v časopisoch. V nich je séria 10-30 otázok s rôzne obodovanými hodnoteniami. Po zodpovedaní kvízu a sčítaní bodov získate hodnotenie nejakej ľudskej vlastnosti, napríklad kreativity, impulzivnosti, rozhodnosti a podobne. V tomto prípade ide o zobrazenie vektora odpovedí \mathbf{r} na číselné hodnotenie pomocou váhového vektora \mathbf{w}

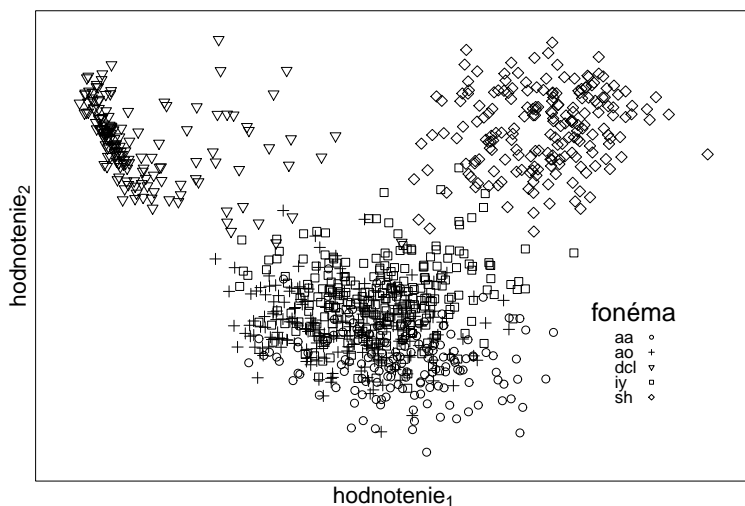
$$\text{hodnotenie charakteristiky} = \mathbf{w} \cdot \mathbf{r}$$

V iných prípadoch analýzy viacrozmerných dát nám jedno číselné hodnotenie nepostačuje. Aby sme pochopili štruktúru dát, môžeme sa pokúsiť vypočítať dve hodnotenia

$$\text{hodnotenie prvej charakteristiky} = \mathbf{w}_1 \cdot \mathbf{r}$$

$$\text{hodnotenie druhej charakteristiky} = \mathbf{w}_2 \cdot \mathbf{r}$$

a nakresliť si rozloženie týchto charakteristík do dvojrozmerného obrázka, ktorý ukazuje dvojrozmerný náhľad do vysoko rozmerných dát.



Obr. 6.1. Zobrazenie spektier anglických foném pomocou dvoch lineárnych funkcií

Znižovanie dimenzií sa často používa pri rozdelení dát na skupiny, ako je to potrebné napríklad pri rozpoznávaní reči. Príklad takého zníženia dimenzií pomocou lineárnych funkcií je znázornený na obrázku 6.1. Pomocou tohto zobrazenia do roviny získavame intuíciu o tom, ako sú vektory rozložené vo vysokorozmernom priestore, ktorý si nevieme predstaviť.

6.1 Dôležité súradnice

Úloha, ktorou sa budeme zaoberať je zjednodušenie popisu skupiny viacrozmerných dátových vektorov tým spôsobom, že sa urobí ich lineárna transformácia. Touto transformáciou získame informáciu o tých súradniciach, v ktorých sa vektory najviac líšia. Na ilustráciu načrtnutej myšlienky uvedieme nasledujúci príklad.

Príklad 6.1.1 *Zákazník si chce vybrať spomedzi 3 mobilných telefónov. Parametre, podľa ktorých sa rozhoduje sú cena, výdrž baterky, veľkosť pamäte a váha mobilu. Hodnoty parametrov pre jednotlivé mobily sú v tabuľke 6.1.*

	cena	baterka	pamäť	váha
M1	3700	81	400	40
M2	3700	82	200	20
M3	3700	81	500	50

Tab. 6.1. Cena, výdrž baterky, pamäť a váha troch mobilných telefónov

Zákazník sa chce rozhodnúť pre kúpu telefónu len podľa jedného parametra. Ktorý z parametrov si má pre svoje rozhodovanie vybrať?

Riešenie:

Riešenie úlohy vidno už z prvého pohľadu na tabuľku. Na cene nezáleží, pretože sa v nej telefóny nelíšia. Podobne je to s výdržou baterky, tu sú síce rozdiely, ale sú veľmi malé. Podľa prvých dvoch parametrov sa rozhodnúť nedá. Tretí a štvrtý parameter už majú telefóny rozdielne. Stačí však uvažovať iba jeden z nich, lebo druhý je jeho násobkom. Na rozlíšenie medzi telefónmi stačí zákazníkovi uvažovať napríklad váhu mobilu (posledný parameter). \square

V príklade 6.1.1 sme riešili úlohu nájsť také premenné, ktoré najlepšie odlíšia viacrozmerné dáta. Našli sme zjednodušenie, ktoré nám umožní zoradiť jednotlivé viacrozmerné dáta podľa jedného parametra.

V nasledujúcom príklade ukážeme ako sa dajú zmeniť viacrozmerné dáta na prehľadnejšie.

Príklad 6.1.2 *Referenti študijného oddelenia vysokej školy sa rozhodli urobiť rebríček úspešnosti študentov bakalárskeho stupňa. Ako kritérium si zvolili body za známky získané počas štúdia, ktoré sú uvedené v tabuľke 6.2. Najlepšie je hodnotenie 5, najhoršie je hodnotenie 1.*

Ako usporiadať študentov podľa študijných výsledkov?

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9
S_1	1	2	3	4	5	3	2	4	1
S_2	3	2	3	2	3	2	3	4	1
S_3	1	1	2	3	2	3	1	1	2
S_4	2	1	4	4	4	3	2	1	2
S_5	3	1	3	2	1	1	1	3	1
S_6	2	1	2	5	4	2	2	5	2
S_7	1	2	4	3	4	2	2	4	1
S_8	2	2	1	5	2	2	2	5	4
S_9	2	2	2	3	2	2	1	1	1
S_{10}	2	1	2	3	1	3	2	5	5
S_{11}	1	2	2	4	3	2	4	3	4

Tab. 6.2. Body 11 študentov z 9 predmetov

Riešenie:

Riešenie úlohy už nie je také evidentné ako v predchádzajúcom prípade. Načrtnutie grafov výsledkov jednotlivých študentov neurobí úlohu prehľadnejšou, ako to vidno na obrázku 6.2.

Zo školskej praxe samozrejme každý vie, že školské výsledky možno sumarizovať aj jediným číslom, ich aritmetickým priemerom. Hodnotenie reprezentované aritmetickým priemerom bude prehľadnejšie a študentov môžeme jednoducho usporiadať. Nevýhodou tohto postupu je, že aritmetický priemer nevytvorí o úspešnosti študenta príliš presne. V našom prípade sú aritmetické priemery jednotlivých študentov často zhodné, ako vidíme v tabuľke 6.3.

S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}
2.27	2.09	1.45	2.09	1.45	2.27	2.09	2.27	1.45	2.27	2.27

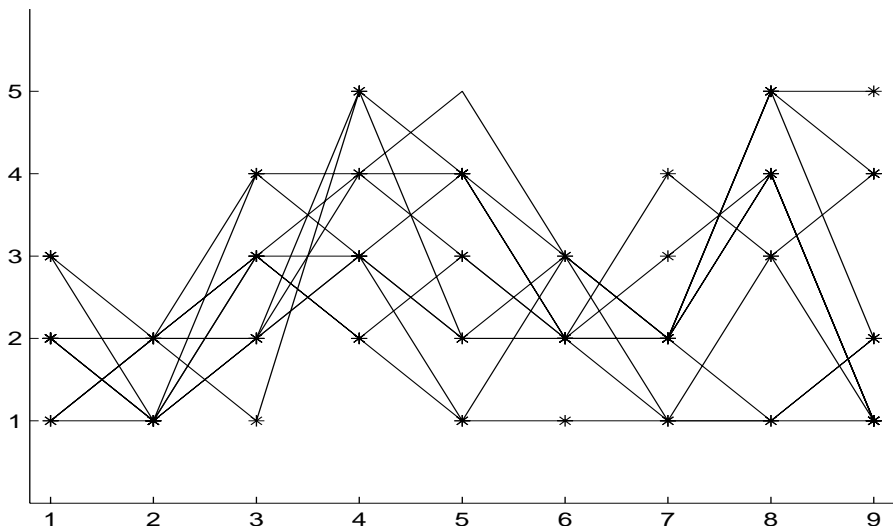
Tab. 6.3. Priemerné hodnotenie 11 študentov

Aritmetickým priemerom sme teda dosiahli rozdelenie študentov na 3 skupiny

$$A_1 = \{S_3, S_5, S_9\}, \quad A_2 = \{S_2, S_4, S_7\} \quad A_3 = \{S_1, S_6, S_8, S_{10}, S_{11}\}$$

Študentov z jednotlivých skupín A_1 , A_2 , A_3 už nedokážeme rozlíšiť. Zo skúsenosti však vieme, že rovnaký priemerný výsledok môžu dosiahnuť aj veľmi rozdielne úspešní študenti, respektíve študenti pri rôznom pracovnom výkone.

Predstavme si, že potrebujeme podstatne jemnejšie usporiadanie študentov podľa ich výsledkov. S riešením nám znovu pomôže bežná prax na vysokých školách, kde sa na ohodnotenie výsledkov štúdia používa takzvaný vážený priemer, ktorým zvýšime význam náročnejších predmetov a potlačíme jednoduchšie predmety, ktoré od študenta vyžadujú menej učenia sa. Predpokladajme, že úlohu riešime na technickej VŠ a z rozloženia úspešnosti v jednotlivých predmetoch vidíme, že predmet P_2 je



Obr. 6.2. Hodnotenie znázornené graficky

zrejme matematika a predmet P_4 je telesná výchova. Študijné oddelenie teda zvýši váhu predmetu P_2 tak, že výsledky tohto predmetu vynásobí dvomi a zníži váhu predmetu P_4 tak, že výsledky tohto predmetu vydělí dvomi. Po takejto úprave je možné rozlíšiť študentov presnejšie, ako vidíme v tabuľke 6.4.

S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}
2.78	2.67	1.72	2.44	1.78	2.61	2.61	2.72	1.83	2.72	2.78

Tab. 6.4. Vážené priemerné hodnotenie 11 študentov

Váženým aritmetickým priemerom sme dosiahli rozloženie študentov podľa prospechu na 8 skupín

$$A_1 = \{S_3\}, \quad A_2 = \{S_5\}, \quad A_3 = \{S_9\}, \quad A_4 = \{S_4\}, \quad A_5 = \{S_6, S_7\},$$

$$A_6 = \{S_2\}, \quad A_7 = \{S_8, S_{10}\}, \quad A_8 = \{S_1, S_{11}\}$$

Študentov z jednotlivých skupín A_5 , A_7 a A_8 nedokážeme rozlíšiť, ale študenti sú usporiadaní lepšie ako v predchádzajúcom prípade.

Znovu, ale prehovorí skúsenosť zo škôl, že váha predmetu sa môže zmeniť (skúša prísnejší učiteľ), alebo nezodpovedá náročnosti predmetu. Skúsme preto len zo známych výsledkov, bez toho, aby sme vedeli o aké predmety ide, určiť také váhy jednotlivých predmetom, že rozlíšenie študentov pomocou váženého priemeru bude čo najlepšie. Táto úloha bude témou nasledujúcej podkapitoly. \square

6.2 Váhové koeficienty pre maximálny rozptyl

Ukázali sme, že v prípade analýzy viacrozmerných dát, môžu byť niektoré súradnice viac a niektoré menej významné. Vhodnou lineárnou kombináciou môžeme z matice dát získať sadu čísel, ktoré popisujú jednotlivé riadky matice. Tieto čísla sú dostatočne rozptýlené na to, aby sa dali pôvodné dáta od seba navzájom dobre odlíšiť.

Ukážeme ako pre dáta z príkladu 6.1.1 rozdeliť váhy medzi jednotlivé parametre telefónu tak, aby výsledné lineárne kombinácie mali čo najväčší rozptyl. Predpokladáme, že dáta, ktoré ideme spracovávať, sú prvkami matice \mathbb{M} . Každý jej riadok sú parametre jedného mobilného telefónu.

Vektor $\mathbf{w} = (w_1, w_2, w_3, w_4)$ je vektor, ktorého zložky sú váhy jednotlivých parametrov.

Chceme, aby vektor $\mathbb{M}\mathbf{w}^T$ mal, čo najväčší rozptyl, teda aby sa vážené priemery jednotlivých telefónov od seba čo najviac líšili.

$$\begin{aligned}\mathbb{M}\mathbf{w}^T &= \begin{pmatrix} 3700 & 81 & 400 & 40 \\ 3700 & 82 & 200 & 20 \\ 3700 & 81 & 500 & 50 \end{pmatrix} \mathbf{w}^T = \begin{pmatrix} 3700 & 81 & 400 & 40 \\ 3700 & 82 & 200 & 20 \\ 3700 & 81 & 500 & 50 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} \\ &= \begin{pmatrix} 3700w_1 + 81w_2 + 400w_3 + 40w_4 \\ 3700w_1 + 82w_2 + 200w_3 + 20w_4 \\ 3700w_1 + 81w_2 + 500w_3 + 50w_4 \end{pmatrix}\end{aligned}$$

Úloha 6.2.1 *Maximalizovať rozptyl budeme tak, že budeme maximalizovať súčet vzdialeností všetkých hodnôt od priemernej hodnoty. Pre pohodlnejšie výpočty budeme ako vzdialenosť používať druhú mocninu rozdielu.*

Ukážte, že rovnaký rozptyl ako vektor $\mathbb{M}\mathbf{w}^T$ má vektor $\mathbb{M}^\circ\mathbf{w}^T$, v ktorom maticu \mathbb{M}° dostaneme z matice \mathbb{M} tak, že od každého stĺpca odčítame jeho aritmetický priemer.

$$\mathbb{M}^\circ = \begin{pmatrix} 0 & -0.3 & 33.3 & 3.3 \\ 0 & 0.7 & -166.7 & -16.7 \\ 0 & -0.3 & 133.3 & 13.3 \end{pmatrix}$$

Príklad 6.2.1 *Ukážte, že aritmetický priemer každého stĺpca matice \mathbb{M}° je 0.*

Riešenie:

Tvrdenie vyplýva priamo z konštrukcie matice \mathbb{M}° . Súčet jedného stĺpca matice \mathbb{M} je vlastne N -násobok jeho aritmetického priemeru. Matica \mathbb{M}° vznikne z \mathbb{M} tak, že od každého prvku v stĺpci odpočítame aritmetický priemer celého stĺpca, spolu N -krát. Súčet jedného stĺpca matice \mathbb{M}° preto je:

$$N \times \text{aritmetický priemer} - N \times \text{aritmetický priemer} = 0$$

Preto priemer každého stĺpca bude 0. □

Hľadáme také \mathbf{w} , ktoré bude maximalizovať rozptyl vektora $\mathbb{M}\mathbf{w}^T$ od jeho priemernej hodnoty, resp. vektora $\mathbb{M}^\circ\mathbf{w}^T$ od hodnoty 0.

$$\mathbb{M}^\circ \mathbf{w}^T = \begin{pmatrix} 0w_1 - 0.3w_2 + 33.3w_3 + 3.3w_4 \\ 0w_1 0.67w_2 - 166.67w_3 - 16.67w_4 \\ 0w_1 - 0.33w_2 + 133.33w_3 + 13.33w_4 \end{pmatrix}$$

Rozptyl ROZ vektora $\mathbb{M} \mathbf{w}^T$ okolo hodnoty 0 je

$$ROZ = (-0.3w_2 + 33.3w_3 + 3.3w_4 - 0)^2 + (0.67w_2 - 166.67w_3 - 16.67w_4 - 0)^2 + \\ + (-0.33w_2 + 133.33w_3 + 13.33w_4 - 0)^2$$

Maticovo môžeme rozptyl ROZ zapísať ako:

$$ROZ = \mathbf{w} \mathbb{M}^{\circ T} \mathbb{M}^\circ \mathbf{w}^T = (-0.3w_2 + 33.3w_3 + 3.3w_4)^2 \\ + (0.67w_2 - 166.67w_3 - 16.67w_4)^2 + (-0.33w_2 + 133.33w_3 + 13.33w_4)^2 \quad (6.2.1)$$

Úloha 6.2.2 Ako vzdialenosť sme použili

$$\|\mathbf{v} - \mathbf{u}\|^2 = (v_0 - u_0)^2 + (v_1 - u_1)^2 + \dots + (v_{N-1} - u_{N-1})^2$$

Ukážte, že táto hodnota bude maximálna vtedy, keď bude maximálna hodnota

$$\|\mathbf{v} - \mathbf{u}\| = \sqrt{(v_0 - u_0)^2 + (v_1 - u_1)^2 + \dots + (v_{N-1} - u_{N-1})^2}$$

Vidíme, že výraz (6.2.1) dosiahne tým väčšiu hodnotu, čím je väčší vektor \mathbf{w} , preto musíme určiť nejakú ďalšiu podmienku, aby sme mohli hľadať maximum. Pretože v našej úlohe vektor \mathbf{w} reprezentoval váhové koeficienty, budeme predpokladať, že veľkosť tohto vektora je 1.

Riešime teda úlohu, nájsť spomedzi všetkých vektorov \mathbf{w} dĺžky 1 taký, ktorý maximalizuje výraz (6.2.1).

$$\mathbf{w} \mathbb{M}^{\circ T} \mathbb{M}^\circ \mathbf{w}^T \rightarrow \max \quad \text{za podmienky} \quad \|\mathbf{w}\| = 1$$

Túto úlohu môžeme riešiť metódami matematickej analýzy. Je to úloha nájsť viazaný extrém funkcie. Riešenie pomocou Lagrangeových multiplikátorov vedie na hľadanie extrému funkcie viac premenných tvaru

$$f(\mathbf{w}) = \mathbf{w} \mathbb{M}^{\circ T} \mathbb{M}^\circ \mathbf{w}^T - \lambda(\mathbf{w} \mathbf{w}^T - 1)$$

Hodnoty, pre ktoré táto funkcia dosiahne maximum, musia spĺňať podmienky pre stacionárne body funkcie:

$$\frac{\partial f}{\partial \mathbf{w}} = 2\mathbb{M}^{\circ T} \mathbb{M}^\circ \mathbf{w}^T - 2\lambda \mathbf{w}^T = 0$$

$$\frac{\partial f}{\partial \lambda} = -\mathbf{w} \mathbf{w}^T + 1 = 0$$

Úpravami dostaneme rovnicu pre neznámy parameter λ a neznámy vektor \mathbf{w} , pre ktorý má byť $\|\mathbf{w}\| = 1$.

$$\mathbb{M}^{\circ T} \mathbb{M}^{\circ} \mathbf{w}^T = \lambda \mathbf{w}^T \quad (6.2.2)$$

Hľadáme taký vektor \mathbf{w}^T , že jeho vynásobenie maticou sa dá nahradiť vynásobením jedným číslom. Takýto vektor sa nazýva vlastný vektor matice a λ sa nazýva vlastné číslo tejto matice.

Vzťah (6.2.2) môžeme zapísať v maticovom tvare:

$$\left(\mathbb{M}^{\circ T} \mathbb{M}^{\circ} - \lambda \mathbb{E} \right) \mathbf{w}^T = \mathbf{0} \quad (6.2.3)$$

Riešením rovnice (6.2.3) je napríklad vektor $\mathbf{w} = (0, 0, \dots, 0)$. Takéto riešenie ale nemôžeme použiť ako váhový vektor. Hľadáme preto nenulové riešenie rovnice (6.2.3). Teda chceme, aby táto rovnica mala aspoň jedno nenulové riešenie. To nastane iba vtedy, ak matica $\mathbb{M}^{\circ T} \mathbb{M}^{\circ} - \lambda \mathbb{E}$ bude singulárna, a teda determinant matice bude 0.

$$\det \left(\mathbb{M}^{\circ T} \mathbb{M}^{\circ} - \lambda \mathbb{E} \right) = 0 \quad (6.2.4)$$

V našom príklade s mobilnými telefónmi dostaneme úpravami rovnice (6.2.4) rovnicu pre neznámu hodnotu λ :

$$\det \begin{pmatrix} 0 - \lambda & 0 & 0 & 0 \\ 0 & 1 - \lambda & -167 & -17 \\ 0 & -167 & 46667 - \lambda & 4667 \\ 0 & -17 & 4667 & 467 - \lambda \end{pmatrix} = 0$$

Vlastné čísla $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ sú riešením rovnice

$$\begin{aligned} & (-\lambda) \left[(1 - \lambda)(46667 - \lambda)(467 - \lambda) + (-17)(-167)(4667) + (-167)(4667)(-17) \right. \\ & \left. - (-17)(46667 - \lambda)(-17) - (1 - \lambda)(4667)(4667) - (-167)(-167)(467 - \lambda) \right] = 0 \end{aligned}$$

Riešenia sú približne

$$\lambda_1 = 0, \lambda_2 = 0.03, \lambda_3 = 0.64, \lambda_4 = 47134.3$$

V nasledujúcej úprave nahradíme násobenie maticou násobením vlastným číslom, dosadíme $\mathbf{w} \mathbf{w}^T = \|\mathbf{w}\|^2 = 1$ a upravíme funkciu $f(\mathbf{w})$:

$$f(\mathbf{w}) = \mathbf{w} \mathbb{M}^{\circ T} \mathbb{M}^{\circ} \mathbf{w}^T - \lambda(\mathbf{w} \mathbf{w}^T - 1) = \mathbf{w}(\lambda)^2 \mathbf{w}^T - \lambda(\mathbf{w} \mathbf{w}^T - 1) = (\lambda)^2$$

Túto úvahu urobíme neskôr presnejšie, aj teraz však intuitívne vidíme, že maximálny rozptyl dosiahneme pre najväčšie vlastné číslo $\lambda = 47134$ a jemu zodpovedajúci vlastný vektor, ktorý vypočítame ako riešenie rovnice (6.2.3) pre konkrétnu hodnotu λ .

Vlastný vektor, zodpovedajúci vlastnému číslu $\lambda = 47134$ je vektor $\mathbf{w} = (0, -0.0036, 0.995, 0.0995)$. Pre najväčší rozptyl lineárnych kombinácií treba použiť v prípade mobilného telefónu váhové koeficienty

$$w_1 = 0, w_2 = -0.0036, w_3 = 0.995, w_4 = 0.0995$$

	cena	baterka	pamäť	váha	vážený priemer
M1	3700	81	400	40	401.7
M2	3700	82	200	20	200.7
M3	3700	81	500	50	502.2

Tab. 6.5. Cena, pamäť, výdrž baterky, váha troch mobilných telefónov

Keby sme chceli telefóny odlíšiť iba podľa jediného parametra, bol by to tretí parameter (pamäť telefónu). S týmto parametrom dosiahneme najväčšie odlíšenie medzi telefónmi. \square

Poznámka:

Z riešenia príkladu o mobilných telefónoch je zrejmé, že najvýraznejší rozptyl dosiahneme, keď dáme vysokú váhu tým premenným, ktoré majú veľký rozptyl a vysoké číselné hodnoty. Ak sú číselné parametre v rôznych jednotkách, ich interpretácia môže byť skreslená.

Lepšie výsledky dá uvedená metóda v prípade, že všetky parametre budú mať rovnaký rozsah hodnôt. Takáto situácia nastane v príklade so známami (príklad 6.1.2) .

Príklad 6.2.2 *Nájdite váhové koeficienty pre najväčší rozptyl. Použite dáta z príkladu 6.1.2, v ktorom sú popísané známky 11 študentov z 9 predmetov.*

Riešenie:

Podobne ako v predchádzajúcej úlohe vytvoríme z dát maticu nameraných hodnôt.

$$M = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 3 & 2 & 4 & 1 \\ 3 & 2 & 3 & 2 & 3 & 2 & 3 & 4 & 1 \\ 1 & 1 & 2 & 3 & 2 & 3 & 1 & 1 & 2 \\ 2 & 1 & 4 & 4 & 4 & 3 & 2 & 1 & 2 \\ 3 & 1 & 3 & 2 & 1 & 1 & 1 & 3 & 1 \\ 2 & 1 & 2 & 5 & 4 & 2 & 2 & 5 & 2 \\ 1 & 2 & 4 & 3 & 4 & 2 & 2 & 4 & 1 \\ 2 & 2 & 1 & 5 & 2 & 2 & 2 & 5 & 4 \\ 2 & 2 & 2 & 3 & 2 & 2 & 1 & 1 & 1 \\ 2 & 1 & 2 & 3 & 1 & 3 & 2 & 5 & 5 \\ 1 & 2 & 2 & 4 & 3 & 2 & 4 & 3 & 4 \end{pmatrix}$$

Z centrovanej matice M° vytvoríme súčin matíc $\mathbb{M}^{\circ T} \mathbb{M}^\circ$:

$$\begin{pmatrix} 5.64 & -0.91 & 0.27 & -3.09 & -4.36 & -2.45 & -1 & 1.55 & -1.64 \\ -0.91 & 2.73 & -0.82 & 0.27 & 2.09 & -0.64 & 2 & 1.36 & -1.09 \\ 0.27 & -0.82 & 8.55 & -4.18 & 4.27 & 1.09 & 0 & -1.91 & -5.27 \\ -3.09 & 0.27 & -4.18 & 10.73 & 5.91 & 1.64 & 2 & 4.64 & 6.09 \\ -4.36 & 2.09 & 4.27 & 5.91 & 17.64 & 2.55 & 4 & 1.55 & -7.64 \\ -2.45 & -0.64 & 1.09 & 1.64 & 2.55 & 4.18 & 0 & -1.82 & 2.45 \\ -1 & 2 & 0 & 2 & 4 & 0 & 8 & 5 & 5 \\ 1.55 & 1.36 & -1.91 & 4.64 & 1.55 & -1.82 & 5 & 26.18 & 8.45 \\ -1.64 & -1.09 & -5.27 & 6.09 & -7.64 & 2.45 & 5 & 8.45 & 21.64 \end{pmatrix}$$

Nájďme vlastné čísla matice $\mathbb{M}^{\circ T} \mathbb{M}^\circ$:

$$\lambda_1 = 0, \lambda_2 = 0.1, \lambda_3 = 0.17, \lambda_4 = 0.37, \lambda_5 = 0.66, \lambda_6 = 0.94, \lambda_7 = 1.8,$$

$$\lambda_8 = 2.69, \lambda_9 = 3.81$$

Nájďme vlastný vektor zodpovedajúci najväčšiemu vlastnému číslu λ :

$$\mathbf{w} = (-0.02, 0.02, -0.21, 0.28, -0.09, 0.01, 0.22, 0.66, 0.62)$$

Váhy jednotlivých vyučovacích predmetov, použiteľné pre vážený priemer s najväčším rozptylom sú:

$$w_1 = -0.02, w_2 = 0.02, w_3 = -0.21, w_4 = 0.28, w_5 = -0.09, w_6 = 0.01,$$

$$w_7 = 0.22, w_8 = 0.66, w_9 = 0.62$$

Vážené priemery jednotlivých riadkov matice M , resp. priemery známok pre jednotlivých študentov sú

$$3.8, \quad 3.58, \quad 2.4, \quad 2.28, \quad 2.63, \quad 5.6, \quad 3.39, \quad 7.25, \quad 1.76, \quad 6.98, \quad 5.82,$$

V tabuľke 6.6 je popísaný prospech študentov pomocou aritmetického priemeru (stĺpec „priem“), váženého priemeru s odhadnutými váhami (stĺpec „vp1“) a váženého priemeru s vypočítanými váhami (stĺpec „vp2“). Výpočet váženého priemeru pomocou vlastného vektora rozptylovej matice zabezpečí nielen najväčší rozptyl, ale aj vyššiu váhu tých hodnôt, ktoré majú väčší rozptyl. Konkrétne v tomto príklade to znamená, že ak majú všetci študenti dobrú známku z telesnej výchovy, tak vo váženom priemere vystupuje telesná výchova s nízkym váhovým koeficientom. Naopak, ak sú z informatiky dobré aj zlé známky, tak informatika bude mať vysoký váhový koeficient. A nakoniec, ak sú z matematiky iba zlé známky, tak matematika bude mať nízky váhový koeficient.

Pri konkrétnom spracovaní dát treba uvážiť všetky výhody aj nevýhody použitia tejto metódy. V každom prípade však touto metódou môžeme zistiť, ktoré parametre sú pre rozlíšenie dát významné a ktoré menej významné. \square

Ukázali sme, ako vybrať váhové koeficienty tak, aby vážený priemer jednotlivých skupín dát mal čo najväčší rozptyl. Zároveň sme však videli, že v prípade rôznych rozsahov jednotlivých parametrov, budú mať väčšiu váhu tie parametre, ktoré sú zadane v jednotkách s väčšími hodnotami. Ak uvedieme váhu v gramoch dosiahneme väčší rozptyl, než keď tú istú váhu uvedieme v kilogramoch.

Z uvedeného dôvodu je zrejmé, že metóda hľadania váhových koeficientov pre najväčší rozptyl bude dávať rozumné výsledky v tom prípade, keď budú mať rôzne parametre rovnaký rozsah hodnôt. To je dôvod, prečo metóda dáva realistickejšie výsledky v príklade o školskom prospechu (6.1.2). V príklade o mobilných telefónoch (6.1.2) bola parametrom „cena“ a „výdrž baterky“ telefónu priradená váha 0. V prípade ceny to bolo opodstatnené, nakoľko sa telefóny v cene nelíšili. V prípade parametra „výdrž baterky“ by sme dostali podstatne väčší váhový koeficient, keby bol tento parameter zadáný v sekundách namiesto v minútach.

Príklad 6.2.3 *Pri kúpe auta si zákazník zostavil tabuľku s hodnotením 14 parametrov pre 5 rôznych áut. V tabuľke si vyznačil mieru svojej spokojnosti s jednotlivými parametrami auta bodmi od 0 po 10.*

Určte váhy jednotlivých parametrov a vypočítajte vážený priemer pre mieru spokojnosti pre jednotlivé autá. Ktoré z áut má pri takomto výpočte najvyššie hodnotenie?

Riešenie:

K matici \mathbb{M} predstavujúcej dáta zostrojíme centrovанú maticu \mathbb{M}° .

Nájdeme vlastné čísla matice $\mathbb{M}^{\circ T} \mathbb{M}^\circ$. Sú to postupne čísla:

$$\begin{aligned} \lambda_0 &= 0, & \lambda_1 &= 0, \dots, & \lambda_{10} &= 0, & \lambda_{11} &= 58, & \lambda_{12} &= 103.79, \\ \lambda_{13} &= 140.18, & \lambda_{14} &= 200.82 \end{aligned}$$

K najväčšiemu vlastnému číslu $\lambda = 200.82$ prislúcha vlastný vektor

$$\mathbf{w} = (0.13, 0.56, 0.23, 0.24, -0.31, -0.26, 0.32, 0.03, 0.33, 0.25, 0.18, 0.03, -0.12, 0.27)$$

Keď tento vektor použijeme ako váhový vektor pre hodnotenie parametrov, dostaneme iné poradie, než pri poradí áut podľa aritmetického priemeru. Z hodnôt v tabuľke 6.8 vidíme, že ak usporiadame autá podľa vážených priemerov, najlepšie hodnotené je auto a_4 a najhoršie auto a_3 . Pri hodnotení podľa aritmetického priemeru je rovnako najlepšie hodnotené auto a_4 ale najhoršie hodnotené je auto a_1 . \square

Pri riešení príkladu 6.2.3 sme vypočítali, že matica má iba 4 nenulové vlastné čísla. V ďalšom odseku ukážeme, ako vypočítať ďalšie vektory $\mathbf{w}_2, \mathbf{w}_3, \dots$, ktoré budú tvoriť spolu s vektorom $\mathbf{w} = \mathbf{w}_1$ ortogonálnu bázu. Na veľmi presnú aproximáciu pôvodných dát stačí malý počet vektorov tejto bázy. Na vyjadrenie sa používajú iba tie vektory, ktoré zodpovedajú nenulovým vlastným číslam matice $\mathbb{M}^{\circ T} \mathbb{M}^\circ$, pričom \mathbb{M} je matica pôvodných hodnôt.

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	priem	vp1	vp2
S_1	1	2	3	4	5	3	2	4	1	2.27	2.78	3.8
S_2	3	2	3	2	3	2	3	4	1	2.09	2.67	3.58
S_3	1	1	2	3	2	3	1	1	2	1.45	1.72	2.4
S_4	2	1	4	4	4	3	2	1	2	2.09	2.44	2.28
S_5	3	1	3	2	1	1	1	3	1	1.45	1.78	2.63
S_6	2	1	2	5	4	2	2	5	2	2.27	2.61	5.6
S_7	1	2	4	3	4	2	2	4	1	2.09	2.61	3.39
S_8	2	2	1	5	2	2	2	5	4	2.27	2.72	7.25
S_9	2	2	2	3	2	2	1	1	1	1.45	1.83	1.76
S_{10}	2	1	2	3	1	3	2	5	5	2.27	2.72	6.98
S_{11}	1	2	2	4	3	2	4	3	4	2.27	2.78	5.82

Tab. 6.6. Body 11 študentov z 9 predmetov a zodpovedajúce vážené priemery

parameter	1	2	3	4	5	6	7	8	9	10	11	12	13	14
a_1	5	2	1	1	4	10	4	4	6	7	1	2	1	0
a_2	8	9	1	3	1	5	0	9	8	4	0	4	6	7
a_3	1	1	5	4	9	6	1	7	2	0	0	5	6	5
a_4	3	9	8	6	2	4	9	5	8	9	3	3	1	9
a_5	6	10	6	8	5	3	4	8	9	1	4	7	4	2

Tab. 6.7. Miera spokojnosti zákazníka so 14 parametrami pre 5 áut

	a_1	a_2	a_3	a_4	a_5
priemer	3.43	4.64	3.71	5.64	5.5
vážený priemer	3.66	10.53	0.38	17.9	13.05

Tab. 6.8. Priemerná a vážená priemerná miera spokojnosti zákazníka

6.3 Analýza hlavných komponentov

S úlohou riešenou v predchádzajúcej kapitole súvisí metóda, ktorá sa nazýva analýza hlavných komponentov (PCA). Túto metódu uviedol Angličan Karl Pearson v roku 1901. Rieši úlohu vyjadriť maticu nameraných hodnôt v inej báze. Táto báza bude taká, aby na dobrú aproximáciu stačil menší počet súradníc. V predchádzajúcom odseku sme ukázali, že maximálny rozptyl dosiahne taká lineárna kombinácia nameraných hodnôt, v ktorej sú ako koeficienty použité zložky vlastného vektora, prislúchajúceho najväčšiemu vlastnému číslu matice

$$\mathbb{R} = \mathbb{M}^{\circ T} \mathbb{M}^{\circ}$$

Uvedená lineárna kombinácia obsahuje najväčšie množstvo neurčitosti, a teda aj informácie o nameraných hodnotách. Váhový vektor z predchádzajúcej úlohy môžeme chápať aj ako bázový vektor. Násobky tohto vektora sú najlepšou aproximáciou spomedzi všetkých jednorozmerných aproximácií pôvodných dát. Preto sa táto lineárna kombinácia nazýva prvým hlavným komponentom. Ak prvý hlavný komponent z dát odstránime, znovu môžeme riešiť rovnakú úlohu. Riešením bude druhé najväčšie vlastné číslo a jemu prislúchajúci vlastný vektor. Priemet pôvodných vektorov do podpriestoru generovaného vlastným vektorom s druhým najväčším vlastným číslom sa nazýva druhý hlavný komponent.

Uvedený výsledok nie je samozrejímavý, ani triviálny. Na jeho odvodenie sa využije nasledujúce tvrdenie.

Tvrdenie 6.3.1 *Vlastné čísla ľubovoľnej matice, ktorá sa dá vyjadriť v tvare $\mathbb{A}^T \mathbb{A}$ sú reálne čísla väčšie alebo rovné 0. Vlastné vektory, ktoré prislúchajú týmto vlastným číslam, tvoria ortogonálny systém vektorov (ktorý sa dá upraviť na ortonormálny).*

Celý postup PCA spočíva v niekoľkých krokoch, v ktorých sa spracuje matica nameraných hodnôt

$$\mathbb{M} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_m \end{pmatrix}$$

Rozmer matice \mathbb{M} je $m \times n$, matica obsahuje dáta pre m objektov, každý s n parametrami.

- z matice \mathbb{M} vytvoríme centrovanú maticu \mathbb{M}° tak, že od každého stĺpca matice odčítame jeho aritmetický priemer
- vytvoríme maticu $\mathbb{R} = \mathbb{M}^{\circ T} \mathbb{M}^{\circ}$, ktorej rozmer je $n \times n$
- nájdeme vlastné čísla λ_k a vlastné vektory \mathbf{w}_k matice \mathbb{R}
- zoradíme vlastné čísla podľa veľkosti a označíme tak, aby platilo

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$$

- zodpovedajúce vlastné vektory označíme

$$\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$$

dá sa overiť, že tieto vektory sú ortogonálne

- riadky matice \mathbb{M}° označíme ako vektory $\mathbf{v}_1^\circ, \mathbf{v}_2^\circ, \dots, \mathbf{v}_m^\circ$
- najpresnejšia aproximácia vektorov $\mathbf{v}_1^\circ, \mathbf{v}_2^\circ, \dots, \mathbf{v}_m^\circ$ spomedzi všetkých jednorozmerných podpriestorov bude v podpriestore generovanom vektorom \mathbf{w}_1
- najpresnejšia aproximácia vektorov $\mathbf{v}_1^\circ, \mathbf{v}_2^\circ, \dots, \mathbf{v}_m^\circ$ spomedzi všetkých dvojrozmerných podpriestorov bude v podpriestore generovanom vektormi $\mathbf{w}_1, \mathbf{w}_2$
- najpresnejšia aproximácia vektorov $\mathbf{v}_1^\circ, \mathbf{v}_2^\circ, \dots, \mathbf{v}_m^\circ$ spomedzi všetkých k -rozmerných podpriestorov bude v podpriestore generovanom vektormi $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k$
- vypočítame koeficienty vektorov $\tilde{\mathbf{v}}_1^\circ, \tilde{\mathbf{v}}_2^\circ, \dots, \tilde{\mathbf{v}}_m^\circ$ v ortogonálnej báze $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k$ ako koeficienty priemetov vektorov \mathbf{v}° na jednotlivé vektory tejto bázy, teda podľa vzorcov

$$\tilde{\mathbf{v}}^\circ = c_1 \cdot \mathbf{w}_1 + c_2 \cdot \mathbf{w}_2 + \dots + c_k \cdot \mathbf{w}_k$$

$$c_j = \frac{\langle \mathbf{v}^\circ, \mathbf{w}_j \rangle}{\langle \mathbf{w}_j, \mathbf{w}_j \rangle}$$

- aproximáciu pôvodných vektorov $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ dostaneme tak, že ku vektorom $\mathbf{v}_1^\circ, \mathbf{v}_2^\circ, \dots, \mathbf{v}_m^\circ$ pripočítame aritmetické priemery jednotlivých stĺpcov matice \mathbb{M}

Metódu PCA odvodíme neskôr, zatiaľ ju iba ilustrujeme na úlohe zmenšenia počtu koeficientov potrebných na vyjadrenie sedemrozmerných vektorov.

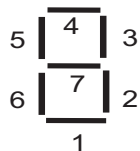
Príklad 6.3.1 Číslce na digitálnych hodinkách sú tvorené siedmimi čiarkami, ktoré buď svietia, alebo nesvietia. Na určenie, o akú číslicu ide, použijeme 7 hodnôt. Hodnota 0 znamená, že príslušná čiarka nesvieti, hodnota 1 znamená, že príslušná čiarka svieti. Nájdite pomocou metódy PCA hlavné komponenty tak, aby bolo možné určiť každú z digitálnych číslíc pomocou menšieho počtu koeficientov než 7.

Riešenie:

Vyobrazenia číslíc na digitálnych hodinkách, tak ako ich budeme v úlohe používať sú na obrázku 6.3. V každej číslici sú niektoré z čiarok zasvietené, iné nie. Hrubá čiara na obrázku znamená zasvietenú čiarku a bude jej priradená hodnota 1. Tenká čiara znamená, že príslušné miesto je nezasvietené a bude mu priradená hodnota 0.



Obr. 6.3. Vyobrazenia číslíc 0, 1, 2, 3, 4, 5, 6, 7, 8, 9



Obr. 6.4. Čísla 1, 2, 3, 4, 5, 6, 7 označujú pozíciu čiarky v digitálnom čísle

Na obrázku 6.4 sú pozície jednotlivých čiarok označené číslami.

Jednotlivé číslice budú reprezentované vektormi, ktorých súradnice sú hodnoty 0, alebo 1 na zodpovedajúcich miestach.

$$\mathbf{v}_0 = (1, 1, 1, 1, 1, 1, 0)$$

$$\mathbf{v}_1 = (0, 1, 1, 0, 0, 0, 0)$$

$$\mathbf{v}_2 = (1, 0, 1, 1, 0, 1, 1)$$

$$\mathbf{v}_3 = (1, 1, 1, 1, 0, 0, 1)$$

$$\mathbf{v}_4 = (0, 1, 1, 0, 1, 0, 1)$$

$$\mathbf{v}_5 = (1, 1, 0, 1, 1, 0, 1)$$

$$\mathbf{v}_6 = (1, 1, 0, 1, 1, 1, 1)$$

$$\mathbf{v}_7 = (0, 1, 1, 1, 0, 0, 0)$$

$$\mathbf{v}_8 = (1, 1, 1, 1, 1, 1, 1)$$

$$\mathbf{v}_9 = (0, 1, 1, 1, 1, 0, 1)$$

Matica \mathbb{M} reprezentujúca namerané dáta, v našom prípade vektory $\mathbf{v}_0, \dots, \mathbf{v}_9$ má tvar:

$$\mathbb{M} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Vypočítame aritmetický priemer každého stĺpca matice. Dostaneme vektor aritmetických priemerov \mathbf{m} .

$$\mathbf{m} = (0.6, 0.9, 0.8, 0.7, 0.6, 0.4, 0.7)$$

Z matice \mathbb{M} vypočítame centrovanú maticu $\mathbb{M}^\circ = \mathbb{M} - \mathbf{m}$.

$$\mathbb{M}^\circ = \begin{pmatrix} 0.4 & 0.1 & 0.2 & 0.3 & 0.4 & 0.6 & -0.7 \\ -0.6 & 0.1 & 0.2 & -0.7 & -0.6 & -0.4 & -0.7 \\ 0.4 & -0.9 & 0.2 & 0.3 & -0.6 & 0.6 & 0.3 \\ 0.4 & 0.1 & 0.2 & 0.3 & -0.6 & -0.4 & 0.3 \\ -0.6 & 0.1 & 0.2 & -0.7 & 0.4 & -0.4 & 0.3 \\ 0.4 & 0.1 & -0.8 & 0.3 & 0.4 & -0.4 & 0.3 \\ 0.4 & 0.1 & -0.8 & -0.7 & 0.4 & 0.6 & 0.3 \\ -0.6 & 0.1 & 0.2 & 0.3 & -0.6 & -0.4 & -0.7 \\ 0.4 & 0.1 & 0.2 & 0.3 & 0.4 & 0.6 & 0.3 \\ -0.6 & 0.1 & 0.2 & 0.3 & 0.4 & -0.4 & 0.3 \end{pmatrix}$$

Vypočítame maticu $\mathbb{R} = \mathbb{M}^{\circ T} \mathbb{M}^\circ$.

$$\mathbb{R} = \begin{pmatrix} 0.27 & -0.04 & -0.09 & 0.09 & 0.04 & 0.18 & 0.09 \\ -0.04 & 0.1 & -0.02 & -0.03 & 0.07 & -0.07 & -0.03 \\ -0.09 & -0.02 & 0.18 & 0.04 & -0.09 & -0.02 & -0.07 \\ 0.09 & -0.03 & 0.04 & 0.23 & -0.02 & 0.02 & 0.01 \\ 0.04 & 0.07 & -0.09 & -0.02 & 0.27 & 0.07 & 0.09 \\ 0.18 & -0.07 & -0.02 & 0.02 & 0.07 & 0.27 & 0.02 \\ 0.09 & -0.03 & -0.07 & 0.01 & 0.09 & 0.02 & 0.23 \end{pmatrix}$$

Nájďme všetky vlastné čísla matice \mathbb{R} a zoradíme ich podľa veľkosti.

$$\lambda_1 = 0.57, \lambda_2 = 0.37, \lambda_3 = 0.22, \lambda_4 = 0.19, \lambda_5 = 0.14, \lambda_6 = 0.05, \lambda_7 = 0.01$$

K vlastným číslam vypočítame zodpovedajúce vlastné vektory.

$$\begin{aligned} \mathbf{w}_1 &= (0.6, -0.09, -0.29, 0.15, 0.36, 0.52, 0.35) \\ \mathbf{w}_2 &= (-0.27, 0.32, -0.36, -0.46, 0.59, -0.28, 0.25) \\ \mathbf{w}_3 &= (0.05, 0.02, -0.04, 0.6, -0.05, -0.59, 0.53) \\ \mathbf{w}_4 &= (0.11, -0.35, -0.27, -0.53, -0.57, -0.09, 0.43) \\ \mathbf{w}_5 &= (-0.36, -0.3, 0.64, -0.08, 0.23, 0.26, 0.49) \\ \mathbf{w}_6 &= (0.49, 0.62, 0.51, -0.27, -0.13, -0.1, 0.15) \\ \mathbf{w}_7 &= (0.43, -0.55, 0.23, -0.2, 0.36, -0.48, -0.27) \end{aligned}$$

(Presvedčíme sa, že vektory $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_7$ sú ortogonálne.)

Vypočítame súradnice vektorov $\tilde{\mathbf{v}}_0^\circ, \tilde{\mathbf{v}}_1^\circ, \tilde{\mathbf{v}}_2^\circ, \tilde{\mathbf{v}}_3^\circ, \tilde{\mathbf{v}}_4^\circ, \tilde{\mathbf{v}}_5^\circ, \tilde{\mathbf{v}}_6^\circ, \tilde{\mathbf{v}}_7^\circ, \tilde{\mathbf{v}}_8^\circ, \tilde{\mathbf{v}}_9^\circ$, v ortogonálnej báze $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_7$ ako koeficienty priemetov vektorov \mathbf{v}° do tejto bázy, teda podľa vzorca

$$\tilde{\mathbf{v}}^\circ = c_1 \cdot \mathbf{w}_1 + c_2 \cdot \mathbf{w}_2 + \dots + c_7 \cdot \mathbf{w}_7$$

$$c_j = \frac{\langle \mathbf{v}^\circ, \mathbf{w}_j \rangle}{\langle \mathbf{w}_j, \mathbf{w}_j \rangle}, \quad j = 1, \dots, 7$$

V prípade, že do bázy zoberieme iba prvý vlastný vektor, vyjadríme vektory ako násobky vektora $\mathbf{w}_1 = (0.6, -0.09, -0.29, 0.15, 0.36, 0.52, 0.35)$, a tak dostaneme najpresnejší jednorozmerný odhad centrovaných vektorov $\mathbf{v}_0^\circ, \dots, \mathbf{v}_9^\circ$.

$$\tilde{\mathbf{v}}_0^\circ = c_{01} \cdot \mathbf{w}_1, \text{ kde } c_{01} = \frac{\langle \mathbf{v}_0^\circ, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} = \frac{\langle \mathbf{v}_0^\circ, \mathbf{w}_1 \rangle}{1} = 0.43$$

$$\tilde{\mathbf{v}}_0^\circ = (0.26, -0.04, -0.12, 0.06, 0.15, 0.22, 0.15)$$

Po pripočítaní vektora stredných hodnôt dostaneme odhad pôvodného vektora \mathbf{v}_0 , reprezentujúceho digitálnu číslicu „0“.

$$\tilde{\mathbf{v}}_0 = \tilde{\mathbf{v}}_0^\circ + \mathbf{m} = (0.86, 0.86, 0.68, 0.76, 0.75, 0.62, 0.85)$$

Pre ostatné vektory rovnakým spôsobom dostaneme hodnoty:

$$\tilde{\mathbf{v}}_1 = (-0.12, 1.01, 1.15, 0.52, 0.17, -0.22, 0.27)$$

$$\tilde{\mathbf{v}}_2 = (0.91, 0.85, 0.65, 0.78, 0.78, 0.67, 0.88)$$

$$\tilde{\mathbf{v}}_3 = (0.54, 0.9, 0.83, 0.69, 0.57, 0.35, 0.67)$$

$$\tilde{\mathbf{v}}_4 = (0.3, 0.95, 0.94, 0.63, 0.42, 0.15, 0.53)$$

$$\tilde{\mathbf{v}}_5 = (0.93, 0.85, 0.64, 0.78, 0.8, 0.69, 0.9)$$

$$\tilde{\mathbf{v}}_6 = (1.15, 0.81, 0.53, 0.84, 0.93, 0.87, 1.02)$$

$$\tilde{\mathbf{v}}_7 = (-0.04, 1, 1.1, 0.54, 0.22, -0.14, 0.33)$$

$$\tilde{\mathbf{v}}_8 = (1.07, 0.83, 0.57, 0.82, 0.88, 0.8, 0.98)$$

$$\tilde{\mathbf{v}}_9 = (0.4, 0.93, 0.9, 0.65, 0.48, 0.22, 0.58)$$

Tým je úloha priemetu do podpriestoru určeného prvým vlastným vektorom ukončená. Každá číslica je v tomto podpriestore určená jediným koeficientom. Keby sme chceli odhadnuté digitálne číslice zakresliť podobne ako na obrázku 6.3, museli by sme použiť odtiene sivej farby.

Teraz urobíme zjednodušenie a hodnoty zaokrúhlime a až potom zakreslíme. Označme $\tilde{\mathbb{M}}_{\mathbf{w}_1}$ maticu zaokrúhlených hodnôt, predstavujúcich priemet na prvý vlastný vektor.

$$\tilde{\mathbb{M}}_{\mathbf{w}_1} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Vyobrazenia číslic, predstavujúce prvý hlavný komponent, sú na obrázku 6.5. Táto aproximácia je dosť nepresná, správne zakreslené sú len 2 čísla z 10, úspešnosť je 20%.

V prípade, že do bázy zoberieme prvé dva vlastné vektory, vyjadríme vektory ako lineárnu kombináciu vektorov \mathbf{w}_1 a \mathbf{w}_2 . Dostaneme najpresnejší dvojrozmerný odhad centrovanych vektorov $\mathbf{v}_0^\circ, \dots, \mathbf{v}_9^\circ$.

$$\tilde{\mathbf{v}}_0^\circ = c_{01} \cdot \mathbf{w}_1 + c_{02} \cdot \mathbf{w}_2, \text{ kde } c_{01} = \frac{\langle \mathbf{v}_0^\circ, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle} = 0.43$$

$$c_{02} = \frac{\langle \mathbf{v}_0^\circ, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle} = -1.2$$



Obr. 6.5. Vyobrazenia číslíc 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 ako prvý hlavný komponent

$$\tilde{\mathbf{v}}_0^\circ = (0.36, -0.16, 0.02, 0.25, -0.07, 0.33, 0.05)$$

Po pripočítaní vektora stredných hodnôt dostaneme odhad pôvodného vektora \mathbf{v}_0 , reprezentujúceho digitálnu číslicu „0“.

$$\tilde{\mathbf{v}}_0 = \tilde{\mathbf{v}}_0^\circ + \mathbf{m} = (0.96, 0.74, 0.82, 0.95, 0.52, 0.73, 0.75)$$

Pre ostatné vektory rovnakým spôsobom dostaneme hodnoty $\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_9$. Vypočítané hodnoty zaokrúhlime a potom zakreslíme.

Označme $\tilde{\mathbb{M}}_{\mathbf{w}_1, \mathbf{w}_2}$ maticu zaokrúhlených hodnôt, predstavujúcich priemet do podpriestoru určeného prvými dvomi vlastnými vektormi.

$$\tilde{\mathbb{M}}_{\mathbf{w}_1, \mathbf{w}_2} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Vyobrazenia číslíc, vyjadrených ako lineárna kombinácia prvých dvoch vlastných vektorov sú na obrázku 6.6. Táto aproximácia je trochu presnejšia, správne zakreslených je 5 čísel z 10, úspešnosť je 50%.



Obr. 6.6: Vyobrazenia číslíc 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 ako lineárna kombinácia prvých dvoch vlastných vektorov

Matica zaokrúhlených hodnôt, predstavujúcich priemet do podpriestoru určeného vektormi \mathbf{w}_1 , \mathbf{w}_2 a \mathbf{w}_3 je matica $\tilde{\mathbb{M}}_{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3}$. Táto matica sa od matice pôvodných

dát líši len na jedinom mieste.

$$\tilde{\mathbb{M}}_{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Vyobrazenia číslíc, vyjadrených ako lineárna kombinácia vektorov \mathbf{w}_1 , \mathbf{w}_2 a \mathbf{w}_3 sú na obrázku 6.7. Táto aproximácia je znovu presnejšia, správne zakreslených je 9 čísel z 10, úspešnosť je 90%.



Obr. 6.7: Vyobrazenia číslíc 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 ako lineárna kombinácia \mathbf{w}_1 , \mathbf{w}_2 a \mathbf{w}_3

Matica zaokrúhlených hodnôt, predstavujúcich priemet do podpriestoru určeného vektormi vektorov \mathbf{w}_1 , \mathbf{w}_2 , \mathbf{w}_3 a \mathbf{w}_4 je zhodná s maticou pôvodných hodnôt \mathbb{M} .

Vyobrazenia číslíc, vyjadrených ako lineárna kombinácia štyroch vlastných vektorov sú na obrázku 6.8. Táto aproximácia je zhodná s pôvodným obrázkom 6.3.



Obr. 6.8: Vyobrazenia číslíc 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 ako lineárna kombinácia vektorov \mathbf{w}_1 , \mathbf{w}_2 , \mathbf{w}_3 a \mathbf{w}_4

To znamená, že je postačujúce vyjadriť vektory $\mathbf{v}_0, \dots, \mathbf{v}_9$ pomocou 4 koeficientov v báze tvorenej z prvých štyroch vlastných vektorov \mathbf{w}_1 , \mathbf{w}_2 , \mathbf{w}_3 a \mathbf{w}_4 . \square

Príklad 6.3.2 Vyjadrite v báze vytvorenej z prvých štyroch vlastných vektorov z príkladu 6.3.1 vektory

$$\mathbf{f}_1 = (0, 1, 1, 0, 1, 1, 0)$$

$$\mathbf{f}_2 = (0, 0, 1, 0, 0, 1, 1)$$

$$\mathbf{f}_3 = (1, 0, 0, 1, 0, 0, 1)$$

Je báza z príkladu 6.3.1 vhodná na ich reprezentáciu?

Riešenie:

Vektorom $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$ zodpovedajú vyobrazenia, ktoré sú na obrázku 6.9.



Obr. 6.9. Vyobrazenia vektorov $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$

Použijeme výpočty z príkladu 6.3.1. Vektor stredných hodnôt bude znovu vektor

$$\mathbf{m} = (0.6, 0.9, 0.8, 0.7, 0.6, 0.4, 0.7)$$

Vypočítame koeficienty priemetov centrovanych vektorov $\mathbf{f}_1^\circ, \mathbf{f}_2^\circ, \mathbf{f}_3^\circ$ do podpriestoru generovaného vektormi $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ a \mathbf{w}_4 .

$$\begin{aligned}\tilde{\mathbf{f}}_1^\circ &= c_{11} \cdot \mathbf{w}_1 + c_{12} \cdot \mathbf{w}_2 + c_{13} \cdot \mathbf{w}_3 + c_{14} \cdot \mathbf{w}_4 \\ &= (-0.39, 0.24, 0.13, -0.74, 0.35, 0.47, -0.83)\end{aligned}$$

$$\begin{aligned}\tilde{\mathbf{f}}_2^\circ &= c_{21} \cdot \mathbf{w}_1 + c_{22} \cdot \mathbf{w}_2 + c_{23} \cdot \mathbf{w}_3 + c_{24} \cdot \mathbf{w}_4 \\ &= (0.02, -0.43, -0.05, -0.8, -0.8, 0.25, -0.08)\end{aligned}$$

$$\begin{aligned}\tilde{\mathbf{f}}_3^\circ &= c_{31} \cdot \mathbf{w}_1 + c_{32} \cdot \mathbf{w}_2 + c_{33} \cdot \mathbf{w}_3 + c_{34} \cdot \mathbf{w}_4 \\ &= (0.41, -0.46, -0.21, 0.14, -0.69, -0.19, 0.73)\end{aligned}$$

Po pripočítaní vektora \mathbf{m} a zaokrúhlení hodnôt na celé čísla dostávame odhady vektorov $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$.

$$\begin{aligned}\tilde{\mathbf{f}}_1 &= (0, 1, 1, 0, 1, 1, 0) \\ \tilde{\mathbf{f}}_2 &= (1, 0, 1, 0, 0, 1, 1) \\ \tilde{\mathbf{f}}_3 &= (1, 0, 1, 1, 0, 0, 1)\end{aligned}$$

Na rozdiel od predchádzajúceho príkladu, v tomto prípade aproximácia v štvorrozmernom podpriestore nie je taká presná, ako vidíme na obrázku 6.10. Dôvodom je, že vlastné vektory boli vytvorené k matici \mathbb{R} zostavenej pre číslce, teda pre vektory rôzne od $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$. Chyba však nebola príliš veľká. V jednom prípade chyba nastala, v dvoch prípadoch bola chyba na jednom mieste. Navyše chyby spôsobili, že sa aproximované obrazce viac podobajú na číslce, než pôvodné vektory.

Úloha 6.3.1 Zopakujte postup z príkladu 6.3.2 pre rôzne vektory \mathbf{f} . Nájdite taký vektor \mathbf{f} , ktorý bude najviac odlišný od svojho priemetu do 4-rozmerného podpriestoru generovaného vektormi $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ a \mathbf{w}_4 z príkladu 6.3.1. Pridajte vektor \mathbf{f} do matice \mathbb{M} pôvodných hodnôt (použite maticu \mathbb{M} z príkladu 6.3.1). Vypočítajte vlastné vektory pre takto vytvorenú maticu. Zistite, koľko koeficientov stačí na presné vyjadrenie všetkých digitálnych číslc a vektora \mathbf{f} .



Obr. 6.10. Vyobrazenia vektorov $\tilde{\mathbf{f}}_1$, $\tilde{\mathbf{f}}_2$, $\tilde{\mathbf{f}}_3$

Úloha 6.3.2 Vyjadrite písmená slovenskej abecedy ako obrázky vytvorené z čiernych a bielych štvorčekov v sieti veľkosti 5×5 . Každé písmeno sa dá vyjadriť ako 25–zložkový vektor, ktorého zložky sú čísla 0 a 1. Nájdite pomocou metódy PCA bázu so 6 vektormi, v ktorej budú jednotlivé písmená najlepšie aproximované (spomedzi všetkých 6–rozmerných podpriestorov 25–rozmerného priestoru).

Metódy, ktoré sme uviedli v tejto kapitole, spočívajú vo vyjadrení viacrozmerých dát ako lineárnych kombinácií vlastných vektorov matice $\mathbb{R} = \mathbb{M}^{\circ T} \mathbb{M}^{\circ}$, ktorú vypočítame z matice nameraných hodnôt \mathbb{M} . Namiesto nameraných hodnôt matice \mathbb{M} môžeme uvažovať o všetkých možných hodnotách a ich teoretických pravdepodobnostiach. V tom prípade nahradíme maticu dát náhodným procesom. Analýze náhodných procesov sa budeme venovať v ďalšej podkapitole.

6.4 Náhodné procesy

Analýza hlavných komponentov je metóda, ktorá pracuje s konkrétnymi dátami. V teoretickom modeli dáta nahradíme hodnotami odhadnutými pomocou pravdepodobnosti výskytu týchto hodnôt. Teoretická metóda zodpovedajúca PCA sa nazýva Karhunen-Loèveova transformácia. Na odvodenie tejto metódy sú potrebné pojmy a tvrdenia z teórie pravdepodobnosti a náhodných procesov, ktoré tu uvedieme bez dôkazov.

6.4.1 Pojmy a tvrdenia z teórie pravdepodobnosti

V teórii pravdepodobnosti je vybudovaný aparát, ktorý prevádza reálne udalosti do reči čísel, množín, ďalších matematických objektov a vzťahov medzi nimi. Namerané hodnoty sú reprezentované teoretickými hodnotami, sú predpovedané pravdepodobnosti uskutočnenia jednotlivých udalostí pri opakovaní pokusu s rovnakými vstupnými hodnotami. Ak pri rovnakých vstupných podmienkach nastane vždy rovnaký výstup, ide o deterministický pokus. Ak pri rovnakých vstupných podmienkach môžu nastať rôzne výsledky, ide o náhodný pokus.

Najjednoduchší model na popis náhodného pokusu pozostáva z množiny elementárnych udalostí Ω a pravdepodobnostnej funkcie $P : \Omega \rightarrow R$, ktorá každej udalosti $\omega \in \Omega$ priradí pravdepodobnosť tejto udalosti $P(\omega) \in \langle 0, 1 \rangle$. Vo všeobecnosti sú náhodné udalosti vyjadrené ako podmnožiny $A \subseteq \Omega$, teda množiny elementárnych udalostí.

Pravdepodobnostnú funkciu môžeme prirodzene rozšíriť na všetky náhodné udalosti nasledovným spôsobom:

$$P(A) = \sum_{\omega \in A} P(\omega).$$

Dvojicu (Ω, P) nazývame **pravdepodobnostný priestor**.

Pre náhodnú udalosť $A \subseteq \Omega$ definujeme opačnú udalosť \bar{A} ako udalosť, ktorá zodpovedá doplnku množiny A do množiny Ω .

Tvrdenie 6.4.1 (Vlastnosti pravdepodobnosti) *Nech A, B sú náhodné udalosti a nech $P(A)$ a $P(B)$ sú pravdepodobnosti týchto udalostí, potom*

$$P(\emptyset) = 0, P(\Omega) = 1$$

$$0 \leq P(A) \leq 1$$

$$\text{ak } A \cap B = \emptyset, \text{ potom } P(A \cup B) = P(A) + P(B)$$

$$P(\bar{A}) = 1 - P(A)$$

$$\text{ak je množina } A \text{ podmnožinou množiny } B \text{ potom } P(A) \leq P(B)$$

Príklad 6.4.1 *Pri náhodnom pokuse spočívajúcom v hádzaní kockou môžeme za náhodnú udalosť považovať to, že padne párne číslo. Množina reprezentujúca túto udalosť je $A = \{2, 4, 6\}$. Predpokladajme, že všetky elementárne udalosti $\omega \in \{1, 2, 3, 4, 5, 6\}$*

majú rovnakú pravdepodobnosť uskutočnenia $P(\omega) = \frac{1}{6}$.
Pre pravdepodobnosť hodenia párneho čísla bude

$$P(A) = \sum_{\omega \in A} P(\omega) = \frac{3}{6} = \frac{1}{2}$$

Náhodná premenná je zobrazenie $\mathbb{X} : \Omega \rightarrow R$, ktoré každej elementárnej udalosti ω priradí reálnu hodnotu $\mathbb{X}(\omega)$.

Ku náhodnej premennej \mathbb{X} a ku podmnožine reálnych čísel $S \subseteq R$ označíme

$$\mathbb{X}^{-1}(S) = \{\omega \in \Omega; \mathbb{X}(\omega) \in S\}$$

vzor množiny S v zobrazení \mathbb{X} . Týmto spôsobom môžeme pravdepodobnosť udalosti $\mathbb{X}(\omega)$, ktorá sa zobrazí do množiny S definovať vzťahom

$$P(\mathbb{X} \in S) = P(\mathbb{X}^{-1}(S))$$

Ak je množinou S interval tvaru $(-\infty, x)$ kde $x \in R$, dostaneme funkciu $F : R \rightarrow R$ danú predpisom

$$F(x) = P(\mathbb{X} < x)$$

nazývanú **distribučná funkcia** náhodnej premennej \mathbb{X} . Podobne pre interval (a, b) bude $P(a < \mathbb{X} \leq b)$ atď. (Reálna hodnota a môže byť interpretovaná ako interval $\langle a, a \rangle$.)

Distribučná funkcia náhodnej premennej vyjadruje všetky dôležité vlastnosti tejto premennej. Podľa tejto úvahy, náhodné premenné s rovnakou distribučnou funkciou zvyknú byť považované za rovnaké. **Hustota náhodnej premennej** \mathbb{X} je definovaná ako derivácia jej distribučnej funkcie

$$\varphi(x) = F'(x)$$

Na jednoduchší popis náhodnej premennej sa používajú **číselné charakteristiky** náhodnej premennej. Najdôležitejšie, ktoré budeme používať, sú **stredná hodnota**, **disperzia**, pre dve náhodné premenné budeme používať kovarianciu a koeficient korelácie.

Pre zjednodušenie označme $\Omega = \{\omega_1, \dots, \omega_M\}$ a položíme $P(\omega_k) = p_k$ pre $k = 1, \dots, M$. Čísla x_k nazveme hodnoty náhodnej premennej \mathbb{X} . Podobne pre náhodnú premennú \mathbb{X} položíme $\mathbb{X}(\omega_k) = x_k$. Týmto spôsobom bude náhodná premenná \mathbb{X} reprezentovaná M -rozmerným vektorom $(x_1, \dots, x_M) \in R^M$.

Nech $x_k \in \mathbb{X}(\Omega)$ sú hodnoty náhodnej premennej a p_k ich pravdepodobnosti. **Strednou hodnotou** náhodnej premennej \mathbb{X} nazveme číslo

$$E(\mathbb{X}) = \sum_{x_k \in \mathbb{X}} x_k \cdot p_k$$

Disperziou náhodnej premennej \mathbb{X} nazveme číslo

$$D(\mathbb{X}) = E[(\mathbb{X} - E(\mathbb{X}))^2] = \sum_{x_k \in \mathbb{X}} x_k^2 \cdot p_k - \left(\sum_{x_k \in \mathbb{X}} x_k \cdot p_k \right)^2$$

Kovariancia dvoch náhodných premenných \mathbb{X}, \mathbb{Y} je definovaná ako

$$\text{cov}(\mathbb{X}, \mathbb{Y}) = E\{(\mathbb{X} - E(\mathbb{X}))(\mathbb{Y} - E(\mathbb{Y}))\} = E(\mathbb{X}\mathbb{Y}) - E(\mathbb{X})E(\mathbb{Y})$$

Korelačný koeficient dvoch náhodných premenných \mathbb{X}, \mathbb{Y} je definovaný ako

$$\begin{aligned} \rho(\mathbb{X}, \mathbb{Y}) &= \frac{\text{cov}(\mathbb{X}, \mathbb{Y})}{(\mathbb{X} - E(\mathbb{X}))(\mathbb{Y} - E(\mathbb{Y}))} = \\ &= \frac{E\{(\mathbb{X} - E(\mathbb{X}))(\mathbb{Y} - E(\mathbb{Y}))\}}{\sqrt{D(\mathbb{X})}\sqrt{D(\mathbb{Y})}} = \frac{E(\mathbb{X}\mathbb{Y}) - E(\mathbb{X})E(\mathbb{Y})}{\sqrt{D(\mathbb{X})}D(\mathbb{Y})} \end{aligned}$$

Dve náhodné premenné \mathbb{X}, \mathbb{Y} nazveme **nekorelované**, ak platí

$$\rho(\mathbb{X}, \mathbb{Y}) = 0 \quad \text{alebo} \quad \text{cov}(\mathbb{X}, \mathbb{Y}) = 0.$$

6.5 Náhodné vektory

Procesy, ktorým sme sa venovali v predchádzajúcich kapitolách, boli deterministické procesy. V **deterministickom procese** je pre každý časový okamih t_k určená jediná hodnota f_k .

$$\mathbf{f} = (f_0, f_1, \dots, f_{N-1})$$

V tejto kapitole budeme vyšetrovať vlastnosti náhodných procesov.

V **náhodnom procese** je každému časovému okamihu t_k priradená náhodná premenná $\mathbb{X}(t_k) = \mathbb{X}_{t_k} = \mathbb{X}_k$. Táto náhodná premenná je funkcia $\mathbb{X}_k : \Omega \rightarrow R$. Obvykle sa označuje $f(\omega, t_k) = \mathbb{X}_{t_k}(\omega) = \mathbb{X}_k(\omega)$ a náhodný proces má potom tvar

$$\mathbf{f} = (\mathbb{X}_0(\omega), \mathbb{X}_1(\omega), \dots, \mathbb{X}_{N-1}(\omega)) = (f(\omega, t_0), f(\omega, t_1), \dots, f(\omega, t_{N-1}))$$

Je to N -rozmerný vektor závislý od náhodnej udalosti $\omega \in \Omega$.

Náhodný proces popisuje náhodné udalosti tak, že výsledku $\omega_0 \in \Omega$ náhodného pokusu priradí funkciu času $f_\omega(t_k)$, táto funkcia je jedna **realizácia** náhodného procesu:

$$f_\omega(t_k) = f(\omega_0, t_k), \quad k = 0, 1, \dots, N-1$$

Všetky realizácie náhodného procesu v nejakom fixnom čase $t_k = t^*$, teda **rez náhodného procesu** v čase, tvorí náhodnú premennú:

$$f(\omega, t_k = t^*) = f_{t^*}(\omega), \quad \omega \in \mathcal{S}$$

Náhodný proces môžeme chápať ako vektor, ktorého zložky sú náhodné premenné:

$$\mathbf{f}(\omega, t_k) = (f(\omega, t_0), f(\omega, t_1), \dots, f(\omega, t_{N-1}))$$

Distribučnou funkciou náhodného procesu nazveme funkciu

$$F : R^N \rightarrow R$$

$$F(x_0, x_1, \dots, x_{N-1}) = P(f(\omega, t_0) < x_0, f(\omega, t_1) < x_1, \dots, f(\omega, t_{N-1}) < x_{N-1})$$

Hovoríme, že dva náhodné procesy **sa rovnajú** ak sa rovnajú ich distribučné funkcie. Ak existujú parciálne derivácie distribučnej funkcie podľa všetkých premenných, môžeme zaviesť pojem hustoty náhodného procesu.

Hustotou náhodného procesu nazveme funkciu N premenných

$$\varphi(x_0, x_1, \dots, x_{N-1}) = F'(x_0, x_1, \dots, x_{N-1}) = \left(\frac{\partial F}{\partial x_0}, \frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_{N-1}} \right)$$

Ak je známa distribučná funkcia náhodného procesu, je proces jednoznačne popísaný, poznáme ho celý. Takýto popis je však komplikovaný a preto často stačí, ak je známa **dvojrozmerná distribučná funkcia** procesu:

$$F(x_i, x_j) = P(f(\omega, i) < x_i, f(\omega, j) < x_j), x_i, x_j \in R \quad i, j \in \{0, 1, \dots, N-1\}$$

Popisom procesov, pri ktorom sa používa iba dvojrozmerná distribučná funkcia, sa zaoberá **korelačná teória**.

Základné charakteristiky, ktoré sme uviedli pre náhodnú premennú sú stredná hodnota, disperzia, kovariancia. Tieto pojmy zavedieme aj pre náhodné procesy:

Strednou hodnotou náhodného procesu nazveme deterministický vektor

$$\mathbf{m} = (m_0, m_1, \dots, m_{N-1}), \quad \text{kde} \quad m_k = E(f(\omega, k)), \quad k = 0, 1, \dots, N-1$$

Disperziou náhodného procesu nazveme deterministický vektor

$$\mathbf{d} = (d_0, d_1, \dots, d_{N-1}), \quad \text{kde} \quad d_k = E(f^2(\omega, k)) - m_k^2 \quad k = 0, 1, \dots, N-1$$

Kovariančnou maticou náhodného procesu nazveme maticu

$$\mathbf{C} = [c_{ij}], \quad \text{kde} \quad c_{ij} = E\left((f(\omega, i) - m_i) \overline{(f(\omega, j) - m_j)}\right)$$

Na určenie strednej hodnoty a disperzie stačí použiť jednorozmernú distribučnú funkciu. Na určenie kovariančnej matice je potrebná dvojrozmerná distribučná funkcia.

Centrovaným náhodným procesom sa nazýva proces, ktorého stredná hodnota je rovná nule. Ak nie je proces $\mathbf{f}(\omega, t)$ centrovaný, stačí odčítať jeho strednú hodnotu a dostaneme centrovaný proces $\mathbf{f}^\circ(\omega, t)$

$$\mathbf{f}^\circ(\omega, t) = \mathbf{f}(\omega, t) - \mathbf{m}$$

$$\mathbf{f}^\circ(\omega, t) = (f(\omega, 0) - m_0, f(\omega, 1) - m_1, \dots, f(\omega, N-1) - m_{N-1})$$

Kvôli zjednodušeniu a bez ujmy na všeobecnosti budeme v ďalšom texte študovať len centrované náhodné procesy.

Tvrdenie 6.5.1 *Pre centrovaný náhodný proces majú korelačná a kovariančná matica rovnaký tvar*

$$\mathbf{C} = [c_{ij}] = \mathbf{R} = [r_{ij}] = \left[E\left(f^\circ(\omega, i) \overline{f^\circ(\omega, j)}\right) \right] \quad \text{kde} \quad i, j = 0, 1, \dots, N-1$$

Tvrdenie 6.5.2 *Pre prvky korelačnej matice a pre prvky kovariančnej matice platí*

$$r_{ij} = \overline{r_{ji}}, \quad c_{ij} = \overline{c_{ji}}$$

Úloha 6.5.1 *Ak je proces popísaný strednou hodnotou, disperziou a korelačnou maticou, môžeme ho rovnako dobre popísať len dvomi z týchto charakteristík. Ktorými a prečo?*

6.5.1 Priestor náhodných vektorov

Definujme N -rozmerný vektorový priestor náhodných vektorov. Prvkami sú **vektory**, ktorých jednotlivé zložky sú náhodné premenné:

$$\mathbf{f} = (f(\omega, 0), f(\omega, 1), \dots, f(\omega, k), \dots, f(\omega, N-1))$$

Skaláry v tomto priestore môžu byť buď reálne čísla, alebo reálne náhodné premenné.

Súčtom dvoch nezávislých náhodných vektorov \mathbf{f} a \mathbf{g} nazveme vektor \mathbf{z} , ktorého hustotu dostaneme ako konvolúciu hustôt vektorov \mathbf{f} a \mathbf{g} .

$$\mathbf{z} = \mathbf{f} + \mathbf{g}$$

$$\mathbf{z} \text{ má hustotu } \varphi_{\mathbf{z}}(\mathbf{y}) = \int_{\mathbf{x} \in \mathbb{R}^N} \varphi_{\mathbf{f}}(\mathbf{y} - \mathbf{x}) \varphi_{\mathbf{g}}(\mathbf{x}) d\mathbf{x}$$

Opačným vektorom k vektoru

$$\mathbf{f} = (f(\omega, 0), f(\omega, 1), \dots, f(\omega, N-1))$$

je vektor

$$-\mathbf{f} = (-f(\omega, 0), -f(\omega, 1), \dots, -f(\omega, N-1))$$

Neutrálnym prvkom vektorového priestoru je nulový náhodný vektor

$$\mathbf{0} = (0, 0, \dots, 0)$$

Súčin vektora so skalárom je definovaný

$$c \cdot \mathbf{f} = (c \cdot f(\omega, 0), c \cdot f(\omega, 1), \dots, c \cdot f(\omega, N-1))$$

Bez ujmy na všeobecnosti budeme v ďalšom texte uvažovať iba centrovane náhodné procesy. Skalárny súčin môžeme definovať viacerými spôsobmi, výsledok môže byť náhodný, alebo nenáhodný.

Uvažujme najprv vektorový priestor, v ktorom je definovaný **nenáhodný skalár**.

Skalárny súčin dvoch náhodných vektorov je skalár vektorového priestoru, teda nenáhodné číslo.

$$\langle \mathbf{f}, \mathbf{g} \rangle = E \left(\sum_{n=0}^{N-1} f(\omega, n) \overline{g(\omega, n)} \right)$$

Dva náhodné vektory nazveme **ortogonálne**, ak je ich skalárny súčin rovný 0:

$$\mathbf{f} \perp \mathbf{g} \Leftrightarrow \langle \mathbf{f}, \mathbf{g} \rangle = 0$$

Veľkosť náhodného vektora definujeme:

$$\|\mathbf{f}\| = \sqrt{\langle \mathbf{f}, \mathbf{f} \rangle}$$

Vzdialenosť dvoch náhodných vektorov definujeme:

$$d(\mathbf{f}, \mathbf{g}) = \|\mathbf{f} - \mathbf{g}\| = \sqrt{\langle \mathbf{f} - \mathbf{g}, \mathbf{f} - \mathbf{g} \rangle}$$

Rovnosť dvoch náhodných vektorov definujeme:

$$\mathbf{f} = \mathbf{g} \quad \Leftrightarrow \quad \|\mathbf{f} - \mathbf{g}\| = 0$$

V inom vektorovom priestore môžeme definovať skalár, ktorý je **náhodná premenná**. **Skalárny súčin** dvoch náhodných vektorov je skalár vektorového priestoru, teda náhodná premenná:

$$\langle \mathbf{f}, \mathbf{g} \rangle = \sum_{k=0}^{N-1} f(\omega, k) \overline{g(\omega, k)}$$

V tomto priestore sú dva náhodné vektory **ortogonálne**, ak je ich skalárny súčin rovný náhodnej premennej so strednou hodnotou 0:

$$\mathbf{f} \perp \mathbf{g} \quad \Leftrightarrow \quad E\langle \mathbf{f}, \mathbf{g} \rangle = 0$$

Veľkosť náhodného vektora definujeme:

$$\|\mathbf{f}\| = \sqrt{E\langle \mathbf{f}, \mathbf{f} \rangle}$$

Vzdialenosť dvoch náhodných vektorov definujeme:

$$d(\mathbf{f}, \mathbf{g}) = \|\mathbf{f} - \mathbf{g}\| = \sqrt{E\langle \mathbf{f} - \mathbf{g}, \mathbf{f} - \mathbf{g} \rangle}$$

Rovnosť dvoch náhodných vektorov definujeme:

$$\mathbf{f} = \mathbf{g} \quad \Leftrightarrow \quad \|\mathbf{f} - \mathbf{g}\| = 0$$

6.5.2 Vlastnosti priestorov s náhodnými vektormi

Tvrdenie 6.5.3 *Pre náhodné vektory platí Pytagorova veta.*

Dôkaz:

Nech sú \mathbf{f} a \mathbf{g} navzájom kolmé náhodné vektory, potom

$$\begin{aligned} \|\mathbf{f} + \mathbf{g}\|^2 &= E \left(\sum_{n=0}^{N-1} (f(\omega, n) + g(\omega, n)) \overline{(f(\omega, n) + g(\omega, n))} \right) = \\ &= E \left(\sum_{n=0}^{N-1} f(\omega, n) \overline{f(\omega, n)} \right) + E \left(\sum_{n=0}^{N-1} g(\omega, n) \overline{g(\omega, n)} \right) = \|\mathbf{f}\|^2 + \|\mathbf{g}\|^2 \end{aligned}$$

Teda platí

$$E \left(\sum_{n=0}^{N-1} |f(\omega, n) + g(\omega, n)|^2 \right) = E \left(\sum_{n=0}^{N-1} |f(\omega, n)|^2 \right) + E \left(\sum_{n=0}^{N-1} |g(\omega, n)|^2 \right)$$

To znamená, že stredná energia súčtu dvoch navzájom kolmých procesov sa rovná súčtu stredných energií týchto procesov \square

Podmienku kolmosti môžeme nahradiť podmienkou nekorelovanosti.

Tvrdenie 6.5.4 *Stredná energia súčtu dvoch navzájom nekorelovaných procesov sa rovná súčtu stredných energií týchto procesov.*

Systém vektorov $\mathcal{B} = \{\mathbf{b}_n, n = 0, 1, \dots, N-1\}$ s N prvkami tvaru

$$\mathbf{b}_n = (b_n(\omega, 0), b_n(\omega, 1), \dots, b_n(\omega, N-1))$$

nazveme **bázou** vektorového priestoru náhodných procesov, ak sa každý náhodný proces \mathbf{f} vektorového priestoru dá jednoznačne napísať ako lineárna kombinácia vektorov bázy:

$$\mathbf{f} = \sum_{n=0}^{N-1} c_n \mathbf{b}_n$$

Tvrdenie 6.5.5 *Ak pre každé dva vektory $\mathbf{b}_n, \mathbf{b}_m$ $n \neq m$ bázy \mathcal{B} , platí, že sú ortogonálne*

$$E(\mathbf{b}_n, \mathbf{b}_m) = 0$$

potom sú aj nekorelované a pre koeficienty c_n rozkladu $\mathbf{f} = \sum_{n=0}^{N-1} c_n \mathbf{b}_n$ platí:

a) v prípade, že skaláry sú deterministické (reálne, alebo komplexné čísla)

$$c_n = \frac{\langle \mathbf{f}, \mathbf{b}_n \rangle}{\langle \mathbf{b}_n, \mathbf{b}_n \rangle} = \frac{E \left(\sum_{k=0}^{N-1} f(\omega, k) \overline{b_n(\omega, k)} \right)}{E \left(\sum_{k=0}^{N-1} |b_n(\omega, k)|^2 \right)} \quad n = 0, 1, \dots, N-1$$

b) v prípade, že skaláry sú náhodné (reálne, alebo komplexné náhodné premenné)

$$c_n(\omega) = \frac{\langle \mathbf{f}, \mathbf{b}_n \rangle}{\langle \mathbf{b}_n, \mathbf{b}_n \rangle} = \frac{\sum_{k=0}^{N-1} f(\omega, k) \overline{b_n(\omega, k)}}{\sum_{k=0}^{N-1} |b_n(\omega, k)|^2} \quad n = 0, 1, \dots, N-1$$

V prípade, že vo vektorovom priestore sú skaláry náhodné čísla, tak sa dá ukázať, že ortogonálnu bázu môžeme nahradiť deterministickou bázou. Výhodou rozkladu do deterministickej bázy je, že do rovnakej bázy je možné rozložiť nielen proces, ale aj každú jeho realizáciu.

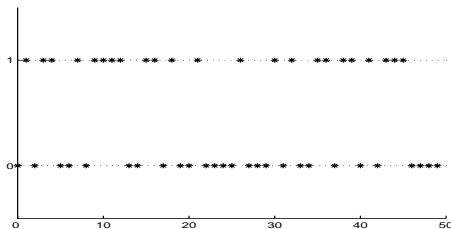
Tvrdenie 6.5.6 (Brown (1970)) *Vzdialenosť medzi procesom \mathbf{f} a jeho rozkladom do deterministickej bázy sa rovná nule práve vtedy, ak je táto báza ortogonálna.*

$$d(\mathbf{f}, \sum_{n=0}^{N-1} c_n(\omega) \cdot \mathbf{b}_n) = 0$$

Ak je daná ortogonálna báza tvorená deterministickými vektormi. Ak zoberieme ku vektorom tejto bázy náhodnú amplitúdu, vznikne náhodný proces.

6.6 Karhunen Loëvov rozklad

Príklad 6.6.1 Uvažujme proces, ktorý vznikne ako výsledok opakovaného hodu mincou. Ak padne hlava, hodnota procesu bude 0, ak padne znak, hodnota procesu bude 1. Na obrázku 6.11 vidíme 50 hodnôt takéhoto procesu. Žiadnou z doteraz uvedených



Obr. 6.11. hod mincou: znak 1, hlava 0

metód analýzy procesov nevieme odhadnúť ďalšiu hodnotu tohto procesu, pretože medzi hodnotami procesu nie je žiadna súvislosť a hodnoty nezávisia od času.

Zopakujme si postup pri analýze procesu \mathbf{f} :

- Odstránenie časovej závislosti (všetky druhy lineárnej regresie, vrátane aproximácie pomocou harmonickej bázy)

$$\tilde{\mathbf{f}}(k) = \sum_{n=0}^{M-1} c_n \varphi_n(k) \quad \text{pre } k = 0, 1, \dots, N-1$$

Dostaneme proces $\mathbf{g} = \mathbf{f} - \tilde{\mathbf{f}}$.

- Odstránenie súvislosti medzi hodnotami (modely kľazových súčtov)

$$\tilde{\mathbf{g}}(t_k) = \sum_{n=0}^{L-1} d_n \mathbf{g}(k-n) \quad \text{pre } k = 0, 1, \dots, N-1$$

Dostaneme proces $\mathbf{z} = \mathbf{g} - \tilde{\mathbf{g}}$.

Proces \mathbf{z} by pri dobre urobenej analýze mal mať charakter podobný ako proces \mathbf{z} príkladu 6.6.1, v ktorom sa nedá predpovedať žiadna hodnota. Ak sme ako prediktor použili model využívajúci jednu predchádzajúcu hodnotu

$$\tilde{f}_{k+1} = c \cdot f_k$$

Koeficient c vypočítame pomocou vzorca $c = \frac{\langle \mathbf{z}, \mathbf{b} \rangle}{\langle \mathbf{b}, \mathbf{b} \rangle}$. Ak je hodnota $c = 0$, tak bola predikcia urobená dobre. Hodnotu c môžeme považovať za koeficient korelácie medzi nasledujúcimi hodnotami. Ak je $c = 0$, tak hodnoty sú nekorelované. Ak je dobre

odhadnutá časová zložka, tak aritmetický priemer hodnôt procesu je rovný nule. Po odstránení časovej závislosti a súvislosti medzi hodnotami ostáva ešte jedna zákonitosť, a to štatistická. Táto zákonitosť sa v príklade 6.6.1 prejaví tak, že pri veľkom počte pokusov padne rovnako veľa 0 a 1.

Ak hodnoty procesu nezávisia od času, tak nemusíme popisovať na časovej osi, ale môžeme výsledky pokusu nanášať v jednom čase. V príklade s mincou dostaneme 7-krát hodnotu 1 a 13-krát hodnotu 0. Dostali sme vlastne 20 realizácií náhodnej premennej. Ak by ostala v procese z nejaká periodická časová závislosť, naskladáme na seba časové periódy, ktoré proces vykazuje a dostaneme náhodný vektor. Tento budeme analyzovať ako náhodný proces.

Tvrdenie 6.6.1 (Karhunen Loëvov rozklad) *Nech \mathbf{f} je centrovanej náhodný vektor z vektorového priestoru s náhodným skalárom. Potom existuje jeho rozklad do deterministickej bázy $\mathcal{B} = \{\mathbf{b}_n, n = 0, 1, \dots, N-1\}$:*

$$\mathbf{f} = \sum_{n=0}^{N-1} c_n(\omega) \mathbf{b}_n$$

taký, že každé pre $n \neq m$ sú vektory $\mathbf{b}_n, \mathbf{b}_m$ na seba kolmé. Navyše sú koeficienty $c_n(\omega), c_m(\omega)$ pre $n \neq m$ nekorelované.

Bázové vektory \mathbf{b}_n sú vlastné vektory kovariančnej matice procesu \mathbf{f} , ktoré zodpovedajú vlastným hodnotám λ_n . Pre λ_n platí

$$\lambda_n = \|\mathbf{b}_n\|^2 E(|c_n(\omega)|^2) = \|\mathbf{b}_n\|^2 \sigma_n^2$$

Dôkaz:

Nech je proces rozložený do bázy podľa tvrdenia vety:

$$\mathbf{f} = \sum_{n=0}^{N-1} c_n(\omega) \mathbf{b}_n$$

a nech platí $\langle \mathbf{b}_n, \mathbf{b}_m \rangle = 0$ a nech $c_n(\omega), c_m(\omega)$ sú nekorelované.

Vypočítame energiu n -tého bázického vektora \mathbf{b}_n a označíme σ_n^2 disperziu náhodného skalára $c_n(\omega)$.

$$(\mathbf{b}_n, \mathbf{b}_n) = \|\mathbf{b}_n\|^2 = \sum_{k=0}^{N-1} |b_n(k)|^2 \quad (6.6.1)$$

Z predpokladu nekorelovanosti $c_n(\omega), c_m(\omega)$ dostaneme

$$E\left(c_n(\omega) \cdot \overline{c_m(\omega)}\right) = \begin{cases} 0 & \text{pre } n \neq m \\ \sigma_n^2 & \text{pre } n = m \end{cases} \quad (6.6.2)$$

Koeficienty $c_n(\omega)$ procesu \mathbf{f} vzhľadom na ortogonálnu deterministickú bázu vypočítame:

$$c_n(\omega) = \frac{\langle \mathbf{f}, \mathbf{b}_n \rangle}{\langle \mathbf{b}_n, \mathbf{b}_n \rangle} = \frac{1}{\|\mathbf{b}_n\|^2} \sum_{k=0}^{N-1} f(\omega, k) \overline{b_n(k)}, \quad n = 0, \dots, N-1$$

Číslo komplexne združené ku $c_m(\omega)$ je

$$\overline{c_m(\omega)} = \frac{1}{\|\mathbf{b}_m\|^2} \sum_{k=0}^{N-1} \overline{f(\omega, k)} b_m(k) \quad m = 0, \dots, N-1$$

Ak dosadíme $c_n(\omega), \overline{c_m(\omega)}$ do vzťahu (6.6.2) dostaneme

$$\begin{aligned} E\left(c_n(\omega) \cdot \overline{c_m(\omega)}\right) &= E\left(\frac{1}{\|\mathbf{b}_n\|^2} \sum_{k=0}^{N-1} f(\omega, k) \overline{b_n(k)} \frac{1}{\|\mathbf{b}_m\|^2} \sum_{l=0}^{N-1} \overline{f(\omega, l)} b_m(l)\right) \\ &= \begin{cases} 0 & \text{pre } n \neq m \\ \sigma_n^2 & \text{pre } n = m \end{cases} \end{aligned}$$

Po ďalších úpravách dostaneme

$$\begin{aligned} \frac{1}{\|\mathbf{b}_n\|^2 \|\mathbf{b}_m\|^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} E\left(f(\omega, k) \overline{f(\omega, l)}\right) \overline{b_n(k)} b_m(l) &= \begin{cases} 0 & \text{pre } n \neq m \\ \sigma_n^2 & \text{pre } n = m \end{cases} \\ \frac{1}{\|\mathbf{b}_n\|^2 \|\mathbf{b}_m\|^2} \sum_{k=0}^{N-1} \overline{b_n(k)} \sum_{l=0}^{N-1} r_{kl} b_m(l) &= \begin{cases} 0 & \text{pre } n \neq m \\ \sigma_n^2 & \text{pre } n = m \end{cases} \end{aligned}$$

nakoniec

$$\sum_{k=0}^{N-1} \overline{b_n(k)} \sum_{l=0}^{N-1} r_{kl} b_m(l) = \begin{cases} 0 & \text{pre } n \neq m \\ \sigma_n^2 \|\mathbf{b}_n\|^2 \|\mathbf{b}_m\|^2 & \text{pre } n = m \end{cases} \quad (6.6.3)$$

pre $n \neq m$ je pravá strana rovná 0, teda ostávajú len $n = m$, pričom platí

$$\|\mathbf{b}_n\|^2 = \sum_{k=0}^{N-1} |b_n(k)|^2 = \sum_{k=0}^{N-1} \overline{b_n(k)} b_n(k)$$

poslednú rovnosť rozšírime výrazom $\|\mathbf{b}_n\|^2 \sigma_n^2$, dostaneme

$$\|\mathbf{b}_n\|^2 \|\mathbf{b}_n\|^2 \sigma_n^2 = \sum_{k=0}^{N-1} \overline{b_n(k)} \|\mathbf{b}_n\|^2 \sigma_n^2 b_n(k)$$

porovnaním poslednej rovnosti s (6.6.3) pre $n = m$ dostaneme

$$\sum_{k=0}^{N-1} \overline{b_n(k)} \sum_{l=0}^{N-1} r_{kl} b_n(l) = \sum_{k=0}^{N-1} \overline{b_n(k)} \|\mathbf{b}_n\|^2 \sigma_n^2 b_n(k) \quad (6.6.4)$$

Rovnosť (6.6.4) je splnená napríklad pre také \mathbf{b}_n a σ_n^2 , pre ktoré platí

$$\sum_{l=0}^{N-1} r_{kl} b_n(l) = \|\mathbf{b}_n\|^2 \sigma_n^2 b_n(k)$$

označme $\lambda_n = \|\mathbf{b}_n\|^2 \sigma_n^2$

$$\sum_{l=0}^{N-1} r_{kl} b_n(l) = \lambda_n b_n(k)$$

alebo ak $\mathbf{b}_n = (b_n(0), b_n(1), \dots, b_n(N-1))$ a $\mathbf{R} = [r_{k,l}]$

$$\mathbf{R}\mathbf{b}_n = \lambda_n \mathbf{b}_n$$

Pre nekorelované $c_n(\omega)$ dostaneme, že bázičné vektory \mathbf{b}_n budú vlastné vektory kovariančnej matice \mathbf{R} náhodného procesu \mathbf{f} .

Vlastné čísla λ_n kovariančnej matice \mathbf{R} určujú strednú energiu n -tej zložky $c_n(\omega)\mathbf{b}_n$ rozkladu náhodného procesu \mathbf{f} .

$$\lambda_n = \|\mathbf{b}_n\|^2 \sigma_n^2 = \langle \mathbf{b}_n, \mathbf{b}_n \rangle E(|c_n(\omega)|^2)$$

Ak zoradíme vlastné vektory podľa veľkosti vlastných čísel, ktoré im prislúchajú

$$\lambda_0 \geq \lambda_1 \geq \dots \geq \lambda_{N-1}$$

najlepšou projekciou do M -rozmerného podpriestoru bude proces

$$\tilde{\mathbf{f}} = \sum_{n=0}^{M-1} c_n(\omega) \mathbf{b}_n$$

a chyba aproximácie bude

$$\|\mathbf{f} - \tilde{\mathbf{f}}\|^2 = \sum_{n=M}^{N-1} \lambda_n$$

Nekorelovanosť koeficientov $c_n(\omega)$ zabezpečuje, že pre zadané $M < N$ má odhad v Karhunen-Loëvovej báze najmenšiu chybu.

Príklad 6.6.2 V tabuľke 6.9 sú váhy piatich novorodencov a váhy týchto detí v deň ich prvých narodení. Nech je náhodný vektor $\mathbf{f}(\omega)$ modelom pre dáta, predstavujúce

dieťa	váha pri narodení	váha v 1. roku
D1	3	11
D2	2.96	9
D3	3.2	10
D4	3.6	12
D5	3.65	11.3

Tab. 6.9. Váha piatich detí pri narodení a v 1. roku

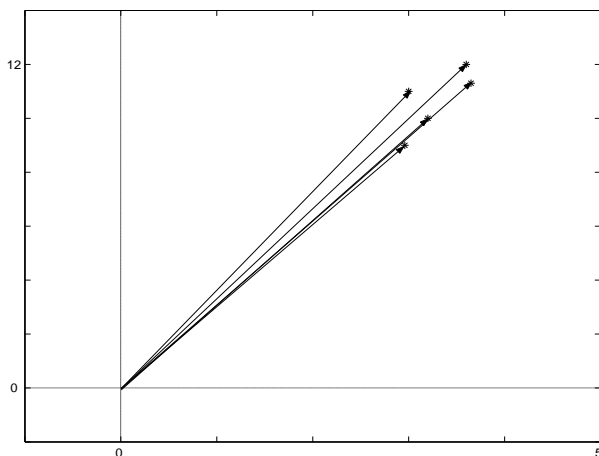
váhy detí pri narodení (prvá zložka vektora) a v deň prvých narodení (druhá zložka vektora). Nájdite ortogonálnu bázu, v ktorej náhodný vektor \mathbf{f} bude mať nekorelované koeficienty $c_0(\omega)$, $c_1(\omega)$.

Riešenie:

Jednotlivé realizácie náhodného procesu $\mathbf{f}(\omega)$ sú reprezentované maticou \mathbf{M} .

$$\mathbf{M} = \begin{pmatrix} 3 & 11 \\ 2.96 & 9 \\ 3.2 & 10 \\ 3.6 & 12 \\ 3.65 & 11.3 \end{pmatrix}$$

Vektory, reprezentujúce jednotlivé realizácie sú zakreslené na obrázku 6.12.

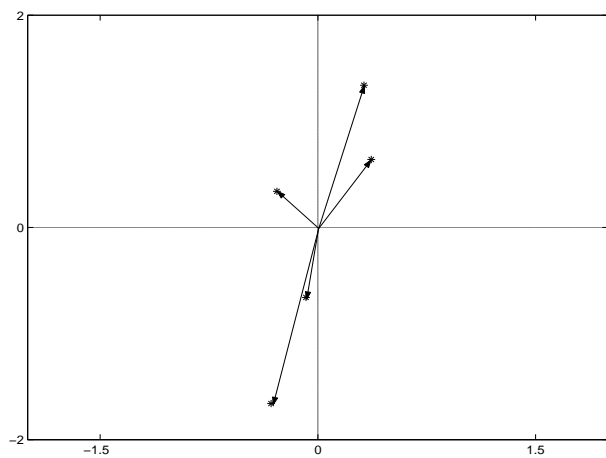
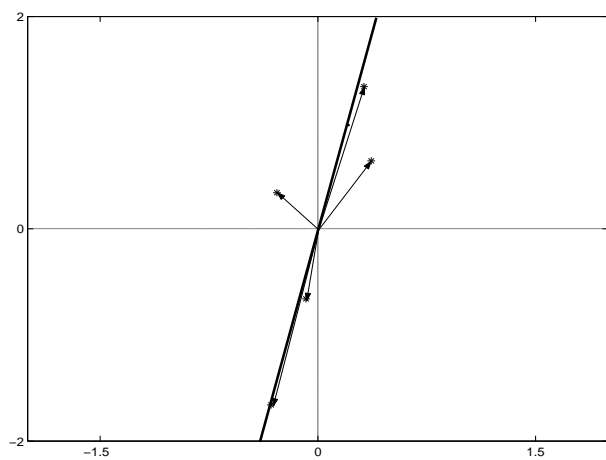


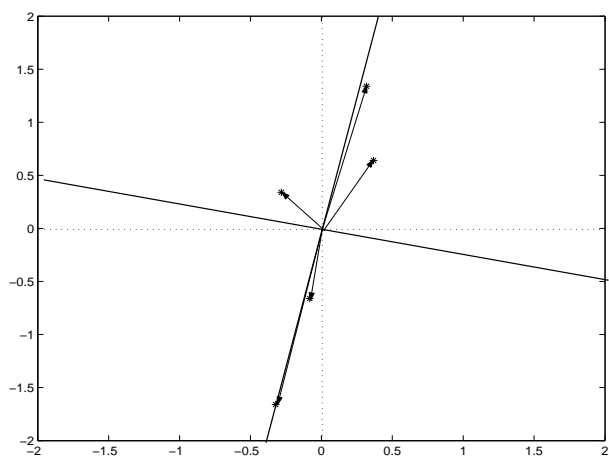
Obr. 6.12. Vektory zodpovedajúce matici \mathbf{M}

Riadky centrovanej matice \mathbf{M}° sú zakreslené ako vektory na obrázku 6.13.

Na obrázku 6.14 je zakreslená priamka predstavujúca bázový vektor v smere najväčšieho rozptylu náhodného vektora $\mathbf{f}(\omega)$.

Na obrázku 6.15 je okrem prvej bázovej priamky nakreslená aj druhá bázová priamka, ktorá je na prvú kolmá. Náhodná premenná $c_1(\omega)$ predstavuje koeficient priemetu vektora $\mathbf{f}(\omega)$ na túto priamku. Koeficient $c_1(\omega)$ je nekorelovaný s koeficientom $c_0(\omega)$.

Obr. 6.13. Vektory zodpovedajúce matici \mathbf{M}° Obr. 6.14. Smer najväčšieho rozptylu náhodného vektora $\mathbf{f}(\omega)$



Obr. 6.15: Dva kolmé smery, v ktorých má náhodný vektor $\mathbf{f}(\omega)$ nekorelované koeficienty

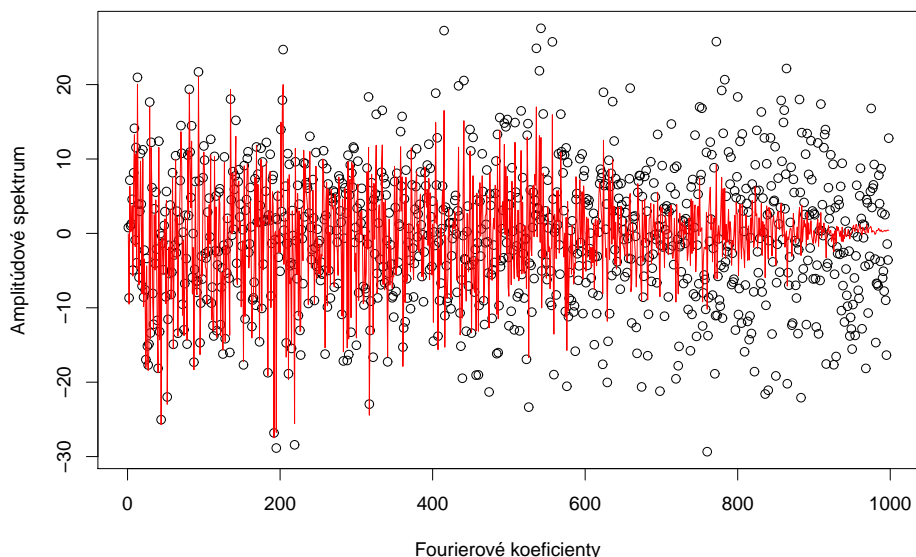
7 Lineárne systémy

Začneme malým experimentom, ktorým ukážeme ako navzájom interagujú kľzavé súčty a diskrétna Fourierova transformácia. Zoberme $N = 2000$ náhodne vybraných reálnych čísel f_1, \dots, f_{2000} . Toto bude náš proces \mathbf{f} . Takýto proces sa nazýva šumom. (Ak sme navyše vybrali f_i navzájom nezávisle, tak sa \mathbf{f} nazýva bielym šumom).

Naň aplikujeme kľzavý súčet $g_k = \frac{f_k + f_{k+1}}{2}$ a vytvoríme proces

$$\mathbf{g} = (g_1, \dots, g_{1999}, 0)$$

Potom si vykreslíme ich amplitúdové spektrum, ako je to ukázané na obrázku 7.1. Na obrázku vidno, že spektrum pôvodného procesu je náhodné pre všetky Fourierove



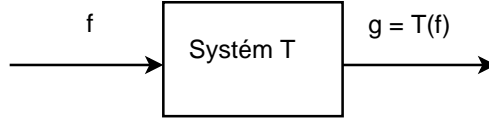
Obr. 7.1: Amplitúdové spektrum procesu \mathbf{f} (čierny kružky) a amplitúdové spektrum procesu \mathbf{g} (červená čiara)

koeficienty. Avšak po aplikácii kľzavého súčtu je vidno jasný trend. Pre nízke frekvencie spektrum je takmer nezmenené, zatiaľ čo vysoké frekvencie sú silno tlmené. Toto

zaujímavé správanie sa využíva napríklad pri tvorbe ekvalizérov do HiFi systémov. Cieľom tejto kapitoly je pochopiť správanie podobných transformácií.

7.1 Lineárny systém

Transformácia jedného procesu na druhý sa nazýva (prenosovým) systémom. Systém \mathcal{T} je schématicky znázornený na obrázku 7.2.



Obr. 7.2. Vstup a výstup zo systému \mathcal{T}

Systém \mathcal{T} sa nazýva **lineárnym systémom**, ak sa správa lineárne, tj.

$$\mathcal{T}(f + g) = \mathcal{T}(f) + \mathcal{T}(g)$$

$$\mathcal{T}(c \cdot f) = c\mathcal{T}(f)$$

$$\mathcal{T}(c_1 \cdot f_1 + c_2 \cdot f_2) = c_1\mathcal{T}(f_1) + c_2\mathcal{T}(f_2)$$

$$\mathcal{T}(c_1 \cdot f_1 + c_2 \cdot f_2 + \dots + c_k \cdot f_k) = c_1\mathcal{T}(f_1) + c_2\mathcal{T}(f_2) + \dots + c_k\mathcal{T}(f_k)$$

Kľzavý súčet $g_k = \frac{f_k + f_{k+1}}{2}$ je typický príklad lineárneho systému.

Úloha 7.1.1 Ktoré z už uvedených vlastností splňajú systémy

- $\mathcal{T}_1(f_1, f_2, f_3) = (f_1, f_2, f_3)$,
- $\mathcal{T}_2(f_1, f_2, f_3) = (f_1 - f_2, f_2 - f_3)$,
- $\mathcal{T}_3(f_1, f_2, f_3) = (\sqrt{f_1 f_2}, \sqrt{f_2^2 + f_3^2})$?

Ak je proces lineárny, tak na jeho popis stačí odozva na jednotkové procesy (vektory). Odozva systému na jednotkový proces \mathbf{e}_n sa nazýva **impulzná charakteristika systému** \mathcal{T} . Označenie $\delta_n = \mathcal{T}(\mathbf{e}_n)$.

Teda platí

$$\tilde{\mathbf{f}} = \mathcal{T}\left(\sum_{k=0}^{N-1} f_k \mathbf{e}_k\right) = \sum_{k=0}^{N-1} f_k \mathcal{T}(\mathbf{e}_k) = \sum_{k=0}^{N-1} f_k \delta_k$$

Príklad 7.1.1 Prenosový systém (lineárna transformácia) \mathcal{T} , je úplne popísaný svojou odozvou na jednotkové impulzy:

$$\delta_0 = \mathcal{T}(\mathbf{e}_0) = \left(\frac{1}{2}, \frac{1}{2}, 0\right),$$

$$\delta_1 = \mathcal{T}(\mathbf{e}_1) = (0, 1, -1),$$

$$\delta_2 = \mathcal{T}(\mathbf{e}_2) = \left(-\frac{1}{2}, 0, \frac{1}{2}\right)$$

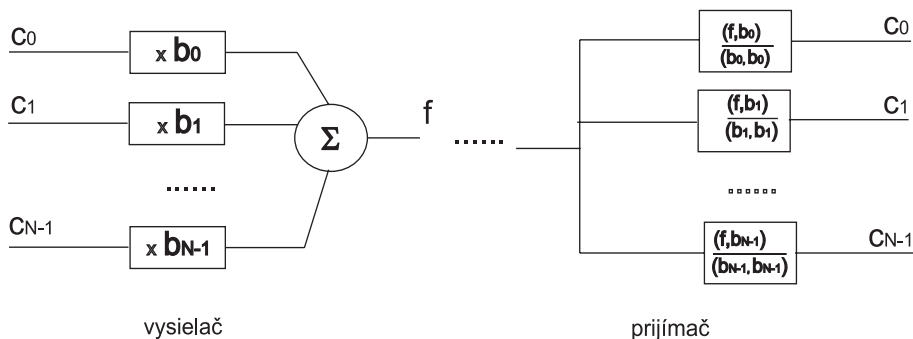
Proces $\mathbf{f} = (1, 2, -1)$ sa prenosovým systémom zobrazí na:

$$\begin{aligned}\tilde{\mathbf{f}} = \mathcal{T}(\mathbf{f}) &= \mathcal{T}(1 \cdot \mathbf{e}_0 + 2 \cdot \mathbf{e}_1 - 1 \cdot \mathbf{e}_2) = 1 \cdot \mathcal{T}(\mathbf{e}_0) + 2 \cdot \mathcal{T}(\mathbf{e}_1) - 1 \cdot \mathcal{T}(\mathbf{e}_2) = \\ &= 1 \cdot \delta_1 + 2 \cdot \delta_1 - 1 \cdot \delta_2 = \left(1, \frac{5}{2}, -\frac{5}{2}\right)\end{aligned}$$

7.2 Modulácia

V snahe čo najlepšie využiť kapacitu prenosového prostredia, používa sa pri prenose dát mechanizmus, ktorý dovoľuje tým istým kanálom prenášať rôzne dáta. Tento mechanizmus sa nazýva **modulácia** a spočíva vo vtláčení určitého príznaku prenášaným dátam tak, aby sa po prenose dali tieto dáta od seba bezpečne oddeliť, demodulovať.

Modulácia teda spočíva v transformácii procesu do tvaru, ktorý dovoľuje prenos. Matematický princíp, ktorý využíva modulácia, je vnorenie procesu do nadpriestoru. Jednotlivé hodnoty, ktoré chceme preniesť, prevedieme na vektory nadpriestoru tak, že ich vynásobíme báзовými vektormi patriacimi nejakej báze vektorového priestoru. Vynásobené vektory sčítame a takto skonštruovaný vektor (proces) prenesieme spoločným komunikačným prostredím. Po prenose, na strane prijímača proces demodulujeme tak, že urobíme jeho priemety do jednorozmerných podpriestorov generovaných jednotlivými báзовými vektormi. Schéma postupu pri modulácii a demodulácii je nakreslená na obrázku 7.3.



Obr. 7.3. Schématický náčrt princípu modulácie a demodulácie

Príklad 7.2.1 Na vstupe A je jedna hodnota z množiny $\{0, 1\}$. Na vstupe B je jedna hodnota z množiny $\{0, 1\}$. Popíšte a zakreslite priebeh modulácie a demodulácie týchto hodnôt pomocou jednotkovej bázy dvojrozmerného vektorového priestoru.

Riešenie:

Hodnotu c_A z množiny $\{0, 1\}$ na vstupe A vynásobíme vektorom $(1, 0)$. Nech je táto hodnota 1.

Hodnotu c_B množiny $\{0, 1\}$ na vstupe B vynásobíme vektorom $(0, 1)$. Nech je táto

hodnota 0.

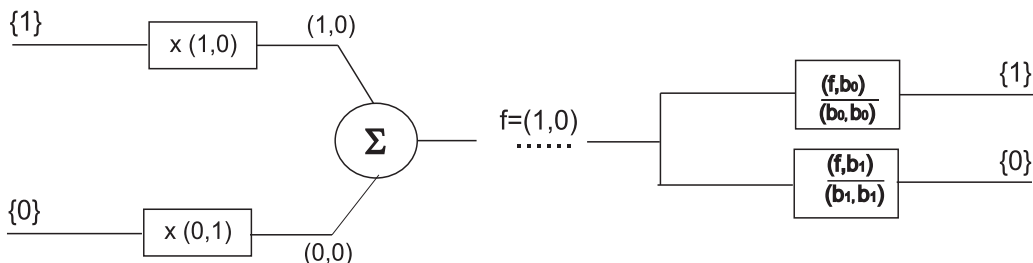
Vektory sčítame a cez spoločné komunikačné prostredie prenesieme niektorý z vektorov $(0, 0)$, $(1, 0)$, $(0, 1)$, $(1, 1)$, v našom prípade to bude vektor $(1, 0)$.

Demodulátor na strane prijímača potom urobí priemet preneseného vektora do dvoch podpriestorov. Do podpriestoru generovaného vektorom $(1, 0)$ pre výstup A a do podpriestoru generovaného vektorom $(0, 1)$ pre výstup B .

$$c_A = \frac{\langle (1, 0), (1, 0) \rangle}{\langle (1, 0), (1, 0) \rangle} = 1$$

$$c_B = \frac{\langle (1, 0), (0, 1) \rangle}{\langle (0, 1), (0, 1) \rangle} = 0$$

Schéma postupu pri modulácii a demodulácii je nakreslená na obrázku 7.4. □



Obr. 7.4. Schématický náčrt modulácie a demodulácie pre dve vstupné hodnoty

Úloha 7.2.1 Pri prenose dvoch hodnôt z množiny $\{0, 1\}$, tak ako je popísaná v príklade 7.2.1, je komunikačným prostredím prenášaný dvojzložkový vektor. Aký je vlastne význam modulácie, keď namiesto dvoch hodnôt 1 a 0 je prenášaný vektor $(1, 0)$? Nestačilo by len povedať, že prvá zložka vektora bude hodnota zo vstupu A a druhá zložka vektora bude hodnota zo vstupu B? Vymyslite nejaké zdôvodnenie!

7.3 Optimálny prijímač

V úlohe modulácie sme riešili technickú úlohu ako cez jedno komunikačné prostredie prenášať naraz väčší počet správ. Teraz budeme riešiť úlohu, ako upraviť dáta pre prenos tak, aby aj v prípade porúch počas prenosu, boli tieto na strane prijímača správne zrekonštruované. V tejto úlohe predpokladáme, že počas prenosu kanálom budú dáta náhodne zmenené. K prenášaným dátam sa pridá šum.

Na vstupe kanála dáta upravíme tak, aby na výstupe bolo možné správne identifikovať, čo bolo vyslané. Úprava na vstupe bude spočívať vo vnorení hodnôt do nadpriestoru tak, aby boli v tomto nadpriestore jednotlivé prenášané prvky od seba dostatočne vzdialené. Aj po prenose potom bude možné zistiť, čo bolo vyslané. Vysielanú hodnotu vynásobíme bázovým vektorom. Počas prenosu sa proces zašumí. Po prenose urobíme priemet procesu do podpriestoru generovaného bázovým vektorom. Priemet porovnáme s množinou prípustných výsledkov a vyberieme najbližší z nich.

Príklad 7.3.1 Na vstupe kanála je jedna hodnota z množiny $\{0, 1\}$. Popíšte a zakreslite spracovanie a optimálny príjem pomocou vektora $(1, 1, 1)$.

Riešenie:

Hodnotu c z množiny $\{0, 1\}$ na vstupe vynásobíme vektorom $\mathbf{b} = (1, 1, 1)$. Ak je táto hodnota 1, budeme prenášať vektor $\mathbf{f} = 1 \cdot \mathbf{b} = (1, 1, 1)$, ak je táto hodnota 0, budeme prenášať vektor $\mathbf{f} = 0 \cdot \mathbf{b} = (0, 0, 0)$.

Vektor prenesieme cez komunikačný kanál, kde naň pôsobí šum a preto dostaneme niektorý z vektorov

$$(0, 0, 0), (1, 0, 0), (0, 1, 0), (1, 1, 0), (0, 0, 1), (1, 0, 1), (0, 1, 1), (1, 1, 1)$$

Nech to v našom prípade bude vektor $\mathbf{f}' = (1, 0, 1)$.

Optimálny prijímač potom urobí priemet preneseného vektora do podpriestoru generovaného vektorom $\mathbf{b} = (1, 1, 1)$.

Koeficient priemetu $\tilde{\mathbf{f}}'$ vektora \mathbf{f}' je

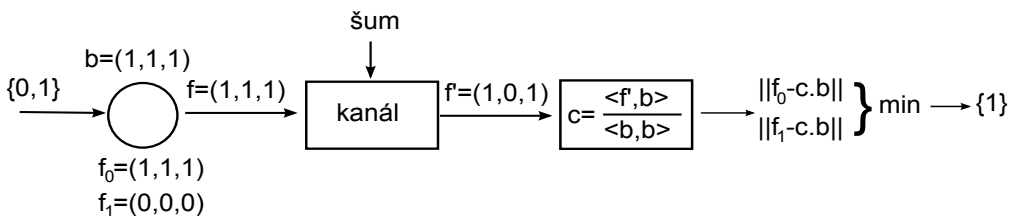
$$c = \frac{\langle \mathbf{f}', \mathbf{b} \rangle}{\langle \mathbf{b}, \mathbf{b} \rangle} = \frac{\langle (1, 0, 1), (1, 1, 1) \rangle}{\langle (1, 1, 1), (1, 1, 1) \rangle} = 0.67$$

Porovnajme teraz výsledný priemet $\tilde{\mathbf{f}}' = (0.67, 0.67, 0.67)$ s dvomi prípustnými možnosťami vyslaného procesu. Mohol to byť $\mathbf{f}_0 = (1, 1, 1)$ alebo $\mathbf{f}_1 = (0, 0, 0)$.

$$\|\tilde{\mathbf{f}}' - \mathbf{f}_0\| = \|(-0.33, -0.33, -0.33)\| = \sqrt{(-0.33)^2 + (-0.33)^2 + (-0.33)^2} = 0.58$$

$$\|\tilde{\mathbf{f}}' - \mathbf{f}_1\| = \|(0.67, 0.67, 0.67)\| = \sqrt{0.67^2 + 0.67^2 + 0.67^2} = 1.16$$

Porovnaním veľkosti odchýliek $\|\tilde{\mathbf{f}}' - \mathbf{f}_0\|$ a $\|\tilde{\mathbf{f}}' - \mathbf{f}_1\|$ usúdime, že pravdepodobne bol vyslaný vektor \mathbf{f}_0 , a teda pôvodná vyslaná hodnota bola 1. Schéma použitia takéhoto optimálneho prijímača je nakreslená na obrázku 7.5. \square



Obr. 7.5. Schéma práce optimálneho prijímača

Úloha 7.3.1 Navrhните optimálny prijímač pre prenos 8 rôznych hodnôt. Koľko rozmerný vektorový priestor treba použiť?

Poznámka:

Optimálny prijímač a samoopravné kódy sú dve technológie, ktoré využívajú rovnaký princíp vnorenia vektora, predstavujúceho prenášané slovo, do nadpriestoru. Komunikačným prostredím sú potom prenášané vektory, ktorých vzájomná vzdialenosť je dostatočne veľká na to, aby ich bolo možné aj po chybnom prenose správne vyhodnotiť. To znamená nahradiť najbližšími vektormi pôvodného priestoru. Optimálny prijímač a samoopravné kódy teda používajú rovnaký princíp v rôznych vektorových priestoroch, s rôznymi metrikami.

7.4 Lineárne časovo invariantné transformácie

Významnou skupinou lineárnych transformácií (systémov) sú **transformácie časovo invariantné** (niekedy nazývané aj stacionárne transformácie), teda transformácie, ktoré nemenia svoje charakteristiky v čase.

Presnejšie, odozva na vstup posunutý v čase, je výstup posunutý v čase:

$$\mathcal{T}(\mathbf{e}_0) = (a_0, a_1, \dots, a_{N-1}); \quad \mathbf{e}_k = \mathbf{e}_{0+k} \Rightarrow$$

$$\mathcal{T}(\mathbf{e}_k) = \mathcal{T}(\mathbf{e}_{0+k}) = (a_{0+k}, a_{1+k}, \dots, a_{N-1+k}) = (a_k, a_{k+1}, \dots, a_{k-1})$$

Teda v \mathbf{e}_0 , resp. v $\mathcal{T}(\mathbf{e}_0)$ robíme kruhový posun, preto na úplný popis časovo invariantnej transformácie stačí, ak poznáme hodnotu $\mathcal{T}(\mathbf{e}_0)$.

Príklad:

Nech impulzná odozva na jednotkový impulz $\mathbf{e}_0 = (1, 0, 0, 0)$ je

$$\mathbf{r} = \mathcal{T}(\mathbf{e}_0) = (1, 1, -1/2, 0)$$

Aká bude odpoveď lineárneho časovo invariantného systému (transformácie, popísanej odozvou \mathbf{r} na vstupný proces $\mathbf{f} = (1, 2, -1, 3)$?

$$\mathbf{e}_0 = (1, 0, 0, 0), \quad \mathcal{T}(\mathbf{e}_0) = (1, 1, -1/2, 0)$$

$$\mathbf{e}_1 = (0, 1, 0, 0), \quad \mathcal{T}(\mathbf{e}_1) = (0, 1, 1, -1/2)$$

$$\mathbf{e}_2 = (0, 0, 1, 0), \quad \mathcal{T}(\mathbf{e}_2) = (-1/2, 0, 1, 1)$$

$$\mathbf{e}_3 = (0, 0, 0, 1), \quad \mathcal{T}(\mathbf{e}_3) = (1, -1/2, 0, 1)$$

$$\begin{aligned} \mathcal{T}(\mathbf{f}) &= \mathcal{T}(1, 2, -1, 3) = 1 \cdot (1, 1, -1/2, 0) \\ &\quad + 2 \cdot (0, 1, 1, -1/2) \\ &\quad - 1 \cdot (-1/2, 0, 1, 1) \\ &\quad + 3 \cdot (1, -1/2, 0, 1) \\ &= (9/2, 3/2, 1/2, 1) \end{aligned}$$

Vyjadríme si teraz vzorce pre koeficienty $\tilde{\mathbf{f}} = \mathcal{T}(\mathbf{f})$. Ak impulzná odozva na jednotkový impulz \mathbf{e}_0 je proces \mathbf{r}

$$\mathcal{T}(\mathbf{e}_0) = \mathbf{r} = (r_0, \dots, r_{n-1}) \tag{7.4.1}$$

tak pre prvú zložku \tilde{f}_0 vektora $\tilde{\mathbf{f}}$ dostaneme:

$$\begin{aligned}\tilde{f}_0 &= r_0 f_0 + r_1 f_{-1} + r_2 f_{-2} + \cdots + r_{N-1} f_{-N+1} \\ &= r_0 f_0 + r_1 f_{N-1} + \cdots + r_{N-1} f_1\end{aligned}$$

a všeobecne

$$\tilde{f}_k = \sum_{n=0}^{N-1} f_n r_{k-n} \quad (7.4.2)$$

Posledný výraz nazývame **konvolúciou vektorov \mathbf{f}, \mathbf{r}**

$$\tilde{\mathbf{f}} = \mathbf{f} * \mathbf{r}$$

Poznámka (model kľzavého súčtu):

$$\tilde{f}_k = \sum_{n=0}^{M-1} c_n f_{k-n}$$

Vidíme, že vzťah (7.4.2) sa podobá predchádzajúcemu vzťahu. Rozdiel je iba v tom, že pri modeli kľzavého súčtu sú nenulové hodnoty c_n len pre $n \leq M - 1$ (ostatné c_n sú rovné nule).

7.5 Spektrálny popis lineárnych, časovo invariantných transformácií

Ukázali sme, že ak \mathcal{T} je lineárna, časovo invariantná transformácia, tak na popis tejto transformácie stačí tzv. impulzná charakteristika.

Pôvodný proces je definovaný nameranými hodnotami:

$$\mathbf{f} = (f_0, f_1, \dots, f_{N-1})\mathcal{E}$$

Jeho spektrálny popis je daný

$$\mathbf{f} = \sum_{k=0}^{N-1} c_k \mathbf{b}_k = (c_0, c_1, \dots, c_{N-1})\mathcal{B}$$

Proces \mathbf{f}' na výstupe lineárnej transformácie (systému) je popísaný

$$\mathbf{f}' = \sum_{k=0}^{N-1} c'_k \mathbf{b}'_k = (c'_0, c'_1, \dots, c'_{N-1})\mathcal{B}'$$

Predpokladajme, že báza \mathcal{B} aj \mathcal{B}' sú obe ortogonálne. Teda platí

$$\begin{aligned}\langle \mathbf{b}_n, \mathbf{b}_m \rangle &= 0, & \text{pre } n \neq m \text{ a takisto} \\ \langle \mathbf{b}'_n, \mathbf{b}'_m \rangle &= 0, & \text{pre } n \neq m.\end{aligned}$$

Pre koeficienty c'_n dostaneme

$$\begin{aligned} c'_n &= \frac{\langle \mathbf{f}', \mathbf{b}'_n \rangle}{\langle \mathbf{b}'_n, \mathbf{b}'_n \rangle} = \frac{\langle \mathcal{T}(\mathbf{f}), \mathbf{b}'_n \rangle}{\langle \mathbf{b}'_n, \mathbf{b}'_n \rangle} = \frac{\langle \mathcal{T}(\sum_{k=0}^{N-1} c_k \mathbf{b}_k), \mathbf{b}'_n \rangle}{\langle \mathbf{b}'_n, \mathbf{b}'_n \rangle} \\ &= \frac{\sum_{k=0}^{N-1} c_k \langle \mathcal{T}(\mathbf{b}_k), \mathbf{b}'_n \rangle}{\langle \mathbf{b}'_n, \mathbf{b}'_n \rangle} = \sum_{k=0}^{N-1} c_k F_{nk} \end{aligned}$$

Čísla F_{nk} tvoria štvorcovú maticu \mathbf{F} a charakterizujú transformovaný proces \mathbf{f}' v spektrálnej oblasti.

$$\mathbf{c}'^T = \mathbf{F} \mathbf{c}^T$$

Nech teraz bázy \mathcal{B} a \mathcal{B}' sú rovnaké, teda

$$\mathbf{b}'_n = \mathbf{b}_n$$

Platí $\langle \mathbf{b}_n, \mathbf{b}_m \rangle = 0$, pre $n \neq m$ a navyše nech platí $\langle \mathcal{T}(\mathbf{b}_n), \mathbf{b}_m \rangle = 0$, pre $n \neq m$. Dostávame

$$c'_n = \sum_{k=0}^{N-1} c_k F_{nk} = F_{nn} c_n$$

Vektor $\mathbf{F} = (F_0, F_1, \dots, F_{N-1})$, (kde použijeme označenie $F_n = F_{nn}$) nazývame **spektrálny prenos lineárneho systému**.

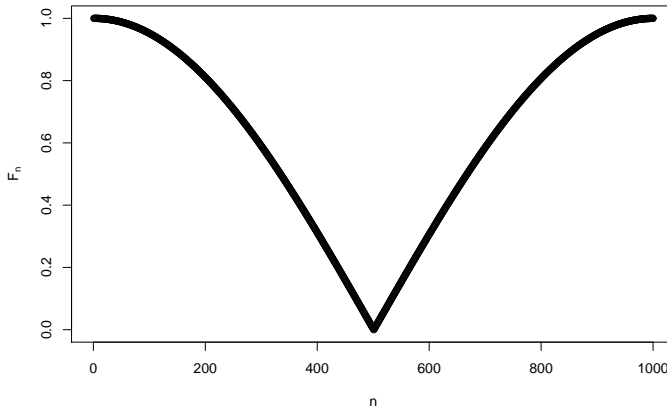
Príklad 7.5.1 Vypočítajme spektrálny prenos lineárneho časovo invariantného systému s impulznou odozvou $\mathbf{r} = (1/2, 1/2, 0, \dots)$ na priestore rozmeru $N = 1000$. Musíme vybrať vhodnú ortogonálnu bázu \mathbf{b}_n . Jednotková báza nie je vhodná, pretože odozva systému nedodržiava ortogonalnosť vektorov

$$\langle \mathcal{T}(\mathbf{e}_0), \mathbf{e}_1 \rangle = \langle (1/2, 1, 2, 0, \dots, 0), (0, 1, 0, \dots, 0) \rangle = 1/2$$

Vhodná je harmonická báza.

$$\begin{aligned} \langle \mathcal{T}(\mathbf{h}_m), \mathbf{h}_n \rangle &= \langle \mathcal{T}(e^{j\frac{2\pi}{N}m}, e^{j\frac{2\pi}{N}2m}, \dots), \mathbf{h}_n \rangle \\ &= \langle (0.5e^{j\frac{2\pi}{N}m} + 0.5e^{j\frac{2\pi}{N}2m}, 0.5e^{j\frac{2\pi}{N}2m} + 0.5e^{j\frac{2\pi}{N}3m}, \dots), \mathbf{h}_n \rangle \\ &= \langle 0.5\mathbf{h}_m + 0.5e^{j\frac{2\pi}{N}m}\mathbf{h}_m, \mathbf{h}_n \rangle \\ &= (0.5 + 0.5e^{j\frac{2\pi}{N}m})\langle \mathbf{h}_m, \mathbf{h}_n \rangle \end{aligned}$$

Na obrázku 7.6 je ukázaná veľkosť jednotlivých koeficientov F_n . Tento výsledok vysvetľuje tlmenie vyšších frekvencií, ktoré bolo vidno v úvode kapitoly na obrázku 7.1.



Obr. 7.6: Spektrálny prenos lineárneho časovo invariantného systému s odozvou $(1/2, 1/2, 0, \dots)$

Ak na vstupe lineárneho, časovo invariantného systému je harmonický proces, tak aj na výstupe je harmonický proces (je to jediný proces s touto vlastnosťou):

$$\begin{aligned}\mathbf{b}'_n = \mathbf{b}_n &= (1, e^{-j\frac{2\pi}{N}n}, \dots, e^{-j\frac{2\pi}{N}n \cdot k}, \dots, e^{-j\frac{2\pi}{N}n \cdot (N-1)}) \\ \mathcal{T}(\mathbf{b}_n) &= (F_n 1, F_n e^{-j\frac{2\pi}{N}n}, \dots, F_n e^{-j\frac{2\pi}{N}n \cdot k}, \dots, F_n e^{-j\frac{2\pi}{N}n \cdot (N-1)}) \\ c'_n &= F_n c_n = |F_n| e^{j\varphi_n} c_n\end{aligned}$$

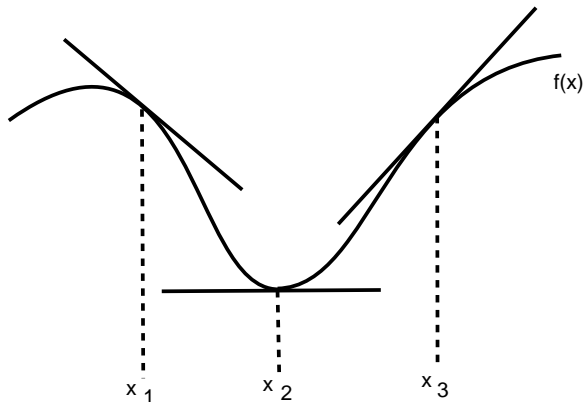
Na to, aby sme zistili $\mathbf{F} = (F_0, F_1, \dots, F_{N-1})$, dávame na vstup postupne všetky vektory harmonickej bázy.

Pomocou spektrálneho popisu teda vieme určiť časový priebeh výstupného procesu:

$$\begin{array}{ccc} \mathbf{f} & \xrightarrow{\mathcal{T}} & \mathbf{f}' \\ DFT \downarrow & & \uparrow DFT^{-1} \\ \mathbf{c} & \xrightarrow{c_n F_n} & \mathbf{c}' \end{array}$$

7.6 Gradientová metóda

Gradientová metóda slúži na hľadanie minima alebo maxima. Začneme najjednoduchším prípadom, keď hľadáme minimum funkciu iba jedného argumentu $f(x)$, ktorá je znázornená na obrázku 7.7. Povedzme, že začíname v bode x_1 . Keďže funkcia v blízkosti x_1 klesá, v hľadaní minima je najlepšie pokračovať doprava. Ak by sme však začali v bode x_3 , v ktorého okolí funkcia rastie, najlepšie je pokračovať doľava. To, či



Obr. 7.7. Hľadanie minima funkcie

funkcia rastie, alebo klesá sa dá určiť z derivácie, ktorá vyjadruje sklon dotýčnice ku grafu funkcie $f(x)$

$$\text{hodnota } f'(x) \quad \begin{cases} > 0 & \text{funkcia rastie, a teda hľadáť vľavo} \\ < 0 & \text{funkcia klesá, a teda hľadáť vpravo} \end{cases}$$

Príčina, prečo hrá určujúcu úlohu derivácia je, že v blízkosti bodu x platí približná aproximácia

$$f(x + \Delta x) \approx f(x) + f'(x)\Delta x$$

V prípade funkcie dvoch premenných $f(x_1, x_2)$ pri hľadaní minima už nemáme na výber len dva smery, ale nekonečne veľa smerov. Keby sme boli na kopci, a hľadáme najnižší bod, prirodzené je vybrať sa najstrmším smerom. Ako však určiť najstrmší smer v prípade funkcie? Použitím Taylorovho rozvoja funkcie a predpokladu, že funkcia je približne lineárna v bode, z ktorého sa plánujeme pohnúť, dostávame

$$f(x_1 + \Delta x_1, x_2 + \Delta x_2) \approx f(x_1, x_2) + \frac{\partial f}{\partial x_1} \Delta x_1 + \frac{\partial f}{\partial x_2} \Delta x_2,$$

alebo inak povedané, prírastok funkcie je (približne) lineárnym systémom smeru pohybu

$$\begin{aligned} f(x_1 + \Delta x_1, x_2 + \Delta x_2) - f(x_1, x_2) &\approx \frac{\partial f}{\partial x_1} \Delta x_1 + \frac{\partial f}{\partial x_2} \Delta x_2 \\ &= \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right) \cdot (\Delta x_1, \Delta x_2) \\ &= \left\| \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right) \right\| \|(\Delta x_1, \Delta x_2)\| \cos \alpha, \end{aligned}$$

kde α je uhol medzi gradientom

$$\text{grad } f := \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right)$$

a vektorom smeru, ktorý má zložky $(\Delta x_1, \Delta x_2)$. Pri hľadaní minima je zjavne najlepšie, aby $\cos \alpha = -1$, a teda vybrali sme sa smerom opačným ku gradientu (vtedy je uhol π). Takže ďalší bod hľadania bude

$$\mathbf{x}' = \mathbf{x} - s \cdot \text{grad } f$$

kde veľkosť kroku s sa môže vybrať na základe experimentálneho testovania.

Naopak pri hľadaní maxima je najlepšie, aby $\cos \alpha = 1$, a teda vybrali sa v smere gradientu $\mathbf{x}' = \mathbf{x} + s \cdot \text{grad } f$.

7.7 Metóda adaptívneho hľadania koeficientov

Príklad 7.7.1 *Aplikujte gradientovú metódu na hľadanie optimálnych koeficientov g_0, g_1 pre predikčný model procesu f*

$$\hat{f}_k = a_k f_{k-1} + b_k f_{k-2}.$$

Riešenie:

Budeme postupne meniť koeficienty a_k, b_k tak, aby sme zmenšili chybu, ktorej sme sa dopustili. Chyba

$$e_k^2 = (f_k - \hat{f}_k)^2,$$

ktorej sa dopustíme je funkciou dvoch premenných $e_k^2 = E(a_k, b_k)$. Menšiu chybu dosiahneme, keď zmeníme a_k, b_k pomocou gradientovej metódy. Gradient je vektor, ktorý má smer najvyššieho nárastu funkcie $e_k^2 = E(a_k, b_k)$, v závislosti od hodnôt a_k, b_k .

$$\begin{aligned} (a_{k+1}, b_{k+1}) &= (a_k, b_k) - s \text{grad } e_k^2 \\ \text{grad } e_k^2 &= \left(\frac{\partial e_k^2}{\partial a_k}, \frac{\partial e_k^2}{\partial b_k} \right) \end{aligned}$$

Pre funkciu E , ktorá je konvexná sa týmto spôsobom dostaneme k minimu. Dostávame:

$$\begin{aligned} \frac{\partial e_k^2}{\partial a_k} &= \frac{\partial}{\partial a_k} (f_k - \hat{f}_k)^2 = \frac{\partial}{\partial a_k} (f_k - (a_k f_{k-1} + b_k f_{k-2}))^2 = \\ &= 2(f_k - (a_k f_{k-1} + b_k f_{k-2}))(-f_{k-1}) = -2e_k f_{k-1} \end{aligned}$$

Nasledujúce a_0 teda položíme (podľa $g_{k+1} = g_k - s \cdot \text{grad } e_k^2$)

$$a_{k+1} := a_k + 2s e_k f_{k-1}$$

Podobne:

$$\begin{aligned} \frac{\partial e_k^2}{\partial b_k} &= \frac{\partial}{\partial b_k} (f_k - \hat{f}_k)^2 = \frac{\partial}{\partial b_k} (f_k - (a_k f_{k-1} + b_k f_{k-2}))^2 = \\ &= 2(f_k - (a_k f_{k-1} + b_k f_{k-2}))(-f_{k-2}) = -2e_k f_{k-2} \end{aligned}$$

Nasledujúce b_{k+1} teda položíme

$$b_{k+1} := b_k + 2s e_k f_{k-2}$$

Celkovo teda dostávame

$$(a_{k+1}, b_{k+1}) = (a_k, b_k) + s \cdot e_k (f_{k-1}, f_{k-2})$$

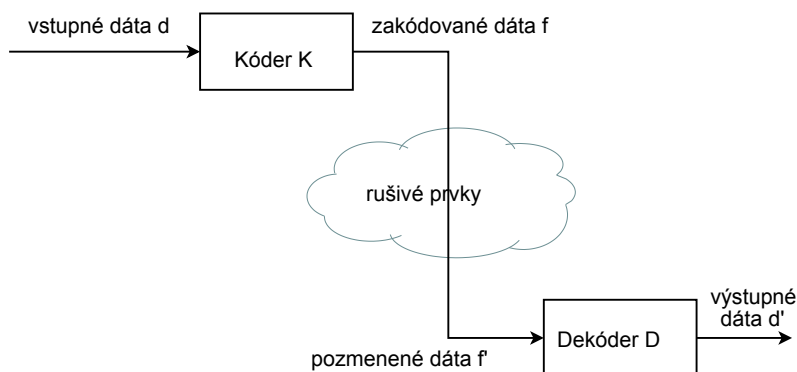
□

Tento postup hľadania optimálneho predikčného modelu je špeciálnym prípadom filtra najmenších stredných štvorcov objaveného B. Widrowom a T. Hoffom [23].

8 Diskrétné lineárne systémy

V informatike, sieťových, mobilných či rádiových komunikáciach sa stretávame so systémami, kde procesy pozostávajú z núl a jednotiek. Systémy, ktoré spracúvajú takéto procesy nazveme diskretnými. Takéto systémy sa používajú napríklad na kompresiu digitálnych dát. Vstupom do systému môže byť binárny súbor alebo nejaký tok dát (povedzme zvuková nahrávka, alebo videosignál z kamery). Výstupom sú tiež binárne dáta, ale vhodne skomprimované, aby sa ušetrilo miesto.

Takpovediac opačný prístup má kódovanie. Pri kódovaní sa vstupné digitálne dáta nejakým spôsobom zakódujú a objem dát väčšinou vzrastie. Kódovanie slúži buď na ochranu dát pred rušivými vplyvmi na komunikačné kanály, alebo na ochranu dát pred neautorizovaným prístupom (šifrovanie).



Pri kódovaní sa často používajú lineárne systémy, ako si teraz ukážeme.

8.1 Detekcia parity

Pri reprezentácii dát v počítači sa používajú N -tice núl a jednotiek. Jednu takúto N -ticu nazveme slovo. Na najjednoduchšie zabezpečenie, ktoré umožňuje detekovanie chyby, sa používa zabezpečenie paritou. Spočíva v pridaní jedného bitu ku každému prenášanému slovu tak, aby celkový počet jednotiek v zabezpečenom slove bol párný. Po prenesení sieťou sa v prenesenom slove spočítajú jednotky. Ak je ich počet nepárny, pri prenose určite nastala chyba.

Príklad 8.1.1 Urobte matematický model, ktorý popisuje zabezpečenie paritou ako transformáciu vektora do nadpriestoru.

Riešenie:

Úlohu vysvetlíme na slovách, ktoré budú pozostávať z N znakov (0 alebo 1). Tieto N -tice tvoria vektorový priestor \mathcal{S} so skalármi 0 a 1. Operácie sčítania a násobenia skalárov 0 a 1 sú definované ako sčítanie a násobenie modulo 2, (tabuľka 8.1).

\oplus_2	0	1
0	0	1
1	1	0

\otimes_2	0	1
0	0	0
1	0	1

Tab. 8.1. Tabuľky sčítania a násobenia modulo 2

Súčet vektorov $\mathbf{f} = (f_0, f_1, \dots, f_{N-1})$ a $\mathbf{g} = (g_0, g_1, \dots, g_{N-1})$ v priestore \mathcal{S} je definovaný predpisom $\mathbf{f} + \mathbf{g} = (f_0 \oplus_2 g_0, f_1 \oplus_2 g_1, \dots, f_{N-1} \oplus_2 g_{N-1})$.

Vektor $\mathbf{f} = (f_0, f_1, \dots, f_{N-1})$ patriaci do N -rozmerného vektorového priestoru \mathcal{S} môžeme zapísať ako lineárnu kombináciu vektorov jednotkovej bázy

$$\mathcal{E} = \{(1, 0, 0, 0, \dots, 0), (0, 1, 0, 0, \dots, 0), \dots, (0, 0, 0, 0, \dots, 1)\}$$

$$\mathbf{f} = f_0(1, 0, 0, 0, \dots, 0) + f_1(0, 1, 0, 0, \dots, 0) + \dots + f_{N-1}(0, 0, 0, 0, \dots, 1)$$

Vytvoríme zabezpečený vektor \mathbf{f}^* ako prvok $N + 1$ -rozmerného vektorového priestoru \mathcal{S}^* . Bude to vektor $\mathbf{f}^* = (f_0, f_1, \dots, f_{N-1}, z)$, kde z je podľa parity číslo 0 alebo 1. Ľubovoľná lineárna kombinácia dvoch zabezpečených kódových slov je zase zabezpečené kódové slovo. Preto zabezpečené kódové slová tvoria N -rozmerný vektorový podpriestor \mathcal{K} vektorového priestoru $\mathcal{S}^*(N + 1$ -rozmerného).

Bázou podpriestoru \mathcal{K} , ktorý pozostáva zo slov kódu (=zabezpečených slov) je množina

$$\mathcal{B}_{\mathcal{K}} = \{(1, 0, 0, 0, \dots, 0, 1), (0, 1, 0, 0, \dots, 0, 1), \dots, (0, 0, 0, 0, \dots, 1, 1)\}$$

Ak vektor

$$\mathbf{f}^* = f_0(1, 0, 0, 0, \dots, 0, 1) + f_1(0, 1, 0, 0, \dots, 0, 1) + \dots + f_{N-1}(0, 0, 0, 0, \dots, 1, 1)$$

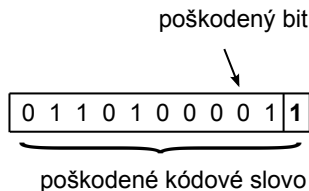
po prenose sieťou nie je prvkom podpriestoru \mathcal{K} , tak hovoríme, že bola detekovaná chyba.

Príklad prenášaného slova a jeho zabezpečenia je na obrázku 8.1.



Obr. 8.1. Zabezpečenie slova paritou

Na obrázku 8.2 vidíme príklad poškodeného slova, teda slova, ktoré nepatrí do podpriestoru kódových slov \mathcal{K} .



Obr. 8.2. Detekcia chyby pri zabezpečení paritou

Úloha 8.1.1 Navrhnite spôsob, ako zistiť, že zvolený vektor \mathbf{f}^* , ktorý je prvkom $N+1$ -rozmerného vektorového priestoru \mathcal{S}^* je vektorom podpriestoru \mathcal{K} všetkých zabezpečených kódových slov. Pri riešení sa môže využiť skutočnosť, že vektor $\mathbf{v}^* = (1, 1, \dots, 1)$ je kolmý na všetky vektory podpriestoru \mathcal{K} .

Pri zabezpečení paritou je možné detekovať chybu, ale nie je to zaručené. Chyba nebude zistená, ak sa zmenia 2, alebo párny počet bitov. Po detekovaní chyby je potrebné poškodené slovo preniesť znovu, pretože táto metóda nedáva odpoveď na to, ktorý bit bol poškodený.

8.2 Ochrana proti reverzii polarity

Pri komunikáciach sa môže stať, že sa zmení fáza vstupu. To spôsobí, že nuly a jednotky sa navzájom vymenia. Toto je vcelku katastrofická chyba, pretože všetky vstupné bity sú nesprávne. Existuje však jednoduchý a spoľahlivý spôsob ochrany.

Povedzme, že na vstupe príde dátový vektor (d_1, d_2, \dots, d_n) . Kódovací systém pošle na výstup vektor pozostávajúci z čiastkových súčtov

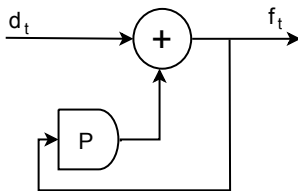
$$(f_1, f_2, \dots) := (0, d_1, d_1 + d_2, d_1 + d_2 + d_3 + \dots)$$

Tieto čiastkové súčty sa dajú počítať aj rekurzívne vzorcami

$$\begin{aligned} f_1 &:= 0 \\ f_2 &:= f_1 + d_1 \\ f_3 &:= f_2 + d_2 \\ &\dots \end{aligned}$$

Schématicky sa takýto výpočet dá znázorniť schémou ako na obrázku 8.3. Tento kodér má jednobitové pamäťové miesto P , kde si ukladá čiastkové súčty f_t . Tie sú potom posielané spolu s novým vstupom do sčítačky, ktorá svoj výstup posiela aj do pamäte P , aj do výstupného kanála.

Dekodér funguje tak, že sa prijaté dáta postupne diferencujú. Keďže $f_{i+1} = f_i + d_i$, vzájomné rozdiely $f_{i+1} - f_i$ nám dajú bity d_i , ak nedôjde k zmene polarity. Zmena polarity však znamená, že ku všetkým bitom sa pripočíta 1, a teda zmena polarity nezmení výsledok dekódovania.



Obr. 8.3. Kodér na opravu reverzie polarity

Príklad 8.2.1 *Povedzme, že sme zakódovali postupnosť $\mathbf{d} = (1, 1, 0, 1, 1)$. Kodér vyšle postupnosť*

$$\mathbf{f} = (0, 1, 1 + 1, 1 + 1 + 0, 1 + 1 + 0 + 1, +1 + 1 + 0 + 1 + 1) = (0, 1, 0, 0, 1, 0).$$

Ak prišla nezmenená, tak diferencovaním získame

$$(1 - 0, 0 - 1, 0 - 0, 1 - 0, 0 - 1) = (1, 1, 0, 1, 1).$$

Ak došlo k reverzii polarity, tak dekodér prijal vektor $(1, 0, 1, 1, 0, 1)$. Diferencovaním dostane

$$(0 - 1, 1 - 0, 1 - 1, 0 - 1, 1 - 0) = (1, 1, 0, 1, 1).$$

Dekodér teda v oboch prípadoch vráti správne dáta.

8.3 Hammingov (7,4) kód

Hamming popísal tento kód vo svojom článku v roku 1950 [8], ktorým otvoril výskum kódovacej teórie. Tento kód funguje tak, že pridáva ku každej štvorici bitov 3 bity navyše (odtiaľ pochádza jeho pomenovanie, sedem zakódovaných bitov zo štyroch dátových bitov). Oproti paritnému kódovaniu je tento kód schopný detekovať nielen jednu chybu, ale aj dve chyby. Navyše, ak nastala chyba iba v jednom bite, tak je schopný určiť, ktorý bit to bol a opraviť ho. Takýto kód sa nazýva *samoopravný*.

Kodér funguje tak, že vstupný vektor zľava vynásobí maticou \mathbf{G} , kde

$$\mathbf{G} = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Všimnime si, že 3, 5, 6 a 7 riadok obsahujú iba jednu jednotku, a teda zakódovaný vektor \mathbf{f} má na pozíciách 3, 5, 6 a 7 pôvodné bity.

Príklad 8.3.1 Zakódujme vektor $\mathbf{d} = (1, 1, 0, 1)$. Tento vektor vynásobíme maticou G a dostaneme

$$\begin{aligned} \mathbf{G} \cdot \mathbf{d} &= \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 \\ 2 \\ 1 \\ 2 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \begin{matrix} \leftarrow d_1 \\ \leftarrow d_2 \\ \leftarrow d_3 \\ \leftarrow d_4 \end{matrix} \end{aligned}$$

Hammingovho kód sa dá použiť dvoma spôsobmi.

- (Samoopravný kód) Ak je pravdepodobnosť chýb jednotlivých bitov nezávislá, potom je omnoho pravdepodobnejšie, že nastane nanajvýš jedna chyba, ako to, že nastanú dve alebo viac chýb. Hammingov kód sa dá použiť na detekciu jednej chyby a opraviť ju.
- (Detekčný kód) Niekedy je takmer rovnako pravdepodobné, že nastane jedna alebo dve chyby vedľa seba. Toto je typický prípad pre výstup z konvolučných kódov, ktoré budú definované v ďalšej časti kapitoly. Vtedy sa Hammingov kód používa na detekciu chyby prenosu.

8.3.1 Samoopravný kód

Prvým krokom pri dekódovaní je výpočet vektora nazývaného **syndrónový vektor**. Tieto *syndrómy* sa vypočítajú ako súčin matice \mathbf{H} s prijatým vektorom \mathbf{f}' .

$$\mathbf{H} := \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Ak dostaneme nulový vektor, tak nenastala žiadna chyba. Ak nám vyjde nenulový vektor, tak chyba nastala v bite, ktorého pozíciu kóduje výsledné dvojkové číslo.

Príklad 8.3.2 Pokračujeme v príklade 8.3.1. Povedzme, že nastala chyba v 6 bite a dostali sme na vstup do dekodéra vektor $\tilde{\mathbf{f}} = (1, 0, 1, 0, 1, 1, 1)$. Vynásobením maticou

\mathbf{H} dostaneme

$$\begin{aligned}\mathbf{H} \cdot \tilde{f} &= \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 4 \\ 3 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\end{aligned}$$

Keďže sme dostali nenulový syndróm, nastala chyba na mieste $110_2 = 6$. Keď túto chybu opravíme, dostaneme vektor $\mathbf{f}' = (1, 0, 1, 0, 1, 0, 1)$, o ktorom už ľahko overíme, že syndrómy $\mathbf{H}\mathbf{f}'$ sú rovné 0. Pôvodné dáta nájdeme na 3., 5., 6. a 7. mieste opraveného slova, a teda dekódovaný výstup je $(1, 1, 0, 1)$.

Na prvý pohľad sa môže zdať neuveriteľné, ako niekto mohol prísť na takéto matice. Ukážeme si teraz, že matica \mathbf{H} je daná jednoznačne, a matica \mathbf{G} sa dá tiež ľahko vybrať. Zhrňme si najprv požiadavky na tieto matice

P1 Kódovanie je dané vzorcom $\mathbf{f} = \mathbf{G}\mathbf{d}$,

P2 Ak nenastala chyba, musia byť syndrómy $\mathbf{H}\mathbf{f}$ rovné 0,

P3 Ak ani jeden bit nebol zmenený, musíme byť schopní dekódovať pôvodné dáta,

P4 Ak práve jeden bit bol pri prenose zmenený, jeho pozícia v dvojkovej sústave je rovná $\mathbf{H}\mathbf{f}$.

Tieto štyri požiadavky zjavne stačia na to, aby kód bol odolný voči jednej nožnej chybe.

Najprv si ukážeme, ako je skonštruovaná matica \mathbf{H} . Ak budeme kódovať nulový vektor \mathbf{d} , tak podľa P1 dostaneme nulový vektor $\mathbf{f} = \mathbf{0}$. Ak nastane chyba na i -tom mieste, tak sme neprijali $\mathbf{0}$ ale vektor \mathbf{e}_i , a podľa P4 dostávame

$$\mathbf{H}\mathbf{e}_i = \text{číslo } i \text{ v dvojkovej sústave.}$$

Lenže podľa pravidiel na násobenie matic, výraz na ľavej strane je rovný i -temu stĺpcu matice \mathbf{H} . Preto matica \mathbf{H} musí mať v i -tom stĺpci zapísané i v dvojkovej sústave, čo aj má. Takže matica \mathbf{H} je jednoznačne určená.

$$\begin{array}{ccccccc} & 1_2 & 10_2 & 11_2 & 100_2 & 101_2 & 110_2 & 111_2 \\ & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ \text{prvky } \mathbf{H} \rightarrow & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{array}$$

Teraz si ukážeme, ako prísť na maticu \mathbf{G} . Z vlastnosti P2 dostávame pre bezchybne prijatý vektor $\mathbf{f} = \mathbf{G}\mathbf{d}$

$$\mathbf{0} = \mathbf{H}\mathbf{f} = \mathbf{H}(\mathbf{G}\mathbf{d}) = (\mathbf{H}\mathbf{G})\mathbf{d}.$$

Toto musí platiť pre všetky vektory, a teda aj pre jednotkové vektory \mathbf{e}_i . Teda musí platiť

$$\mathbf{0} = (\mathbf{H}\mathbf{G})\mathbf{e}_i \leftarrow \text{výsledkom násobenia je } i\text{-ty stĺpec } \mathbf{H}\mathbf{G},$$

a teda každý stĺpec $\mathbf{H}\mathbf{G}$ je nulový a preto $\mathbf{H}\mathbf{G}$ musí byť nulová matica. Toto je zaujímavý fakt, ale ešte sme sa nezaoberali vlastnosťou P3. Tú môžeme dosiahnuť napríklad tak, že všetky pôvodné bity \mathbf{d} sa objavajú vo vektore \mathbf{f} . Povedzme si napríklad, že budú na 1., 2., 3. a 4. mieste vektora \mathbf{f} . To zaručia aj matice vyzerajúce nasledovne:

$$\mathbf{G}' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}, \quad \mathbf{G}'' = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

kde na mieste bodiek môže byť alebo 0 alebo 1. Teraz určíme hodnoty bodiek pre maticu \mathbf{G}' . Treba zabezpečiť rovnosť $\mathbf{H}\mathbf{G}' = \mathbf{0}$. Ak si označíme bodky neznámymi g_{51}, \dots, g_{74} dostávame

$$\mathbf{H}\mathbf{G}' = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ g_{51} & g_{52} & g_{53} & g_{54} \\ g_{61} & g_{62} & g_{63} & g_{64} \\ g_{71} & g_{72} & g_{73} & g_{74} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Toto si môžeme prepísať ako sústavu 12 rovníc o 12 neznámých. Napíšme si prvé tri rovnice zodpovedajúce násobeniu prvým stĺpcom matice \mathbf{G}' .

$$\begin{aligned} 1 + g_{51} + g_{71} &= 0 \\ g_{61} + g_{71} &= 0 \\ g_{51} + g_{61} + g_{71} &= 0 \end{aligned}$$

Z druhej rovnice dostávame $g_{61} = -g_{71} = g_{71}$, keďže v dvojkovej sústave násobenie -1 nemení prvky. Z tretej rovnice dostávame $g_{51} + 2g_{61} = 0$, a teda $g_{51} = 0$. Dosadením do prvej rovnice dostávame $g_{71} = 1$, a teda aj $g_{61} = 1$.

Úloha 8.3.1 *Dopočítajte všetky prvky matice \mathbf{G}' . Na aký vektor nám táto matica zakóduje vektor $(1, 1, 0, 1)$?*

Úloha 8.3.2 *Dopočítajte všetky prvky matice \mathbf{G}'' . Na aký vektor nám táto matica zakóduje vektor $(1, 1, 0, 1)$?*

8.3.2 Detekný kód

Hammingov kód je na rozdiel od paritného kódu schopný detekovať aj dve chyby a to tým, že syndrónový vektor bude nenulový.

Príklad 8.3.3 *Povedzme, že sme namiesto vektora $\mathbf{f} = (1, 0, 1, 0, 1, 0, 1)$ prijali vektor $\mathbf{f}' = (1, 1, 1, 0, 1, 1, 1)$, ktorý sa od neho líši v dvoch bitoch (na 2. a 6. mieste). Vypočítame syndróm*

$$\mathbf{H}\mathbf{f}' = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Keďže nám vyšiel nenulový syndróm, dostávame, že pri prenose nastala chyba alebo v jednom bite alebo v dvoch bitoch.

Všimnime si, že tieto dva prípady nemožno rozlíšiť. Zakódovaním vektora $(1, 1, 1, 1)$ totiž dostaneme

$$\begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

ktorý sa tiež líši od \mathbf{f}' iba na jednom mieste. Preto nie sme schopní chybný vektor dekódovať a rozlíšiť, či na vstupe bol vektor $(1, 1, 0, 1)$ a nastali chyby v dvoch bitoch, alebo na vstupe bol vektor $(1, 1, 1, 1)$ a nastala chyba v jednom bite. Sme schopní iba určiť, že pri prenose nastala nejaká chyba, podobne ako pri paritnom kóde.

Ak nastala chyba na i -tom mieste, tak po vynásobení \mathbf{H} dostaneme

$$\mathbf{H}(\mathbf{f} + \mathbf{e}_i) = \mathbf{H}\mathbf{f} + \mathbf{H}\mathbf{e}_i = 0 + \mathbf{H}\mathbf{e}_i = \mathbf{s}_i$$

kde \mathbf{s}_i je dvojkovo zapísaná pozícia i .

Ak nastali chyby na i -tom a j -tom mieste tak po vynásobení prijatého vektora maticou \mathbf{H} dostaneme

$$\mathbf{H}(\mathbf{f} + \mathbf{e}_i + \mathbf{e}_j) = 0 + \mathbf{s}_i + \mathbf{s}_j = \mathbf{s}_i + \mathbf{s}_j$$

a $\mathbf{s}_i \neq \mathbf{s}_j$ a tak ich súčet v dvojkovej sústave nemôže byť nula.

Všimnime si, že linearita výpočtu syndrónov bola kľúčová na dokázanie schopnosti detekcie dvoch chýb. Ďalej si treba uvedomiť, že schopnosť detekcie chyby bola daná tým, že ľubovoľné dve rôzne kódové slová sa líšili aspoň v 3 bitoch. Počet

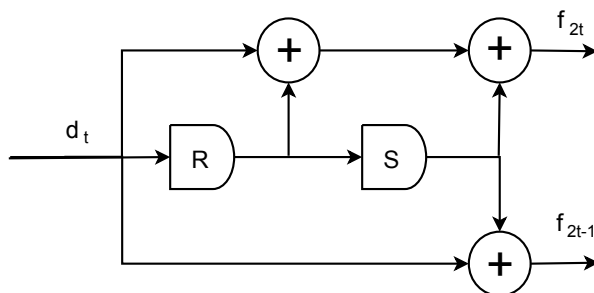
rozdielnych bitov sa nazýva aj **Hammingova vzdialenosť** dvoch slov. Hammingov (7,4) kód funguje ako detekčný kód pre 1 alebo 2 bitové chyby preto, lebo minimálna Hammingova vzdialenosť medzi dvoma rôznymi kódovými slovami je aspoň 3.

Úloha 8.3.3 *Koľko chybových bitov bude schopný detekovať kód, ktorý má minimálnu vzdialenosť medzi rôznymi kódovými slovami rovnú 5? Skúste navrhnúť lineárny kód, ktorý má minimálnu Hammingovu vzdialenosť aspoň 5.*

8.4 Konvolučné kódy

Konvolučný kodér pozostáva z binárnych sčítačiek a pamäťových registrov, podobne ako kodér z kapitoly 8.1. Na rozdiel od blokových kódov, tento kodér spracuje ľubovoľne dlhý reťazec bitov.

Na obrázku 8.4 je znázornený príklad obvodu konvolučného kodéra. Ak si označíme



Obr. 8.4. Nerekurzívny konvolučný kodér s dvoma pamäťovými bitmi R, S .

R_t a S_t hodnoty pamäťových registrov R a S v čase t , tak zo schémy vyplývajú rovnosti:

$$\begin{aligned} R_t &= d_{t-1} \\ S_t &= R_{t-1} = d_{t-2} \\ f_{2t} &= S_t + (d_t + R_t) = d_{t-2} + d_{t-1} + d_t \\ f_{2t-1} &= S_t + d_t = d_{t-2} + d_t \end{aligned} \quad (8.4.1)$$

Z týchto rovníc vidno, že konvolučný kodér je diskretný lineárny systém.

V Exceli, alebo inom tabuľkovom procesore si môžeme na základe rovníc (8.4.1) pripraviť aj tabuľku rôznych stavov v závislosti od vstupných bitov d_t :

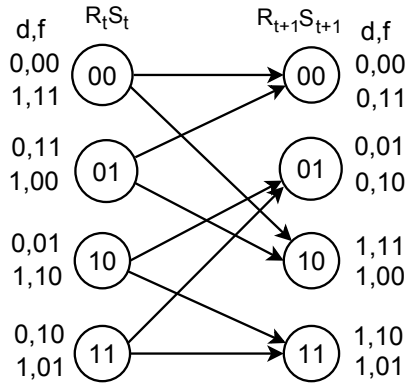
Kódovanie

Rovnice (8.4.1) umožňujú ľahko naprogramovať kodér. Pre lepšie pochopenie činnosti konvolučného kodéra je však lepšie popísať ho pomocou prechodového diagramu, ktorý znázorňuje zmenu stavu pamäťových registrov.

Na ľavej strane sú znázornené zakódované bity f pre každú možnosť vstupného bitu d_t . Podobne pre každý stav na pravej strane $R_{t+1}S_{t+1}$ sú znázornené dve možnosti páru

vstup			registre		registre		výstup	
d_{t-2}	d_{t-1}	d_t	R_t	S_t	R_{t+1}	S_{t+1}	f_{2t}	f_{2t-1}
0	0	0	0	0	0	0	0	0
0	0	1	0	0	1	0	1	1
0	1	0	1	0	0	1	1	0
0	1	1	1	0	1	1	0	1
1	0	0	0	1	0	0	1	1
1	0	1	0	1	1	0	0	0
1	1	0	1	1	0	1	0	1
1	1	1	1	1	1	1	1	0

Tab. 8.2. Výpočtová tabuľka pre konvolučný kodér z obrázka 8.4



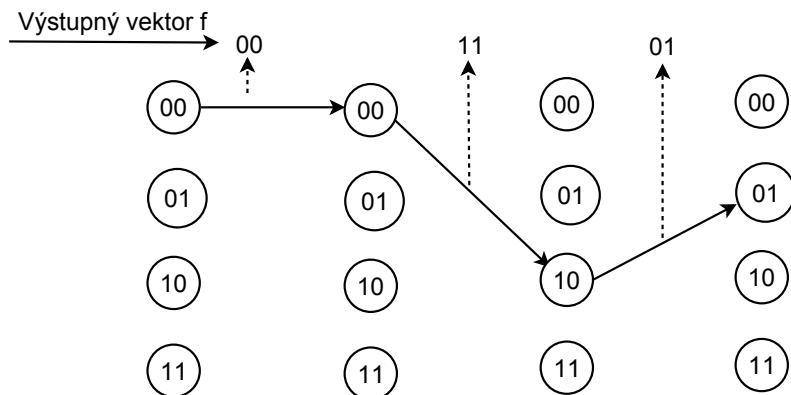
Obr. 8.5. Prechodový diagram konvolučného kodéra z obrázka 8.4

bitov, ktoré kodér vyslal, a bit, ktorý prijal, aby sa dostal do tohto stavu. Všimnime si, že bitové páry f pri každom stave, či už na ľavej strane, alebo na pravej strane sa odlišujú. To je v súlade s predpokladom, že kodér zakóduje dve rôzne postupnosti \mathbf{d}, \mathbf{d}' do dvoch rôznych postupností \mathbf{f}, \mathbf{f}' (za predpokladu, že počiatočné registre boli v rovnakom stave).

Ilustrujme si teraz použitie tohto prechodového diagramu na kódovanie postupnosti bitov $\mathbf{d} = (0, 1, 0)$. Budeme predpokladať, že na začiatku boli registre R a S vynulované. Stavy, ktorými postupne prejde kodér sú znázornené na obrázku 8.6. Tento obrázok sa nazýva *trellis*.

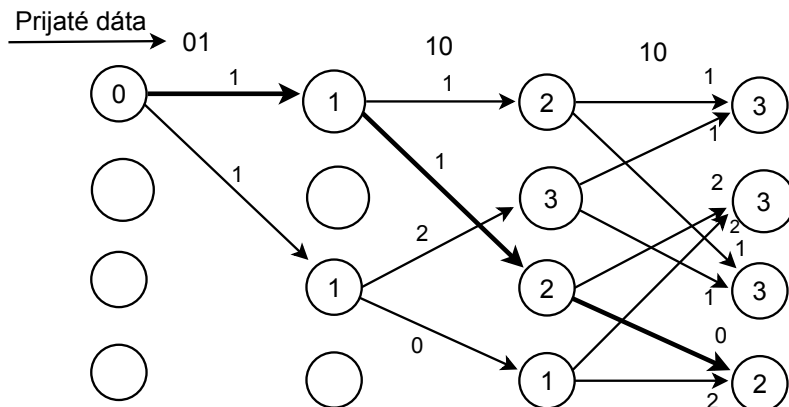
Dekódovanie

Na dekodovanie sa najčastejšie používa Viterbiho algoritmus. Ten sa snaží nájsť takú postupnosť zmien stavov kodéra, ktorá zodpovedá najmenšiemu počtu chýb. Pri dekodovaní sa pre každý stav pamätá jedna postupnosť, ktorá má najmenší počet bitových zmien od prijatých dát. Začne sa od počiatočného stavu $(R, S) = (0, 0)$ a postupne



Obr. 8.6: Postupnosť stavov (R, S) , ktorými kodér prejde pri kódovaní vektora $(0, 1, 0)$

sa rozoberú stavy kodéra po každom prijatom písmene, pričom sa pre každý stav vypočíta najmenší možný počet chýb na to, aby sa kodér dostal do daného stavu. Vyhráva koncový stav s najmenším počtom chýb. Toto je znázornené na obrázku 8.7 pre prijatú postupnosť $\mathbf{f} = (01\ 10\ 10)$.



Obr. 8.7. Vypočítané minimálne počty chýb pre každý stav v trellise

Úloha 8.4.1 Pozrite si schému konvolučného kodéra na obrázku 8.8. Napíšte vzorce pre správanie tohto kodéra a nakreslite príslušný prechodový diagram. Použite Viterbiho dekódovací algoritmus na dekódovanie nasledujúcich postupností bitov

- 01,
- 01 10,

Ak chceme zabezpečiť číslo 124 tak, aby sme ho neporušili, pripíšeme zaň ešte jednu cifru 124x. Číslo x bude také, aby bolo číslo 124x deliteľné siedmimi.

$$124x : 7 = 1246 : 7 = 178 \quad \text{zvyšok} \quad 0 \quad (8.5.2)$$

Číslo $x = 6$ nájdeme tak, že za 124 pripíšeme 9, potom vydelíme $1249 : 7 = 178$ zvyšok 3. Nakoniec odpočítame $1249 - 3$ a dostaneme zabezpečené slovo 1246. Toto číslo je skutočne deliteľné číslom 7 bezo zvyšku.

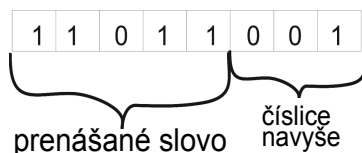
Úloha 8.5.1 *Fungoval by tento postup kódovania, ak by sme pripisovali*

- cifru 2,
- cifru 8?

Ukázkou blokového kódu je cyklický kód z nasledujúceho príkladu, ktorý sa v praxi naozaj používa.

Príklad 8.5.1 (Cyklický kód s generujúcim polynómom, CRC)

Nech slovo, ktoré má byť zakódované je päťica pozostávajúca z núl a jednotiek, napríklad 11011. Za toto slovo pridáme nejaký počet binárnych číslic (v tomto prípade sú to číslice 001) tak, aby polynóm, reprezentovaný celým zabezpečeným slovom bol deliteľný nejakým vopred dohodnutým generujúcim polynómom, napríklad $q(x) = x^3 + x + 1$. Zabezpečené slovo je na obrázku 8.9.



Obr. 8.9. Blokový kód, ktorý zabezpečí slovo 11011 blokom 001

Zabezpečené slovo bude polynóm s koeficientmi 1, 1, 0, 1, 1, 0, 0, 1 pri jednotlivých mocninách x , konkrétne

$$1 \cdot x^7 + 1 \cdot x^6 + 0 \cdot x^5 + 1 \cdot x^4 + 1 \cdot x^3 + 0 \cdot x^2 + 0 \cdot x^1 + 1 \cdot x^0 = x^7 + x^6 + x^4 + x^3 + 1$$

Sčítanie a delenie polynómov definujeme tak, aby koeficienty polynómov boli vždy z množiny $\{0, 1\}$. Pri súčte polynómov budú koeficienty pri zodpovedajúcich mocninách súčtom modulo 2. Pri algoritme delenia rovnako použijeme násobenie a sčítanie modulo 2.

Pre uvedené polynómy potom naozaj platí

$$(x^7 + x^6 + x^4 + x^3 + 1) : (x^3 + x + 1) = x^3 + x^2 + x + 1 \quad (\text{zvyšok } 0)$$

Ako sa vypočítajú hodnoty 001 v zabezpečovacom bloku? Ako sa detekuje chyba v prenesenom slove?

Riešenie:

Prenášanému slovu 11011 zodpovedá polynóm

$$f(x) = 1 \cdot x^4 + 1 \cdot x^3 + 0 \cdot x^2 + 1 \cdot x^1 + 1 \cdot x^0 = x^4 + x^3 + x + 1$$

Podobne ako pri príklade s celými číslami 8.5.1, zistíme zvyšok po delení vopred určeným generujúcim polynómom $q(x) = x^3 + x + 1$. Tento zvyšok potom odpočítame, aby zabezpečený polynóm bol deliteľný polynómom $q(x)$ bezo zvyšku. Prepis do tvaru polynómu navyše umožňuje oddeliť informačnú a zabezpečovaciu časť kódového slova. Vynásobením polynómu $f(x)$ polynómom x^k dostaneme polynóm $f^{\oplus}(x)$, ktorý bude mať nulové koeficienty pri mocninách x^0, x^1, \dots, x^{k-1} a koeficienty pri vyšších mocninách budú zodpovedať koeficientom polynómu $f(x)$.

$$f^{\oplus}(x) = 1 \cdot x^{k+4} + 1 \cdot x^{k+3} + 0 \cdot x^{k+2} + 1 \cdot x^{k+1} + 1 \cdot x^{k+0}$$

Zodpovedajúce slovo polynómu $f^{\oplus}(x)$ je 11011 $\underbrace{00 \dots 0}_k$.

Hodnota k v polynóme x^k závisí od veľkosti zvyškov po delení generujúcim polynómom. V prípade polynómu $q(x) = x^3 + x + 1$ sú zvyšky po delení

polynóm	slovo
0	000
1	001
x	010
$x + 1$	011
x^2	100
$x^2 + x$	110
$x^2 + x + 1$	111

Z tvaru všetkých zvyškov vidíme, že na zabezpečenie stačia tri miesta. Preto polynóm $f(x) = x^4 + x^3 + x + 1$ (teda slovo 11011) vynásobíme x^3 .

Po vynásobení dostaneme polynóm

$$f^{\oplus}(x) = x^7 + x^6 + x^4 + x^3$$

Zodpovedajúce slovo k polynómu $f^{\oplus}(x)$ je 11011000.

Tento polynóm však nie je deliteľný bezo zvyšku generujúcim polynómom $q(x) = x^3 + x + 1$. Aby sme zistili, aké hodnoty má mať zabezpečovací blok, napodobníme postup podľa výpočtu (8.5.2).

$$(x^7 + x^6 + x^4 + x^3) : (x^3 + x + 1) = x^4 + x^3 + x^2 + x + 1 \quad (\text{zvyšok } 1)$$

Zvyšok 1 po delení generujúcim polynómom odpočítame (odpočítat binárne číslo modulo 2 znamená pripočítat toto číslo) od polynómu $f^{\oplus}(x)$ a dostaneme zabezpečený polynóm $f^{\star}(x)$.

$$f^{\star}(x) = f^{\oplus}(x) - 1 = f^{\oplus}(x) \oplus_2 1 = x^7 + x^6 + x^4 + x^3 + 1$$

Polynóm $f^*(x)$ je skutočne bezo zvyšku deliteľný generujúcim polynómom $q(x) = x^3 + x + 1$ a navyše, jeho binárny zápis má informačnú časť oddelenú od zabezpečovacej časti.

Kódové slovo bude obsahovať chybu, ak polynóm $f^*(x)$, prislúchajúci tomuto slovu, nie je deliteľný generujúcim polynómom $q(x) = x^3 + x + 1$. \square

Úloha 8.5.2 Dokážte, že pre ľubovoľné polynómy $f(x)$, $q(x)$ (z priestoru polynómov ako v príklade 8.5.1), polynóm $z(x)$ (zvyšok po delení $f(x) : q(x)$) a číslo k (ktoré udáva stupeň polynómu $q(x)$) platí:

$$q(x) \text{ je deliteľom polynómu } f(x) \cdot x^k + z(x)$$

Úloha 8.5.3 Ako by sa zmenil algoritmus zabezpečovania slova generujúcim polynómom, keby sa na konci slova nepridalo k núl, ale k jednotiek?

Úloha 8.5.4 Pri CRC kódovaní je potrebné deliť polynómy zodpovedajúce kódovým slovám. Slová pozostávajú z 0 a 1, preto by mohli byť považované za zápisy čísel v dvojkovej sústave. Je možné delenie polynómu polynómom nahradiť podielom čísel v dvojkovej sústave, ktoré robí kalkulačka? Ak áno, vysvetlite prečo, ak nie, nájdite príklad, v ktorom nebude výsledok pri týchto spôsoboch delenia rovnaký.

Úloha 8.5.5 Aké slová sa skrývajú za skratkou CRC?

Príklad 8.5.2 Urobte matematický model, ktorý popisuje cyklický kód s generujúcim polynómom ako transformáciu vektora do nadpriestoru.

Riešenie:

Použijeme slová a generujúci polynóm ako v príklade 8.5.1. Naďalej budeme používať súčin a súčet modulo 2. Slová sú prvky vektorového priestoru S_5 všetkých 5-zložkových vektorov zložených z 0 a 1. Tieto slová chceme preniesť nejakým prostredím a predtým zakódovať tak, aby boli deliteľné generujúcim polynómom. Bázu B_5 priestoru S_5 tvoria vektory

$$\mathbf{e}_0 = (1, 0, 0, 0, 0) \quad \mathbf{e}_1 = (0, 1, 0, 0, 0) \quad \mathbf{e}_2 = (0, 0, 1, 0, 0)$$

$$\mathbf{e}_3 = (0, 0, 0, 1, 0) \quad \mathbf{e}_4 = (0, 0, 0, 0, 1)$$

Každý vektor $\mathbf{f} \in S_5$ možno napísať ako lineárnu kombináciu vektorov bázy B . Napríklad

$$\begin{aligned} \mathbf{f} &= (1, 1, 0, 1, 1) = \\ &= 1 \cdot (1, 0, 0, 0, 0) + 1 \cdot (0, 1, 0, 0, 0) + 0 \cdot (0, 0, 1, 0, 0) + 1 \cdot (0, 0, 0, 1, 0) + \\ &+ 0 \cdot (0, 0, 0, 0, 1) = 1 \cdot \mathbf{e}_0 + 1 \cdot \mathbf{e}_1 + 0 \cdot \mathbf{e}_2 + 1 \cdot \mathbf{e}_3 + 1 \cdot \mathbf{e}_4 \end{aligned}$$

Vektorový priestor vnoríme ako podpriestor \mathcal{S} do vektorového priestoru S_8 . Bázu B vektorového podpriestoru \mathcal{S} budú 8 zložkové vektory, ktoré budú rozšírením vektorov bázy B_5 a budú zodpovedať polynómom, ktoré sú deliteľné generujúcim polynómom $q(x) = x^3 + x + 1$.

$$\mathbf{b}_0 = (1, 0, 0, 0, 0, 0, 0, 1) \quad \mathbf{b}_1 = (0, 1, 0, 0, 0, 1, 0, 1) \quad \mathbf{b}_2 = (0, 0, 1, 0, 0, 1, 1, 1)$$

$$\mathbf{b}_3 = (0, 0, 0, 1, 0, 1, 1, 0) \quad \mathbf{b}_4 = (0, 0, 0, 0, 1, 0, 1, 1)$$

Zabezpečiť vektor $\mathbf{f} = (1, 1, 0, 1, 1)$ cyklickým kódom znamená vyjadriť ho v báze \mathcal{B} , teda

$$\mathbf{f}^* = 1 \cdot \mathbf{b}_0 + 1 \cdot \mathbf{b}_1 + 0 \cdot \mathbf{b}_2 + 1 \cdot \mathbf{b}_3 + 1 \cdot \mathbf{b}_4 = (1, 1, 0, 1, 1, 0, 0, 1)$$

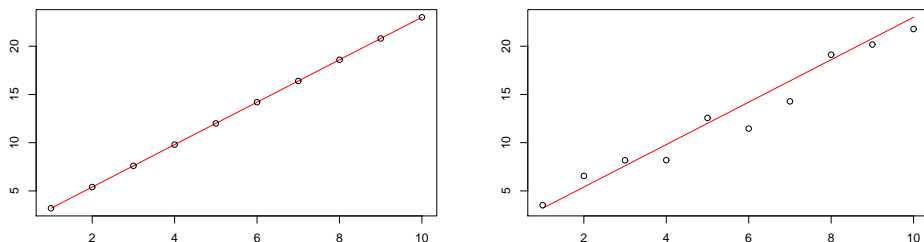
Vektoru $(1, 1, 0, 1, 1, 0, 0, 1)$ zodpovedá polynóm $x^7 + x^6 + x^4 + x^3 + 1$, ktorý je naozaj bezo zvyšku deliteľný polynómom generujúcim polynómom $q(x) = x^3 + x + 1$.

V prípade, že pri prenose komunikačným prostredím, teda transformácii vektora \mathbf{f} na vektor \mathbf{f}' nastane chyba a výsledný polynóm nebude deliteľný polynómom $q(x)$, tak vektor $\mathbf{f}' \notin \mathcal{S}$. \square

Úloha 8.5.6 *Ukážte, že ak sú dva polynómy deliteľné polynómom $q(x)$, tak aj každá ich lineárna kombinácia je deliteľná polynómom $q(x)$.*

9 Modelovanie náhodnosti v procesoch

Pozrime sa na dva grafy na obrázku 9.1, na ktorých sú zobrazené namerané dáta spolu s regresnou priamkou. Ak by študent, vedec, lekár, ekonóm, inžinier či klimatológ



Obr. 9.1. Ideálny a reálny proces

tvrdil, že reálne namerané dáta sú z ľavého obrázka, je pravdepodobné, že nehovorí pravdu. Prakticky namerané dáta sa takmer nikdy nesprávajú ideálne.

V predchádzajúcich kapitolách sme sa preto venovali hľadaniu procesov, ktoré v istom zmysle zachycujú dôležité črty procesov. Typická úloha bola nájsť „zjednodušený“ proces \hat{f} , ktorý najlepšie aproximuje daný proces f :

- v lineárnej regresii sme hľadali lineárnu priamku (alebo krivku), ktorá najlepšie aproximuje proces
- pomocou Fourierovej transformácie sme hľadali periodické vlastnosti procesu
- pomocou Karhunen-Loèveovej transformácie sme hľadali dôležité smery v náhodných dátach

V tejto kapitole obrátíme otázku a skúsime vytvoriť model, z ktorého prichádzajú namerané dáta. Hlavnou myšlienkou je pridanie náhodnosti, avšak ako presne ju pridať si vyžaduje vhodný pohľad na problém.

9.1 Modelovanie lineárnej regresie

Povedzme, že sme dostali dáta vo forme dvoch vektorov $\mathbf{x} = (x_1, \dots, x_m)$ a $\mathbf{y} = (y_1, \dots, y_m)$. Predpokladajme, že medzi y_i a x_i existuje lineárna závislosť

$$y_i = ax_i + b$$

Ak však hodnoty y_i nie sú úplne presne rovné $ax_i + b$, do meraní sa premietla nejaká chyba. Toto si uvedomil už začiatkom 19. storočia nemecký matematik K. F. Gauss, ktorý robil geodetické merania. Ten navrhol, že táto chyba je reálna náhodná premenná, ktorú označíme ϵ , a teda v skutočnosti presne platí:

$$y_i = ax_i + b + \epsilon_i, \quad (9.1.1)$$

kde ϵ_i sú hodnoty náhodnej premennej ϵ .

Ak náhodná premenná ϵ má nenulovú strednú hodnotu $\mu \neq 0$, tak chyby v dátach sa nazývajú systematické. Ak napríklad tachometer na aute ukazuje od 0 do 7 km viac než je skutočná rýchlosť, tak meria rýchlosť so systematickou chybou. To sa dá napraviť tým, že od dát y_i odpočítame $\mu = E(\epsilon)$, a potom už chyby majú strednú hodnotu 0 a prestaneme robiť systematickú chybu.

Keď je známe pravdepodobnostné rozdelenie $p(x)$ chýb ϵ , tak sa koeficienty a a b dajú vypočítať metódou, ktorú štatistickí nazývajú metódou maximálnej vierohodnosti (ML metóda z anglického Maximum likelihood). Metóda maximalizuje vierohodnostnú funkciu modelu, ktorá je definovaná vzorcom

$$V(\mathbf{y}|\mathbf{x}; a, b) = \prod_{i=1}^m p(\epsilon_i) = \prod_{i=1}^m p(y_i - ax_i - b).$$

Ukážeme si jej použitie v nasledujúcom príklade.

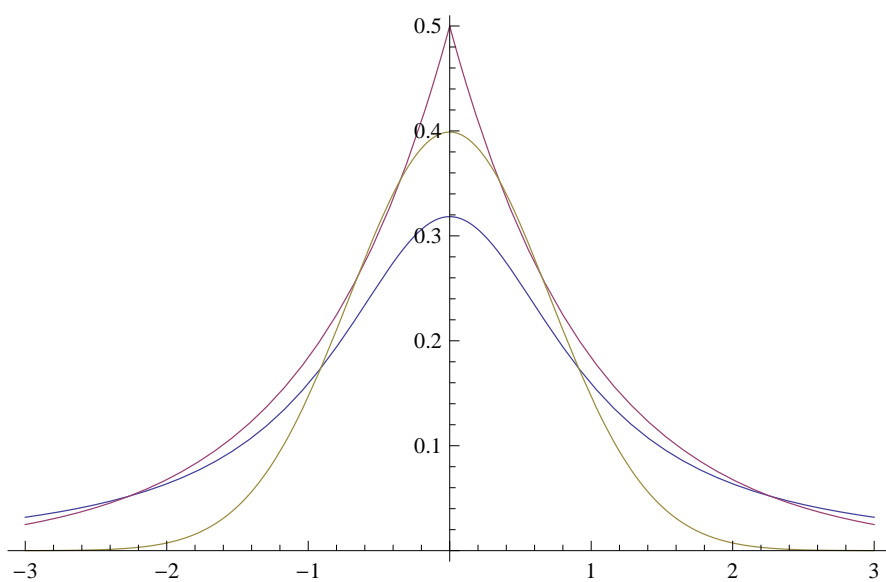
Príklad 9.1.1 *Predpokladajme, že sme namerali nejaké dáta: Našou úlohou je nájsť*

x_i	1	2	3	4	5	6	7	8	9	10
y_i	3.52	5.53	7.7	10.4	11.6	13.9	16.2	18.9	20.9	23.5

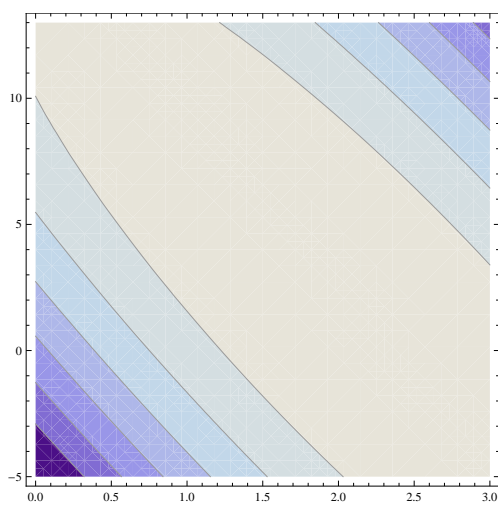
pomocou metódy maximálnej vierohodnosti model $y = ax + b$, ktorý má najväčšiu vierohodnosť. Na to však musíme spraviť predpoklad o pravdepodobnostnom rozdelení chýb ϵ . Úlohu budeme riešiť pre tri rôzne rozdelenia, aby sme si ukázali, že výsledok závisí od toho, ako sú chyby rozdelené.

- normálne rozdelenie s pravdepodobnostnou funkciou $p(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$
- Laplacovo rozdelenie s pravdepodobnostnou funkciou $p(x) = \frac{1}{2}e^{-|x|}$
- Cauchyho rozdelenie s pravdepodobnostnou funkciou $p(x) = \frac{1}{\pi(x^2+1)}$

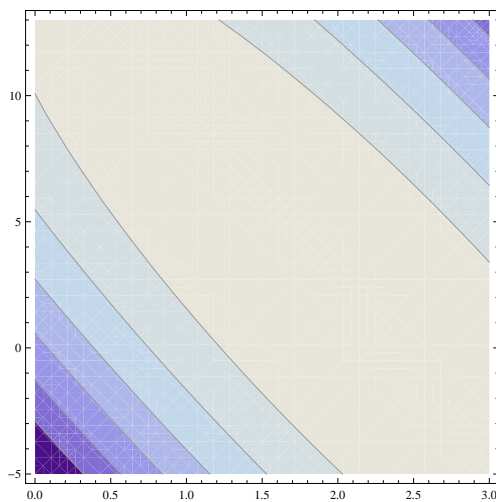
Priebeh týchto funkcií je zobrazený na obrázku 9.2. Na obrázkoch 9.3, 9.4 a 9.5 je zobrazená veľkosť vierohodnosti pre rôzne hodnoty a, b . Čím bledšia oblasť, tým väčšia



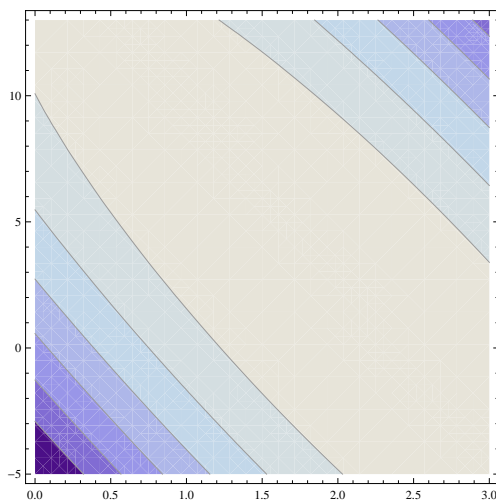
Obr. 9.2. Grafy funkcií normálneho, Laplaceovho a Cauchyho rozdelenia



Obr. 9.3. Veľkosť funkcie vierohodnostnosti pre Gaussovo (normálne) rozdelenie



Obr. 9.4. Veľkosť funkcie vierohodnostnosti pre Laplacovo rozdelenie



Obr. 9.5. Veľkosť funkcie vierohodnostnosti pre Cauchyho rozdelenie

vierohodnosť modelu. V prípade Cauchyho rozdelenia je poloha maxima najzreteľnejšia

Na určenie presnej hodnoty, kde sa nachádza maximum vierohodnosti, použijeme matematický softvér alebo gradientovú metódu z kapitoly 7.6, ktorý nám zistí tieto hodnoty.

$(a, b) = (2.20067, 1.11133)$ v prípade normálneho rozdelenia

$(a, b) = (2.20286, 1.13066)$ v prípade Laplaceovho rozdelenia

$(a, b) = (2.20022, 1.1191)$ v prípade Cauchyho rozdelenia

Vidíme, že výsledky sú podobné, čo sa dá vysvetliť tým, že pravdepodobnostné funkcie pre chyby ϵ_i mali veľmi podobný tvar.

Keby sme vypočítali regresnú priamku $y = ax + b$ pre dáta z predchádzajúceho príkladu, vyšlo by nám $(a, b) = (2, 20067; 1, 11133)$ teda presne ten istý výsledok ako dáva ML metóda pre normálne rozdelenie. Toto nie je náhoda, pretože platí nasledujúca veta.

Veta 9.1.1 Predpokladajme, že až na chyby ϵ_i platí lineárna závislosť medzi procesom \mathbf{f} a procesmi \mathbf{b}_i , t.j

$$\mathbf{f} = a_1 \mathbf{b}_1 + \dots + a_n \mathbf{b}_n + \overbrace{(\epsilon_1, \dots, \epsilon_m)}^{\text{chyby}}$$

Ak chyby ϵ_i majú Gaussovské (normálne) rozdelenie so strednou hodnotou 0 a disperziou σ^2 , tak ML metóda dáva rovnaké výsledky ako metóda najmenších štvorcov.

Dôkaz:

Nech je daný vektor $\mathbf{f} = (f_1, \dots, f_m)$, ku ktorému chceme nájsť koeficienty a_1, \dots, a_n lineárnej kombinácie vektorov $a_1 \mathbf{b}_1 + \dots + a_n \mathbf{b}_n$ pomocou ML metódy. Rozpíšme si vektory \mathbf{b}_i po zložkách

$$\mathbf{b}_1 = (b_{11}, \dots, b_{m1})$$

...

$$\mathbf{b}_n = (b_{1n}, \dots, b_{mn})$$

Podľa predpokladu ϵ_i , zložky rozdielu $\mathbf{f} - a_1 \mathbf{b}_1 - \dots - a_n \mathbf{b}_n$, sú z normálneho rozdelenia s pravdepodobnostnou funkciou $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2}$. Toto sa značí $\epsilon_i \sim N(0, \sigma^2)$. Teraz ideme maximalizovať vierohodnostnú funkciu

$$V(\mathbf{f}|\mathbf{b}_i; a_1, \dots, a_n) = \prod_{i=1}^m p(\epsilon_i).$$

Avšak bude šikovnejšie maximalizovať jej logaritmus namiesto samotnej funkcie. To nezmení miesto maxima, pretože logaritmus je rastúca funkcia. Pre logaritmus vier-

hodnostnej funkcie máme nasledujúce vyjadrenie:

$$\begin{aligned}
 \log V(\mathbf{f}|\mathbf{b}_i; a_1, \dots, a_n) &= \log \prod_{i=1}^m p(\epsilon_i) \\
 &= \log \prod_{i=1}^m \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\epsilon_i^2/2\sigma^2} \\
 &= \left(\sum_{i=1}^m \log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) \right) + \sum_{i=1}^m \log e^{-\epsilon_i^2/2\sigma^2} \\
 &= -\frac{m}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^m \epsilon_i^2
 \end{aligned}$$

Prvý člen nezávisí od výberu koeficientov lineárnej kombinácie a_1, \dots, a_n a druhý si môžeme napísať

$$\begin{aligned}
 \sum_{i=1}^m \epsilon_i^2 &= \sum_{i=1}^m (f_i - a_1 b_{k1} - \dots - a_n b_{kn})^2 \\
 &= \|\mathbf{f} - a_1 \mathbf{b}_1 - \dots - a_n \mathbf{b}_n\|^2
 \end{aligned}$$

Táto suma je najmenšia, ak lineárna kombinácia vektorov $a_1 \mathbf{b}_1 + \dots + a_n \mathbf{b}_n$ je najbližšia k \mathbf{f} , a teda metóda najmenších štvorcov dáva ten istý výsledok, ako ML v prípade gaussovského rozdelenia chýb.

Zaujímavým faktom je, že výsledok ML metódy nezávisí od rozptylu gaussovského rozdelenia, ale len od toho, že má strednú hodnotu 0.

K výpočtu rozptylu chyby môžeme pristúpiť dvoma rôznymi spôsobmi.

Základom prvého je vzorec $\sigma^2 = E(X^2)$, ktorý platí pre náhodnú premennú s priemerom 0. Keďže predpokladáme, že ϵ_i sú gausovsky rozdelené so strednou hodnotou 0, na odhad σ^2 môžeme použiť vzorec

$$\text{odhad } \sigma^2 = \frac{1}{m-1} \sum \epsilon_i^2 = \frac{1}{m-1} \|\mathbf{f} - a_1 \mathbf{b}_1 - \dots - a_n \mathbf{b}_n\|^2.$$

Druhý spôsob spočíva v maximalizácii vierohodnosti ako funkcie rozptylu. Spočítajme deriváciu vierohodnosti podľa σ

$$\begin{aligned}
 \frac{\partial \log V(\mathbf{f}|\mathbf{b}_i; a_1, \dots, a_k)}{\partial \sigma} &= \frac{\partial}{\partial \sigma} \left(-\frac{m}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^m \epsilon_i^2 \right) \\
 &= -\frac{m}{2} \frac{4\pi\sigma}{2\pi\sigma^2} + \frac{1}{\sigma^3} \sum_{i=1}^m \epsilon_i^2 \\
 &= -\frac{m}{\sigma} + \frac{1}{\sigma^3} \|\mathbf{f} - a_1 \mathbf{b}_1 - \dots - a_n \mathbf{b}_n\|^2.
 \end{aligned}$$

Maximálna vierohodnosť bude v tom bode, kde derivácia je rovná 0, tj.

$$0 = -\frac{m}{\sigma} + \frac{1}{\sigma^3} \|\mathbf{f} - a_1 \mathbf{b}_1 - \dots - a_n \mathbf{b}_n\|^2$$

odkiaľ

$$\text{ML odhad } \sigma^2 = \frac{1}{m} \|\mathbf{f} - a_1 \mathbf{b}_1 - \dots - a_n \mathbf{b}_n\|^2$$

Vidíme, že ML odhad sa pre malé hodnoty m (počet pozorovaní) mierne líši od prvého odhadu, avšak so zväčšujúcim sa počtom dát sa tieto odhady k sebe približujú.

Ako môžeme využiť túto informáciu o rozptyle? Povedzme, že sa snažíme identifikovať **odľahlé hodnoty**, ktoré majú podozrivo vysokú chybu. Čebyševova veta napríklad hovorí, že aspoň $8/9 \approx 89\%$ hodnôt leží nanajvýš vo vzdialenosti 3σ od strednej hodnoty. O normálnom rozdelení vieme silnejšie tvrdenie, že viac ako 95% pozorovaní leží nanajvýš vo vzdialenosti 2σ od očakávanej hodnoty. Vediať hodnotu σ môžeme posúdiť, či dané meranie je odľahlé.

Úloha 9.1.1 Zistite, či meranie $(x, y) = (5.5, 11)$ leží vo vzdialenosti viac než 3σ pre dáta z príkladu 9.1.1.

9.2 Stacionárny proces

Analýza procesov často odhalí deterministickú črtu procesu, či už lineárny trend, alebo nenulový priemerný trend, alebo periodické správanie. Zvyšok procesu má veľmi nedeterministické, zdanlivo náhodné správanie. Na štúdium a modelovanie nedeterministických procesov sa používa teória stacionárnych procesov [3], [4]. Táto teória sa rozvinula v súvislosti so štúdiom geofyzikálnych javov, no našla široké uplatnenie aj v ekonómii pri modelovaní trhu cenných papierov a ekonomiky, ale aj v technike a to najmä v spracovaní zvuku a rozpoznávaní reči.

Porovnajme dva grafy procesov zobrazených na obrázku 9.6. Obidva procesy sú náhodné procesy. Spoločné majú to, že ani jeden nemá nejaký lineárny trend, ktorým by sa oddľaloval od 0. Avšak vidíme, že ten dolu sa „mení“ podstatne menej. Ako keby predchádzajúca hodnota mala veľký vplyv na nasledujúcu hodnotu. Toto je typické napríklad pre hodnoty komodít a akcií na burzách. Cena z predošlého dňa je podstatným faktorom pre určenie dnešnej obchodnej ceny.

Teraz si tieto vlastnosti sformalizujeme definíciou stacionárneho procesu.

Stacionárny proces je náhodný proces s nasledujúcimi vlastnosťami

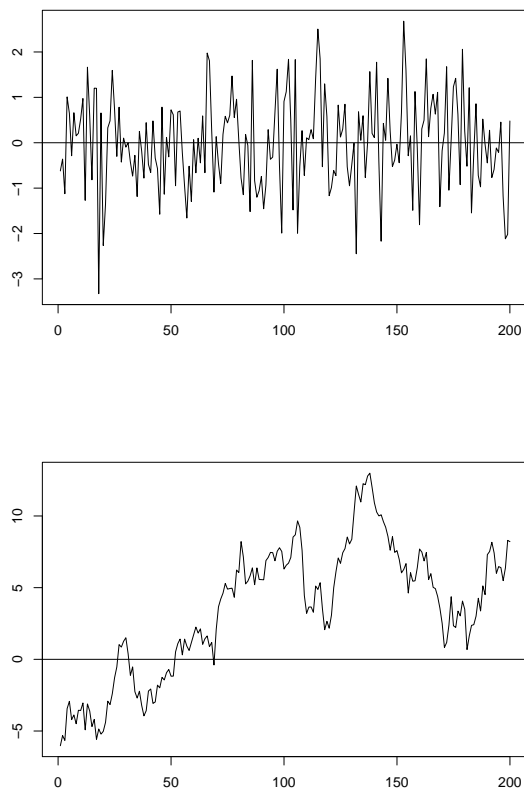
- stredná hodnota sa nemení v čase a je konečná:

$$E(X_s) = E(X_t)$$

pre všetky časy s, t

- kovarianca medzi rozdeleniami procesu v rôznych časoch je konečná, a závisí len od rozdielu časov, tj.

$$\text{Cov}(X_s, X_t) = \text{Cov}(X_{s+d}, X_{t+d})$$



Obr. 9.6. Dve realizácie náhodných stacionárnych procesov

Keďže kovariancia závisí iba od vzdialenosti $|s-t|$, môžeme definovať autokovariančnú funkciu

$$g(h) = \text{Cov}(X_s, X_{s+h})$$

Špeciálne, pre $s = t$ dostávame, že disperzia v každom čase je rovnaké konečné číslo $g(0) = \sigma^2$.

Zoberme si napríklad hladinu Dunaja. Toto nie je stacionárny proces, pretože v jarňých mesiacoch býva v priemere viac vody, ako v zimných. Avšak ak meriame len hladinu Dunaja 1. januára v jednotlivých rokoch, tak toto už je pravdepodobne dobrý príklad na stacionárny proces.

Ako ďalší príklad si zoberte cenu akcie. Keďže fiškálna politika centrálnej banky je udržiavať mierny rast, ceny v priemere mierne rastú o niekoľko percent ročne. Preto ani cena akcie nie je stacionárny proces. Aby sme mohli hovoriť o stacionárnom procese, museli by sme ceny v jednotlivých rokoch upraviť o infláciu.

Ak má stacionárny proces nenulovú strednú hodnotu μ , tj. $E(X_t) = \mu$ pre každý čas t , tak odčítaním tejto strednej hodnoty dostaneme **centrovaný** stacionárny proces $X_t - \mu$, pričom vzájomné kovariancie sa nezmenia.

Pre stacionárny proces je dôležitou charakteristikou jeho autokorelačná funkcia

$$\begin{aligned} r(h) &= \frac{\text{Cov}(X_s, X_{s+h})}{\sqrt{\text{Cov}(X_s, X_s) \cdot \text{Cov}(X_{s+h}, X_{s+h})}} \\ &= \frac{\text{Cov}(X_s, X_{s+h})}{\text{Cov}(X_s, X_s)} \end{aligned}$$

Autokorelačná funkcia nadobúda hodnoty od -1 do 1 , pričom $r(0)$ je vždy 1 .

Na obrázku 9.7 je znázornená autokorelačná funkcia nahrávky hovorenej reči. Vidíme, že tento zvuk má vysokú autokoreláciu pre $h = 400$. Keďže zvuk bol vzorkovaný na frekvencii 44100 Hz, môžeme z toho usúdiť, že je takmer periodický s frekvenciou $44100/400 \approx 110$ Hz.

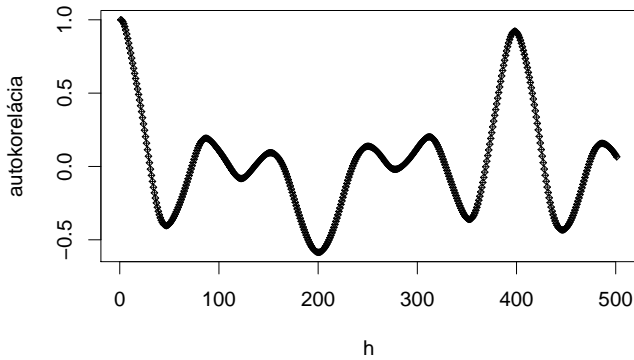
Väčšina štatistických softvérov má funkciu na vykreslenie autokorelačnej funkcie. Typicky sú zobrazené aj dve vodorovné čiary, ktoré ukazujú hranice, medzi ktorými hodnoty autokorelácií sú štatisticky zanedbateľné. Hodnoty v týchto medziach môžeme považovať za nulové.

9.3 Využitie stacionarity na predikciu a interpoláciu

Hlavné použitie predpokladov o stacionárnosti je v predpovedi ďalších hodnôt procesu. Typický príklad takejto predpovede je, že poznáme niekoľko posledných hodnôt procesu a chceme predpovedať ďalšiu hodnotu. Napríklad, ak poznáme posledných päť hodnôt $X_{t-5}, X_{t-4}, X_{t-3}, X_{t-2}, X_{t-1}$, bude nás zaujímať, či neexistuje lineárna kombinácia

$$a_5 X_{t-5} + a_4 X_{t-4} + a_3 X_{t-3} + a_2 X_{t-2} + a_1 X_{t-1},$$

ktorá by čo najpresnejšie predpovedala hodnotu X_t procesu v čase t . Takáto lineárna kombinácia sa nazýva aj **lineárny prediktor**.



Obr. 9.7. Autokorelačná funkcia hovorenej samohlásky

Príklad 9.3.1 Predpokladajme, že priemerná denná teplota počas januára v Brezne sa správa ako centrovanej stacionárny proces T_t . Našou úlohou je odhadnúť teplotu dnes, ak vieme, že včerajšia teplota bola T_{t-1} . Ak by včerajšia teplota bola nulová, je prirodzené odhadnúť aj dnešnú teplotu nulou. Ak však včerajšia teplota bola povedzme nad nulou, dá sa očakávať, že aj dnešná teplota bude nad nulou, napriek tomu, že dlhodobý teplotný priemer je nula. Preto sa budeme pozeráť na odhady dnešnej teploty v tvare $\hat{T}_t = k \cdot T_{t-1}$. Vhodnosť odhadu budeme merať očakávanou chybou

$$\begin{aligned} E((T_t - \hat{T}_t)^2) &= E((T_t - kT_{t-1})^2) \\ &= E(T_t^2 - 2kT_tT_{t-1} + k^2T_{t-1}^2) \\ &= E(T_t^2) - 2k\text{Cov}(T_t, T_{t-1}) + k^2E(T_{t-1}^2) \end{aligned}$$

pretože $E(T_t) = E(T_{t-1}) = 0$.

Vidíme, že veľkosť chyby je kvadratickou funkciou v závislosti od k . Derivovaním nájdeme miesto lokálneho minima

$$\frac{\partial E(T_t - \hat{T}_t)^2}{\partial k} = -2\text{Cov}(T_t, T_{t-1}) + 2kE(T_{t-1}^2)$$

a tým aj vzorec pre najlepší odhad

$$\hat{T}_t = kT_{t-1} = \frac{\text{Cov}(T_t, T_{t-1})}{E(T_{t-1}^2)}T_{t-1}$$

a keďže rozptyl T_t je v každom čase rovnaký

$$\hat{T}_t = \frac{\text{Cov}(T_t, T_{t-1})}{\sqrt{E(T_{t-1}^2)E(T_t^2)}} T_{t-1}$$

inými slovami

$$\hat{T}_t = \left(\text{autokorelačný koeficient } r(1) \text{ procesu } T_t \right) \cdot T_{t-1}$$

Úloha 9.3.1 22. januára bola v Brezne teplota -5 stupňov. Na základe predchádzajúceho príkladu odhadnite, aká teplota bola v Brezne 23. januára, ak viete, že

- a) disperzia $E(T_t^2) = 17$ a autokorelácia $r(1) = 0.6$,
- b) disperzia $E(T_t^2) = 15$ a kovariancia $\text{Cov}(T_t, T_{t-1}) = 10$.

Príklad 9.3.2 Predpokladajme, že výmenný kurz dolára voči euru má dlhodobý priemer 1.5 a odchýlka od tohto priemeru je centrovanej stacionárny proces, ktorý si označíme K_t , t.j. platí vzorec

$$\text{kurz v deň } t = 1.5 + K_t.$$

Ukážme, si aký vplyv má autokorelácia na predikciu. Našou úlohou je odhadnúť kurz D hodnotou \hat{D} , s ktorým dnes budeme vymieňať doláre s eurami za predpokladu, že vieme včerajší výmenný kurz V . Tri príklady takých odhadov sú

- odhad odvodený od včerašieho kurzu $\hat{D}_1 = V$,
- odhad odvodený od dlhodobého priemeru $\hat{D}_2 = 1.5$,
- odhad, ktorý je priemerom predošlých dvoch odhadov $\hat{D}_3 = \frac{1}{2}(V + 1.5)$.

Samozrejme, tieto tri odhady nie sú jediné možné. Obmedzíme sa len na lineárne odhady typu $\hat{D} = a + bV$, ktoré zahŕňajú všetky tri už uvedené prípady. Kritériom kvality odhadu bude priemerná veľkosť chyby $E(D - \hat{D})^2$.

Dosadíme

$$\begin{aligned} E((D - \hat{D})^2) &= E(D^2 - 2D\hat{D} + \hat{D}^2) \\ &= E(D^2) - 2E(D \cdot (a + bV)) + E((a + bV))^2 \\ &= E(D^2) - 2aE(D) - 2bE(D \cdot V) + a^2 + 2abE(V) + b^2E(V^2) \end{aligned}$$

Z definície stacionárneho procesu vyplýva, že $E(D) = E(V) = 1.5$, a teda

$$E((D - \hat{D})^2) = E(D^2) - 3a - 2bE(D \cdot V) + a^2 + 3ab + b^2E(V^2)$$

Táto hodnota bude minimalizovaná, keď gradient bude nulový, t.j.

$$\begin{aligned} \frac{E(D - \hat{D})^2}{\partial a} &= -3 + 2a + 3b = 0 \\ \frac{E(D - \hat{D})^2}{\partial b} &= -2E(D \cdot V) + 3a + 2bE(V^2) = 0 \end{aligned}$$

Odtiaľ

$$\begin{aligned} a &= \frac{1}{2}(3 - 3b) \\ -2\text{Cov}(D, V) + \frac{3}{2}(3 - 3b) + 2bE(V^2) &= 0 \\ b &= \frac{2E(D \cdot V) - 9/2}{2E(V^2) - 9/2} \end{aligned}$$

Dostali sme síce výsledok, ale výsledné vzorce nám veľa neobjasňujú. To sa zmení, ak výpočet mierne modifikujeme. Budeme hľadať odhad v tvare

$$\hat{D}' = 1.5 + c + d(V - 1.5) = 1.5 + c + dK_{t-1}$$

Odhad v takomto tvare dosadíme do vzorca pre chybu

$$\begin{aligned} E(D - \hat{D}')^2 &= E((1.5 + K_t - (1.5 + c + dK_{t-1}))^2) \\ &= E((K_t - c - dK_{t-1})^2) \\ &= E(K_t^2) + c^2 + d^2E(K_{t-1}^2) \\ &\quad - 2cE(K_t) - 2cdE(K_{t-1}) - 2dE(K_tK_{t-1}) \end{aligned}$$

a keďže K_t je centrovanej stacionárny proces, dostávame

$$E(D - \hat{D}')^2 = E(K_t^2) + c^2 + d^2E(K_{t-1}^2) - 2d\text{Cov}(K_t, K_{t-1})$$

Minimum nastáva v bode, v ktorom je gradient nulový, a teda sú splnené rovnice

$$\begin{aligned} \frac{E(D - \hat{D}')^2}{\partial c} &= 2c = 0 \\ \frac{E(D - \hat{D}')^2}{\partial d} &= 2dE(K_{t-1}^2) - 2\text{Cov}(K_t, K_{t-1}) = 0 \end{aligned}$$

Odtiaľ dostaneme $c = 0$ a

$$d = \frac{\text{Cov}(K_t, K_{t-1})}{E(K_{t-1}^2)} = \frac{\text{Cov}(K_t, K_{t-1})}{\sqrt{E(K_{t-1}^2)E(K_t^2)}},$$

a teda d je korelácia medzi nasledujúcimi hodnotami stacionárneho procesu K_t , tj. prvá hodnota autokorelačnej funkcie.

Úloha 9.3.2 Včerajší výmenný kurz dolára k euru bol 1.7. Vypočítajte autokoreláciu $r(1)$, hodnoty a, b, c, d a predpovede \hat{D}, \hat{D}' dnešného podľa predchádzajúceho príkladu v nasledujúcich prípadoch:

- a) $E(K_t^2) = 3, \text{Cov}(K_t, K_{t-1}) = 0.9,$
- b) $E(K_t^2) = 5, \text{Cov}(K_t, K_{t-1}) = 1.5,$

$$c) E(K_t^2) = 5, \text{Cov}(K_t, K_{t-1}) = 4.5.$$

Podme na záver vyriešiť problém interpolácie pre stacionárny proces. Nech P_t je stacionárny proces, o ktorom budeme predpokladať, že je centrovanej. Budeme sa snažiť odhadnúť hodnotu v čase t , ak sú známe hodnoty v časoch $t-1$ a $t+1$. Budeme hľadať koeficienty a, b tak, aby lineárny prediktor $\hat{P}_{a,b} = aP_{t-1} + bP_{t+1}$ bol čo najpresnejší, tj. aby chyba

$$E((P_t - \hat{P}_{a,b})^2)$$

bola čo najmenšia. Dosadíme a upravíme

$$\begin{aligned} E((P_t - aP_{t-1} - bP_{t+1})^2) &= E(P_t^2) + E(a^2 P_{t-1}^2) + E(b^2 P_{t+1}^2) \\ &\quad - 2E(P_t \cdot aP_{t-1}) - 2E(P_t \cdot bP_{t+1}) + 2E(aP_{t-1} \cdot bP_{t+1}) \\ &= (1 + a^2 + b^2)E(P_t^2) - 2(a + b)\text{Cov}(P_t, P_{t+1}) + 2ab\text{Cov}(P_{t-1}, P_{t+1}) \\ &= E(P_t^2) \left((1 + a^2 + b^2) - 2(a + b)r(1) + 2abr(2) \right) \end{aligned}$$

Minimum nastáva v bode, kde gradient funkcie vo veľkých zátvorkách je nulový, tj. ak

$$\begin{aligned} 2a - 2r(1) + 2br(2) &= 0 \\ 2b - 2r(1) + 2ar(2) &= 0 \end{aligned} \tag{9.3.1}$$

čo je sústava lineárnych rovníc pre koeficienty a, b v závislosti od autokorelačných koeficientov $r(1), r(2)$. Vidíme teda, že predikcia a interpolácia sú dobre riešiteľné pre stacionárne procesy, a najdôležitejšími údajmi pre návrh lineárnych prediktorov sú autokorelačné koeficienty.

9.4 Biely šum

V modelovaní javov sa v praxi často robí predpoklad o stacionárnosti procesu. Ako sme si ukázali v predchádzajúcej kapitole, tento predpoklad umožňuje nájsť dobré lineárne prediktory ďalších hodnôt procesu. Toto však mnohokrát nestačí. Pri dopadových štúdiách sa napríklad musí určiť, či pravdepodobnosť opakujúcich sa dažďov, ktoré by viedli k povodniam, je viac než raz za 200 rokov. Alebo je cieľom určiť, aká je pravdepodobnosť, že cena zlata, alebo inej suroviny, ostane nad určitou hladinou počas obdobia výroby nejakého produktu. Preto je dôležité byť schopný simulovať stacionárne procesy, tj. v počítači vyprodukovať mnohé realizácie jedného stacionárneho procesu. Najzákladnejší stacionárny proces, ktorý sa používa, je proces bieleho šumu.

Stacionárny proces Z_t nazveme **bielym šumom**, ak je centrovanej, a jeho hodnoty sú v rôznych časoch navzájom lineárne nekorelované. Inakšie povedané $E(Z_t) = 0$, a autokorelačná funkcia procesu Z_t je okrem hodnoty v bode 0 všade nulová

$$r(h) = \begin{cases} 0 & \text{ak } h \neq 0 \\ 1 & \text{ak } h = 0 \end{cases}$$

Podobne jeho autokovariančná funkcia závisí od iba od jedného parametra σ^2

$$g(t-s) = \text{Cov}(Z_t, Z_s) = \begin{cases} 0 & \text{ak } s \neq t \\ \sigma^2 & \text{ak } s = t \end{cases}$$

čo sa niekedy značí pomocou hranatých zátvoriek takto

$$g(t-s) = \text{Cov}(Z_t, Z_s) = \sigma^2[s=t]. \quad (9.4.1)$$

Dohoda je taká, že hodnota výrazu v zátvorkách je 1, ak je výraz pravdivý. Ak výraz je nepravdivý, hodnota výrazu je nula.

Ako príklad bieleho šumu si môžeme uviesť hladinu Dunaja meranú v ročných intervaloch. Keby sme chceli simulovať biely šum, stačí zobrať postupnosť náhodne vygenerovaných čísel. Ak sme použili dobrý generátor náhodných čísel, tak výsledný proces je bielym šumom.

To, že je proces Z_t bielym šumom sa značí $Z_t \sim WN(0, \sigma^2)$, vyjadrujúc, že stredná hodnota je 0, a disperzia v každom čase t je σ^2 .

9.5 $MA(q)$ procesy

Teraz si predstavíme novú triedu stacionárnych procesov. Na rozdiel od bieleho šumu, tieto procesy môžu mať nenulovú autokoreláciu medzi náhodnými premennými X_t a X_{t+1} . Názov týchto procesov je skratka z anglického pojmu *moving average*, t.j. kľzavý priemer. Najjednoduchším príkladom $MA(q)$ procesu je $MA(1)$ proces. Ten vznikne z bieleho šumu aplikovaním metódy kľzavých súčtov. Konkrétne, proces X_t je $MA(1)$ proces, ak existuje proces bieleho šumu $Z_t \sim WN(0, \sigma^2)$ taký, že platí

$$X_t = Z_t + aZ_{t-1},$$

kde a je reálna konštanta nezávisiaca od času t . Ľahko sa overí, že tento proces je centrovateľný

$$E(X_t) = E(Z_t + aZ_{t-1}) = E(Z_t) + aE(Z_{t-1}) = 0$$

Ďalej vypočítame kovariančnú funkciu medzi rozdeleniami procesu v rovnakom čase

$$\begin{aligned} \text{Cov}(X_t, X_t) &= \text{Cov}(Z_t + aZ_{t-1}, Z_t + aZ_{t-1}) \\ &= \text{Cov}((Z_t + aZ_{t-1}) \cdot (Z_t + aZ_{t-1})) \\ &= \text{Cov}(Z_t^2 + 2aZ_tZ_{t-1} + a^2Z_{t-1}^2) \\ &= \text{Cov}(Z_t, Z_t) + 2a\text{Cov}(Z_t, Z_{t-1}) + a^2\text{Cov}(Z_{t-1}, Z_{t-1}) \\ &= \sigma^2 + 2a \cdot 0 + a^2\sigma^2 = \sigma^2(1 + a^2) \end{aligned}$$

Podobne spravíme výpočet pre časy vzdialené o 1:

$$\begin{aligned} \text{Cov}(X_{t+1}, X_t) &= \text{Cov}(Z_{t+1} + aZ_t, Z_t + aZ_{t-1}) \\ &= \text{Cov}(Z_{t+1}, Z_t) + a\text{Cov}(Z_{t+1}, Z_{t-1}) \\ &\quad + a\text{Cov}(Z_t, Z_t) + a^2\text{Cov}(Z_t, Z_{t-1}) \end{aligned}$$

čo môžeme zjednodušiť využíjúc (9.4.1)

$$\begin{aligned}\text{Cov}(X_{t+1}, X_t) &= \sigma^2([t+1=t] + a[t+1=t-1] + a[t=t] + a^2[t=t-1]) \\ &= 0 + 0 + a\sigma^2 + 0 = a\sigma^2\end{aligned}$$

Ak počítame kovarianciu distribúcií vzdialených o čas $d > 1$, tak dostaneme výrazy podobné predchádzajúcemu prípadu

$$\begin{aligned}\text{Cov}(X_{t+d}, X_t) &= \text{Cov}(Z_{t+d} + aZ_{t+d-1}, Z_t + aZ_{t-1}) \\ &= \text{Cov}(Z_{t+d}, Z_t) + a\text{Cov}(Z_{t+d}, Z_{t-1}) \\ &\quad + a\text{Cov}(Z_{t+d-1}, Z_t) + a^2\text{Cov}(Z_{t+d-1}, Z_{t-1}) \\ &= \sigma^2([t+d=t] + a[t+d=t-1] \\ &\quad + a[t+d-1=t] + a^2[t+d-1=t-1])\end{aligned}$$

Prvý výraz v zátvorke je nenulový iba keď $d = 0$, druhý keď $d = -1$, tretí keď $d = 1$ a posledný keď $d = 0$. Preto pre $d > 1$ dostávame

$$\text{Cov}(X_{t+d}, X_t) = 0$$

Tieto výpočty ukazujú, že $MA(1)$ proces je stacionárny proces. Z uvedených výpočtov odvodíme vzorce pre autokovariančnú funkciu.

$$g(h) = \begin{cases} \sigma^2(1 + a^2) & \text{ak } h = 0 \\ \sigma^2 a & \text{ak } h = \pm 1 \\ 0 & \text{ak } |h| > 1 \end{cases}$$

Pre autokorelačnú funkciu platí vzťah

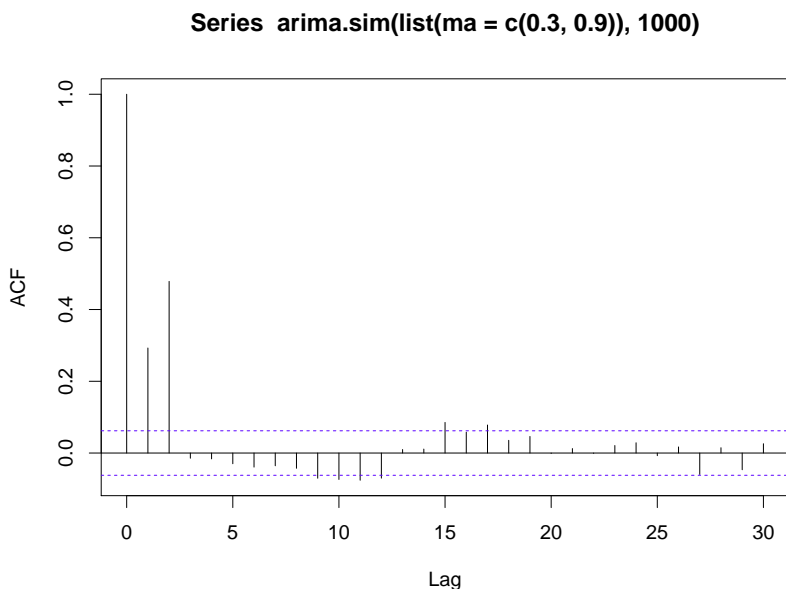
$$r(h) = \begin{cases} 1 & \text{ak } h = 0 \\ \frac{a}{1 + a^2} & \text{ak } h = \pm 1 \\ 0 & \text{ak } |h| > 1 \end{cases}$$

Úloha 9.5.1 Nakreslite graf $r(1)$ v závislosti od parametra a . Ukážte, že pre prvú autokoreláciu $r(1)$ $MA(1)$ procesu platí $-\frac{1}{2} \leq r(1) \leq \frac{1}{2}$.

Všeobecný $MA(q)$ proces vznikne aplikáciou kľazového súčtu s q členmi na biely šum.

$$X_t = Z_t + a_1 Z_{t-1} + \dots + a_{q-1} Z_{t-q+1}$$

Dá sa ukázať, že autokorelačná funkcia $r(h)$ je nulová, ak $h > q$. Tento fakt sa využíva pri modelovaní procesu. Rád modelu q sa vyberá tak, aby zahrnul všetky zjavné nenulové korelácie.



Obr. 9.8. Autokorelačná funkcia realizácie procesu $X_t = Z_t + 0.3Z_{t-1} + 0.9Z_{t-2}$.

Príklad 9.5.1 Nech X_t je $MA(2)$ proces daný vzorcom $X_t = Z_t + 0.3Z_{t-1} + 0.9Z_{t-2}$ a $Z_t \sim WN(0, 1)$. Takýto proces môžeme nasimulovať v programe R, a vykresliť jeho autokorelačnú funkciu, ako vidno na obrázku 9.8. Keďže ide o $MA(2)$ proces, iba autokorelácie $r(1)$ a $r(2)$ sú nenulové.

Úloha 9.5.2 Nech $MA(1)$ proces je daný vzorcom $X_t = Z_t + 3Z_{t-1}$, a $Z_t \sim WN(0, 1)$. Vygenerujte realizáciu (f_1, \dots, f_{1000}) tohto procesu o dĺžke 1000 vzoriek.

- (predikcia) Nájdite koeficienty a , b , také, aby lineárna kombinácia vektorov

$$a(f_1, \dots, f_{998}) + b(f_2, \dots, f_{999})$$

bola čo najbližšie k vektoru (f_3, \dots, f_{1000}) .

- (interpolácia) Nájdite koeficienty a , b , také, aby lineárna kombinácia vektorov

$$a(f_1, \dots, f_{998}) + b(f_3, \dots, f_{1000})$$

bola čo najbližšie k vektoru (f_2, \dots, f_{999}) .

Zopakujte obidve úlohy pre realizáciu procesu o dĺžke 10000. Skúste teoreticky zdôvodniť výsledky.

9.6 $AR(p)$ procesy

Skratka AR pochádza z pomenovania *autoregressive*. Najjednoduchším príkladom je $AR(1)$ proces, ktorý spĺňa rovnicu

$$X_t = bX_{t-1} + Z_t, \quad \text{kde } Z_t \sim WN(0, \sigma^2)$$

pričom sa robia dva predpoklady:

- b je konštanta v absolútnej hodnote menšia ako 1, tj. $|b| < 1$,
- X_s pre $s < t$ nie je korelované s Z_t .

Všimnite si podobnosť s prípadom lineárnej regresie (9.1.1). Spočítajme autokovariančnú funkciu $g(h)$ pre $h \geq 1$.

$$\begin{aligned} g(1) &= \text{Cov}(X_{t+1}, X_t) = \text{Cov}(bX_t + Z_{t+1}, X_t) \\ &= b\text{Cov}(X_t, X_t) + \text{Cov}(Z_{t+1}, X_t) = bg(0) \end{aligned}$$

pretože druhý člen je podľa druhého predpokladu nulový. Ďalej

$$\begin{aligned} g(d) &= \text{Cov}(X_{t+d}, X_t) \\ &= \text{Cov}(bX_{t+d-1} + Z_{t+d}, X_t) = b\text{Cov}(X_{t+d-1}, X_t) + \text{Cov}(Z_{t+d}, X_t) \\ &= bg(d-1), \quad \text{pre } d > 1 \end{aligned}$$

Z týchto rovníc dostávame autokorelačnú funkciu $AR(1)$ procesu

$$r(h) = \frac{g(h)}{g(0)} = b^h$$

Úloha 9.6.1 Spočítajte autokovariáciu $g(0) = E(X_t^2)$ pre $AR(1)$ proces v závislosti od b a σ^2 .

Simulácia $AR(p)$ procesov je o čosi zložitejšia, než simulácia $MA(q)$ procesov. Keďže sme spravili predpoklad $|b| < 1$, $AR(1)$ sa dá vyjadriť ako nekonečný súčet, ktorého členy veľmi rýchlo konvergujú.

$$X_t = Z_t + bZ_{t-1} + b^2Z_{t-2} + \dots$$

Táto vlastnosť procesu, sa nazýva *kauzalita*. Znamená, že hodnoty procesu sa dajú vyjadriť z predošlých hodnôt bieleho šumu.

Úloha 9.6.2 Nech $AR(1)$ proces je daný vzorcom $X_t - 0.3X_{t-1} = Z_t$, a $Z_t \sim WN(0, 1)$. Vygenerujte realizáciu (f_1, \dots, f_{1000}) tohto procesu o dĺžke 1000 vzoriek.

- (predikcia) Nájdite koeficienty a, b , také, aby lineárna kombinácia vektorov

$$a(f_1, \dots, f_{998}) + b(f_2, \dots, f_{999})$$

bola čo najbližšie k vektoru (f_3, \dots, f_{1000}) .

- (interpolácia) Nájdite koeficienty a, b , také, aby lineárna kombinácia vektorov

$$a(f_1, \dots, f_{998}) + b(f_3, \dots, f_{1000})$$

bola čo najbližšie k vektoru (f_2, \dots, f_{999}) .

Zopakujte úlohu pre realizáciu procesu o dĺžke 10000. Porovnajte s výsledkom úlohy 9.5.2. Skúste teoreticky zdôvodniť výsledky.

Vo všeobecnosti $AR(p)$ proces je proces, ktorý spĺňa rovnicu

$$X_t - b_1 X_{t-1} - \dots - b_p X_{t-p} = Z_t, \quad \text{kde } Z_t \sim WN(0, \sigma^2),$$

pričom Z_t nie je korelované s X_s pre $s < t$, a všetky (aj komplexné riešenia) rovnice

$$z^p - b_1 z^{p-1} - \dots - b_p = 0$$

sú v absolútnej hodnote menšie ako 1.

$AR(p)$ procesy, na rozdiel od $MA(q)$ procesov, majú nekonečne veľa nenulových autokorelácií. Ak sa snažíme namodelovať AR model na danú realizáciu procesu, nedá sa určiť vhodné p z grafu autokorelačnej funkcie. Náhradou je zobrazenie čiastkovej autokorelácie (angl. partial autocorrelation, alebo partial ACF), ktorá vykresľuje koreláciu procesu s procesom, z ktorého bola odstránená najbližšia lineárna kombinácia predchádzajúcich členov. Rád modelu je určíme z posledného významne nenulového člena čiastkovej autokorelačnej funkcie (viď obrázok 9.9).

9.7 Yule-Walkerove rovnice

Autoregresné modely sú populárne, pretože sú ľahšie vypočítateľné oproti MA modelom. Teraz si ukážeme jeden zo spôsobov, ako odhadnúť AR model z dát. Povedzme, že chceme namodelovať $AR(3)$ model. Pre hodnoty $AR(3)$ modelu platia vzorce:

$$X_{t+1} = b_1 X_t + b_2 X_{t-1} + b_3 X_{t-2} + Z_{t+1}$$

$$X_{t+2} = b_1 X_{t+1} + b_2 X_t + b_3 X_{t-1} + Z_{t+2}$$

$$X_{t+3} = b_1 X_{t+2} + b_2 X_{t+1} + b_3 X_t + Z_{t+3}$$

Vynásobíme si tieto vzorce X_t , a zoberieme strednú hodnotu. Dostaneme sústavu rovníc

$$E(X_t \cdot X_{t+1}) = E(b_1 X_t^2) + E(X_t \cdot b_2 X_{t-1}) + E(X_t \cdot b_3 X_{t-2}) + E(X_t \cdot Z_{t+1})$$

$$E(X_t \cdot X_{t+2}) = E(X_t \cdot b_1 X_{t+1}) + E(b_2 X_t^2) + E(X_t \cdot b_3 X_{t-1}) + E(X_t \cdot Z_{t+2})$$

$$E(X_t \cdot X_{t+3}) = E(X_t \cdot b_1 X_{t+2}) + E(X_t \cdot b_2 X_{t+1}) + E(b_3 X_t^2) + E(X_t \cdot Z_{t+3})$$

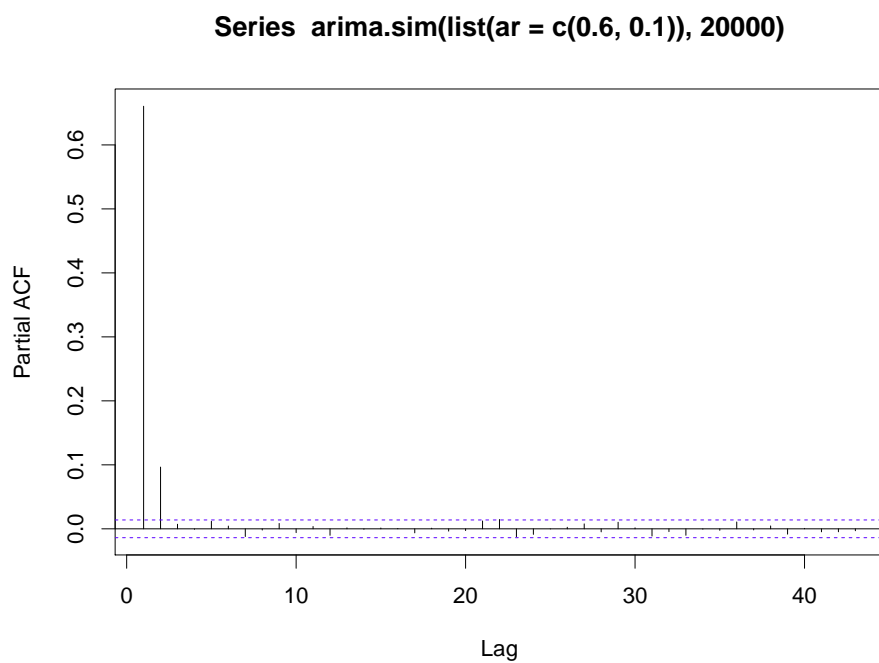
Keďže X_t je centrováný proces, $E(X_s \cdot X_t)$ je autokovariancia $g(s - t)$, a posledné členy sú rovné nule, pretože X_t nie je korelované s nasledujúcimi hodnotami procesu Z . Preto dostávame sústavu rovníc nazývanou **Yule-Walkerovou**:

$$g(1) = g(0)b_1 + g(1)b_2 + g(2)b_3$$

$$g(2) = g(1)b_1 + g(0)b_2 + g(1)b_3$$

$$g(3) = g(2)b_1 + g(1)b_2 + g(0)b_3$$

Tieto rovnice sa dajú použiť dvojakým spôsobom.



Obr. 9.9: Graf čiastkovej autokorelácie simulácie procesu $X_t - 0.6X_{t-1} - 0.1X_{t-2} = Z_t$

- Prvý spôsob použitia je, keď hľadáme koeficienty procesu aproximujúce danú realizáciu procesu. V tom prípade vypočítame odhad autokorelácií, a dosadíme ich do Yule-Walkerových rovníc.
- Ak sú dané koeficienty b_i procesu, tak sa z nich dajú vypočítať hodnoty autokorelačnej funkcie tohto procesu.

Príklad 9.7.1 *Vypočítajte prvé tri hodnoty autokorelačnej funkcie pre autoregresný $AR(2)$ proces s rovnicou $X_t - 0.6X_{t-1} - 0.1X_{t-2} = Z_t$.*

Riešenie:

V tomto prípade Yule-Walkerove rovnice vyzerajú takto

$$\begin{aligned} g(1) &= 0.6g(0) + 0.1g(1) \\ g(2) &= 0.6g(1) + 0.1g(0) \end{aligned}$$

Vydělíme obe rovnice g_0 a dostaneme

$$\begin{aligned} r(1) &= 0,6 + 0,1r(1) \\ r(2) &= 0,6r(1) + 0,1 \end{aligned}$$

Odtiaľ vyplýva $r(1) = 2/3$ a $r(2) = 0,5$ a triviálne samozrejme $r(0) = 1$.

9.8 Durbin-Levinsonova rekurzia

Gaussova eliminačná metóda sústavu p rovníc o p neznámych vyrieši v rádovo p^3 krokoch. Avšak zvláštny tvar sústavy Yule-Walkerových rovníc ju umožňuje vyriešiť rýchlejšie Durbin-Levinsonovou rekuriou. Táto metóda je aplikovateľná, keď matica sústavy lineárnych rovníc je toeplitzovská, t.j. ak má matica rovnaké prvky na každej diagonále vedúcej zo severozápadu na juhovýchod.

Matica sústavy Yule-Walkerových rovníc je nielen Toeplitzovská, ale navyše aj symetrická. Konkrétne, Yule-Walkerove rovnice majú tvar

$$\begin{pmatrix} g(1) \\ g(2) \\ \vdots \\ g(p) \end{pmatrix} = \begin{pmatrix} g(0) & g(1) & \dots & g(p-1) \\ g(1) & g(0) & \dots & g(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ g(p-1) & g(p-2) & \dots & g(0) \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix}$$

Príklad 9.8.1 *Vyriešme sústavu*

$$\begin{aligned} g(1) &= g(0)b_1 + g(1)b_2 + g(2)b_3 \\ g(2) &= g(1)b_1 + g(0)b_2 + g(1)b_3 \\ g(3) &= g(2)b_1 + g(1)b_2 + g(0)b_3 \end{aligned}$$

Predpokladajme, že sme vyriešili jednoduchšiu sústavu s dvoma neznámymi x_1, x_2 , ktorá vznikne vynechaním tretej rovnice a tretej neznámej.

$$\begin{aligned} g(1) &= g(0)x_1 + g(1)x_2 \\ g(2) &= g(1)x_1 + g(0)x_2 \end{aligned}$$

Všimnime si, že

$$\begin{pmatrix} g(0) & g(1) \\ g(1) & g(0) \end{pmatrix} \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} g(0)x_2 + g(1)x_1 \\ g(1)x_2 + g(0)x_1 \end{pmatrix} = \begin{pmatrix} g(2) \\ g(1) \end{pmatrix}$$

Preto z prvých dvoch rovníc dostávame

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - b_3 \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$$

Ostáva určiť len b_3 a to dostaneme dosadením do poslednej rovnice:

$$\begin{aligned} g(3) &= (g(2) \ g(1)) \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + g(0)b_3 \\ &= (g(2) \ g(1)) \left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - b_3 \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} \right] + g(0)b_3 \end{aligned}$$

Vo všeobecnosti Durbin-Levinsonov algoritmus definuje postupnosť vektorov zväčšujúcej sa dĺžky

$$\begin{aligned} &(b_{11}) \\ &(b_{21}, b_{22}) \\ &(b_{31}, b_{32}, b_{33}) \\ &\vdots \end{aligned}$$

nasledujúcimi rekurzívnymi vzťahmi

$$\begin{aligned} b_{11} &= \frac{g(1)}{g(0)}, \quad v_0 = g(0) \\ b_{nn} &= \frac{1}{v_{n-1}} \left(g(n) - \sum_{j=1}^{n-1} b_{n-1,j} g(n-j) \right) \\ \begin{pmatrix} b_{n1} \\ \vdots \\ b_{n,n-1} \end{pmatrix} &= \begin{pmatrix} b_{n-1,1} \\ \vdots \\ b_{n-1,n-1} \end{pmatrix} - b_{nn} \begin{pmatrix} b_{n-1,n-1} \\ \vdots \\ b_{n-1,1} \end{pmatrix} \\ v_n &= v_{n-1} [1 - b_{nn}]^2 \end{aligned}$$

Premenné v tomto algoritme majú nasledujúcu interpretáciu (pre každé $n = 1, 2, \dots$):

- vektor (b_{n1}, \dots, b_{nn}) je riešením Yule-Walkerove rovnice n -teho rádu, a teda predstavuje najlepší $AR(n)$ model pre dané autokovariancie,
- v_n je disperzia procesu bieleho šumu pre $AR(n)$ model,
- b_{nn} je hodnota čiastkovej autokorelačnej funkcie.

Príklad 9.8.2 Predpokladajme, že proces X_t má nasledujúce autokovariancie:

$$\begin{aligned} g(0) &= 2.1626356 & g(1) &= 1.1658003 \\ g(2) &= 0.8775780 & g(3) &= 0.7894288 \end{aligned}$$

Počiatočný krok Durbin-Levinsonovej rekurzie definuje

$$b_{11} = \frac{g(1)}{g(0)} = 0.5390646 \quad v_1 = v_0(1 - b_{11}^2) = 1.534194$$

V ďalšej iterácii dostávame

$$\begin{aligned} b_{22} &= \frac{1}{v_1}(g(2) - b_{11}g(1)) = 0.2484740 & v_2 &= v_1(1 - b_{22}^2) = 1.493737, \\ (b_{21}) &= (b_{11}) - b_{22}(b_{11}) = 0.4515264 \end{aligned}$$

Konečne v treťom kroku dostávame

$$\begin{aligned} b_{33} &= \frac{1}{v_2}(g(3) - b_{21}g(2) - b_{22}g(1)) = 0.1364805 \\ v_3 &= v_2(1 - b_{33}^2) = 1.465913 \\ \begin{pmatrix} b_{31} \\ b_{32} \end{pmatrix} &= \begin{pmatrix} b_{21} \\ b_{22} \end{pmatrix} - b_{33} \begin{pmatrix} b_{22} \\ b_{21} \end{pmatrix} = \begin{pmatrix} 0.4293635 \\ 0.1007645 \end{pmatrix} \end{aligned}$$

Z týchto výpočtov vyplýva, že daný proces môžeme modelovať týmito spôsobmi:

- $AR(1)$ procesom $X_t = 0.5390646X_{t-1} + Z_t$, kde $Z_t \sim WN(0, 1.534194)$
- $AR(2)$ procesom $X_t = 0.4515264X_{t-1} + 0.2484740X_{t-2} + Z_t$, kde $Z_t \sim WN(0, 1.493737)$
- $AR(3)$ procesom $X_t = 0.4293635X_{t-1} + 0.1007645X_{t-2} + 0.1364805X_{t-3} + Z_t$, kde $Z_t \sim WN(0, 1.465913)$.

Úloha 9.8.1 Nájdite $AR(4)$ model z Yule-Walkerových rovníc pre predošlý príklad, ak viete, že $g(4) = -0.01218353$.

Úloha 9.8.2 Naprogramujte Durbin-Levinsonov algoritmus, a overte správnosť hodnôt z príkladu 9.8.2 a úlohy 9.8.1.

9.9 Woldova dekompozícia

Modely stacionárnych procesov $AR(p)$ a $MA(q)$ dokážu namodelovať širokú škálu náhodných procesov. Vyvstáva otázka, či existujú iné stacionárne náhodné procesy, ktoré sa takýmito modelmi nedajú vystihnúť. Začneme ukážkou jedného takého procesu.

Nech A, B sú nekorelované náhodné premenné so strednou hodnotou 0, a varianciou σ^2 . Definujme proces $Y_t = A \cos(\omega t) + B \sin(\omega t)$. Hoci tento proces je náhodný,

má zaujímavú vlastnosť. Keď vieme dve jeho po sebe idúce hodnoty, vieme určiť všetky nasledujúce, pretože platí

$$Y_n = (2 \cos \omega)Y_{n-1} - Y_{n-2}.$$

Takýto proces sa nazýva *deterministicky predikovateľný*. Všeobecne, ak sa hodnoty procesu dajú presne (lineárne) predpovedať na základe predošlých hodnôt, tak tento nazveme deterministicky predikovateľný.

Úloha 9.9.1 *Vypočítajte autokovariačnú a autokorelačnú funkciu procesu Y_t .*

Woldova dekompozícia hovorí, že ľubovoľný nedeterministický stacionárny proces X_t sa dá rozložiť na súčet

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j} + V_t,$$

kde

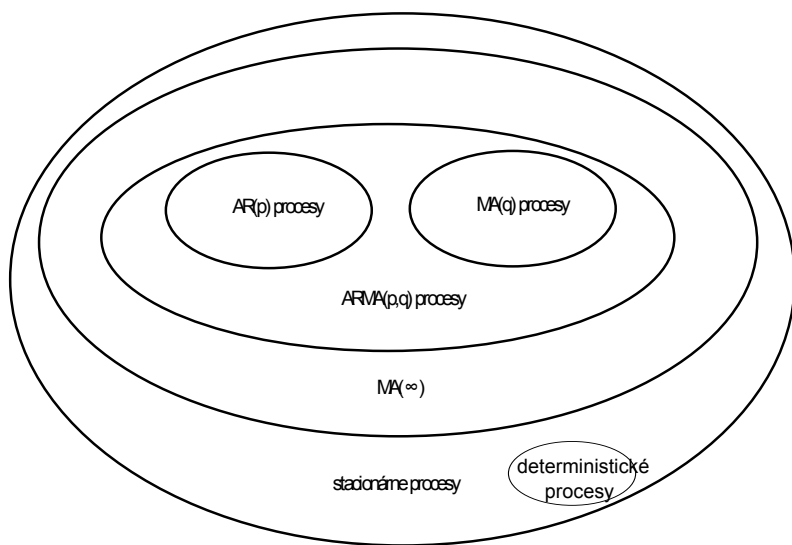
1. $b_0 = 1$ a $\sum_{j=0}^{\infty} \psi_j^2 < \infty$,
2. $Z_t \sim WN(0, \sigma^2)$,
3. $\text{Cov}(Z_s, V_t) = 0$ pre všetky časy s, t ,
4. V_t je deterministicky predikovateľný proces.

Prvá zložka vo Woldovom rozklade, tj. $\sum_{j=0}^{\infty} \psi_j Z_{t-j}$ sa nazýva aj $MA(\infty)$ procesom, keďže je zjavným zovšeobecnením $MA(q)$ procesov z konečných na nekonečné kľzavé súčty.

Woldova dekompozícia sa dá interpretovať tak, že $MA(q)$ modely dokážu dobre aproximovať ľubovoľný stacionárny proces, ktorý nemá deterministicky predikovateľnú zložku. V praxi sa osvedčilo použitie zovšeobecnených $ARMA(p, q)$ modelov. V týchto modeloch proces X_t spĺňa rovnicu

$$X_t - b_1 X_{t-1} - \dots - b_p X_{t-p} = Z_t + a_1 Z_{t-1} + \dots + a_q Z_{t-q}$$

Väčšina štatistických softvérov dokáže nájsť najvhodnejší $ARMA$ model na základe štatistických kritérií, ako je napríklad Akaikeho kritérium.



Obr. 9.10. Vzťah medzi kategóriami procesov

10 Matematický softvér

V súčasnosti sú programy Matlab, R a Mathematica najpoužívanějšíe softvéry vo vedeckej i technickej praxi. V tomto apendixe si ukážeme použitie týchto programov na nasledujúce úlohy:

1. výpočet veľkosti lineárnej kombinácie dvoch vektorov,
2. riešenie sústavy lineárnych rovníc (riešiteľnej ale aj neriešiteľnej),
3. výpočet lineárnej regresie pomocou lineárnej a kvadratickej funkcie,
4. výpočet PCA
5. simulácia $MA(1)$ procesu, tj. procesu, ktorý vznikne aplikáciou kĺzavých súčtov na biely šum a výpočet autokorelačnej funkcie výsledného procesu.

Pre tretiu úlohu predpokladáme, že dáta, ktoré chceme namodelovať, sú uložené v ASCII súbore `data.txt`, v takzvanom CSV (comma separated values) formáte:

```
x,y
2.6,4.8
4.1,2.4
8.2,3
10.3,7.2
13.3,10.9
14.4,12.4
16,15.8
20.6,30
23,34.4
24.9,43.6
```

Tento formát je podporovaný najväčšou škálou softvérov, a tak je uprednostňovaný napríklad aj pred excelovskými súbormi. Excel, ale aj iné softvéry, má možnosť exportovať dáta do tohto formátu. Treba si však dať pozor na lokalizáciu. Niektoré lokalizácie (vrátane slovenskej) používajú čiarku v úlohe desatinnej čiarky a používajú iný symbol, napríklad bodkočiarku, namiesto čiarky na oddeľovanie dát v riadkoch.

10.1 Matlab

Matlab je softvér, ktorý je veľmi populárny medzi inžiniermi. Vydáva ho firma MathWorks, ale existujú aj voľne širiteľné (open-source) programy Octave (www.gnu.org/software/octave) a SciLab (www.scilab.org), ktoré používajú rovnakú syntax ako Matlab. Má veľmi dobrú podporu na prácu s vektormi, maticami a takisto sa ľahko rozširuje pomocou C knižníc.

10.1.1 Základná práca s vektormi

Skúsme vypočítať veľkosť lineárnej kombinácie $3v_1 + 2v_2$ vektorov $v_1 = (3, -1, 4)$ a $v_2 = (1, 1, 0)$. Výsledok týchto príkazov možno vidieť na nasledujúcom obrázku

```
>> v1 = [3 -1 4]

v1 =

     3     -1     4

>> v2 = [1 1 0]

v2 =

     1     1     0

>> v3 = 3 * v1 + 2 * v2

v3 =

    11     -1    12

>> sqrt(v3(1)^2 + v3(2)^2 + v3(3)^2)

ans =

    16.3095
```

Všimnime si, že Matlab vypisuje všetky medzivýsledky. Ak príkazy ukončíme bodkočiarkou, tieto medzivýsledky nie sú zobrazované. Posledný výpočet sa dal spraviť aj týmito alternatívnymi príkazmi

- `sqrt(v3 * v3')`
- `sqrt(sum(v3 .* v3))`
- `norm(v3)`

10.1.2 Riešenie sústavy rovníc

Ak má sústava riešenie, stačí vynásobiť ľavú stranu inverznou maticou.


```
A = [1 0 1;
      0 1 1;
      1 1 0];
y = [3;4;5];
x = inv(A) * y
```

Ak sústava nemá riešenie, používa sa operátor delenia maticou.

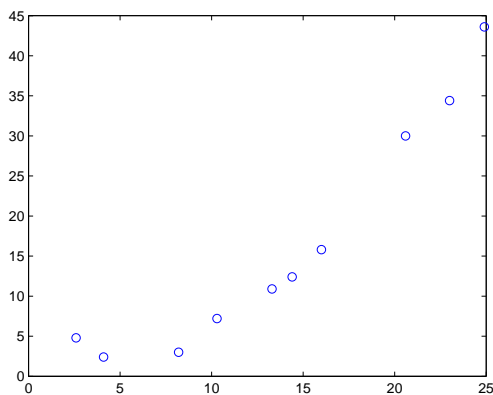
```
A = [1 0 1;
      0 1 1;
      1 1 0;
      1 1 1]
y = [3;4;5;13];
x = A \ y
```

10.1.3 Metóda najmeních tvorcov

Na načítanie dát môžeme použiť príkaz `textread`, ktorý načíta dáta do dvoch stĺpcových vektorov `x,y`.

```
[x,y]= textread('data.txt', '%f,%f', 'headerlines', 1)
plot(x,y, 'o')
```

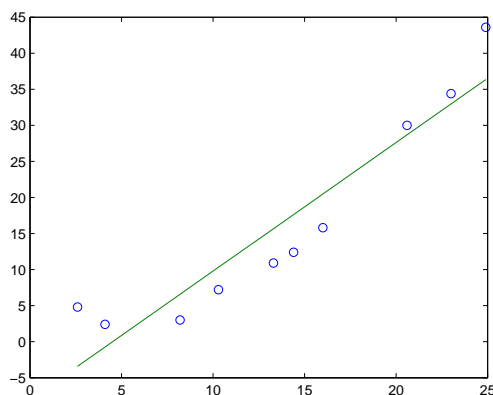
Výsledkom je obrázok 10.1.3.



Obr. 10.1. Vykreslenie nameraných hodnôt

Na výpočet regresie môžeme použiť funkciu `polyfit`, ktorá nájde najlepšiu aproximáciu polynómom daného stupňa.

```
p1 = polyfit(x,y,1)
plot(x,y, 'o', x, polyval(p1,x), '-')
```



Obr. 10.2. Lineárna regresná krivka

Na obrázku 10.2 je zobrazený výsledok lineárnej regresie.

Zmena pre nájdenie najlepšej aproximácie pomocou kvadratického polynómu je triviálna

```
p2 = polyfit(x,y,2)
plot(x,y, 'o', x, polyval(p2,x) , '-')
```

10.1.4 Výpočet PCA

Nech je daná matica M , v ktorej sú dáta uložené po riadkoch. Prvým krokom PCA je vytvorenie matice M° , ktorej stĺpce majú priemer 0.

```
M0 = M-ones(length(M),1) * mean(M)
```

Ak dáta v jednotlivých stĺpcoch majú veľmi rozdielny rozptyl, odporúča sa upraviť maticu tak, aby jednotlivé stĺpce mali jednotkový rozptyl.

```
M0 = (M-ones(length(M),1) * mean(M)) * inv(diag(std(M)))
```

Posledný výpočet sa dá vykonať aj funkciou `zscore` z balíčka *Statistics toolbox*.

Na výpočet vektorov PCA je možné použiť funkciu `eig` používanú na výpočet hlavných vektorov, avšak výsledné poradie vektorov nemusí zodpovedať dôležitosti. Preto je vhodnejšie použiť funkciu `svd`, ktorá vektory zoradí podľa dôležitosti. Obidva z nasledujúcich príkazov

```
[U, S, V] = svd (M0' * M0)
[U, S, V] = svd (cov(M0))
```

vypočítajú rozklad kovariančnej matice $\text{Cov}(M^\circ) = M^{\circ T} M^\circ$. Na výpočet prvých dvoch najdôležitejších súradníc môžeme použiť príkaz

```
M0 * V(:, 1:2)
```

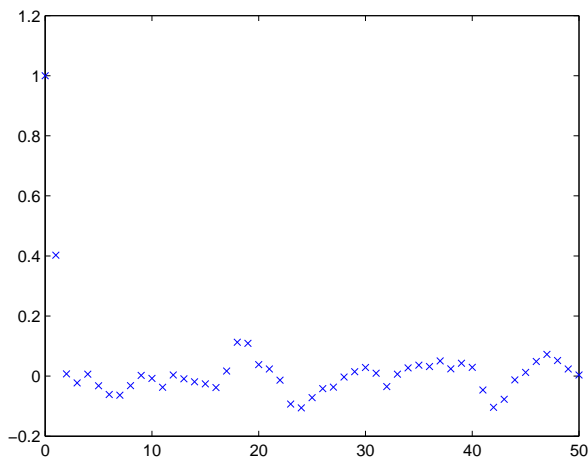
Alternatívne možno použiť funkciu `princomp` z balíčka *Statistics toolbox*. Matlab má tiež funkciu `biplot` umožňujúcu názorné vykreslenie dát v súradniciach danými prvými dvoma hlavnými komponentmi.

10.1.5 Autokorelácia $MA(1)$ procesu.

Vygenerujme si 1001 výberov z bieleho šumu, ktorý vygenerujeme pomocou gaussovského rozdelenia $N(0, 1)$.

```
bielySum = randn(1, 1001);
maProces = bielySum(1:1000) + 0.5 * bielySum(2:1001);
for i = 0:50
    kovMatica = cov(maProces(1:(1000-i)), ...
                    maProces((i+1):1000));
    autoKorelacia(i+1) = kovMatica(1,2) / ...
                        sqrt(kovMatica(1,1) * kovMatica(2,2));
end
plot(0:50, autoKorelacia, 'x')
```

Výsledkom je graf, ktorý je na obrázku 10.3. Je to prvých 51 autokorelácií danej realizácie procesu.



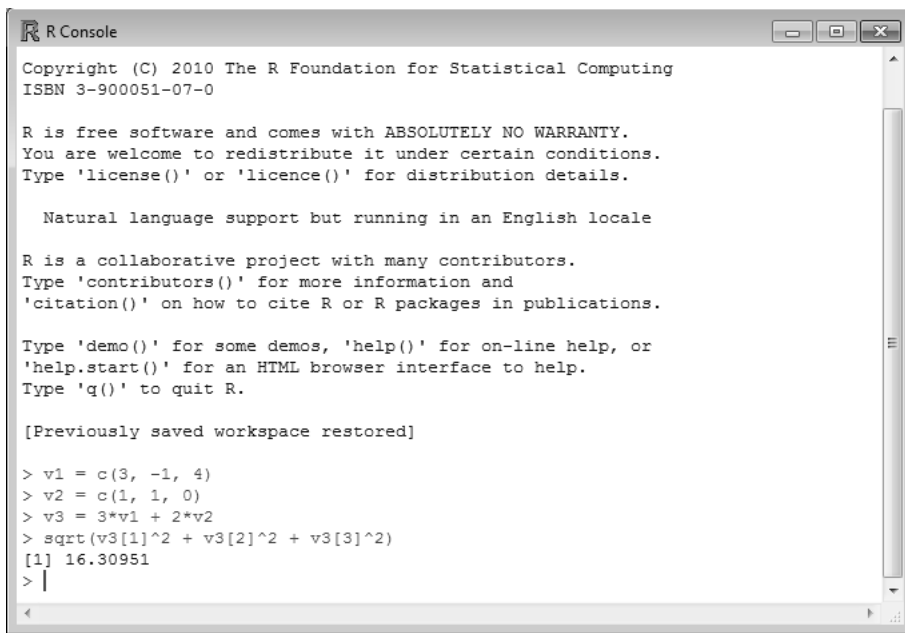
Obr. 10.3. Prvých 51 autokorelácií danej realizácie procesu

10.2 R

R je open-source softvér (www.r-project.org) používaný najmä štatistikmi. Má o poznanie menej pohodlné značenie pre prácu s vektormi a maticami, no na druhej strane je silnejší v pravdepodobnosti, štatistike a tvorbe názorných obrázkov vyjadrujúcich štatistické charakteristiky dát.

10.2.1 Základná práca s vektormi

Skúsme vypočítať veľkosť lineárnej kombinácie $3v_1 + 2v_2$ vektorov $v_1 = (3, -1, 4)$ a $v_2 = (1, 1, 0)$. Výsledok týchto príkazov možno vidieť na nasledujúcom obrázku:



```
R Console
Copyright (C) 2010 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Previously saved workspace restored]

> v1 = c(3, -1, 4)
> v2 = c(1, 1, 0)
> v3 = 3*v1 + 2*v2
> sqrt(v3[1]^2 + v3[2]^2 + v3[3]^2)
[1] 16.30951
> |
```

Všimnime si, že na tvorbu vektorov sa používa funkcia `c()`, ktorá spája svoje argumenty. Na rozdiel od Matlabu vektory nie sú matice a treba použiť funkcie `matrix`, `as.matrix`, ak s nimi chceme pracovať ako s maticami. Program R, znova na rozdiel od Matlabu, vypisuje iba tie medzivýsledky, ktoré sa nepriradia do nejakej premennej. Ak sme zabudli priradiť výsledok výpočtu, nájdeme ho v preddefinovanej premennej `.Last.value`. Veľa funkcionality programu R je v knižniciach, ako napríklad aj `norm()` funkcia, ktorá sa nachádza v knižnici `Matrix`. Ak ju chceme využiť, musíme príslušnú knižnicu najprv načítať pomocou príkazu `library(Matrix)`, alebo použiť celý názov funkcie `Matrix::norm`.

```
library(Matrix)
v1 = c(3, -1, 4)
v2 = c(1, 1, 0)
```

```
v3 = 3*v1 + 2*v2
norm (as.matrix(v3), "F")
```

10.2.2 Rieenie sústavy rovníc

Ak má sústava riešenie, môžeme použiť funkciu `solve()`.

```
A = matrix(c(1, 0, 1,
             0, 1, 1,
             1, 1, 0), nrow = 3)
y = matrix(c(3, 4, 5), ncol = 1)
x = solve(A, y)
```

Ak sústava nemá riešenie, používa sa funkcia `qr.solve()`.

```
A = matrix ( c(1,0, 1,
              0, 1, 1,
              1, 1, 0,
              1, 1, 1), nrow = 4)
y = c(3, 4, 5, 13);
x = qr.solve(A, y)
```

10.2.3 Metóda najmeních tvorcov

Na načítanie dát môžeme použiť príkaz `read.csv`, ktorý načíta dáta do dvoch stĺpcových vektorov `x`, `y`.

```
data1 = read.csv('data.txt')
plot(y ~ x, data = data1)
```

Výsledkom je obrázok 10.4.

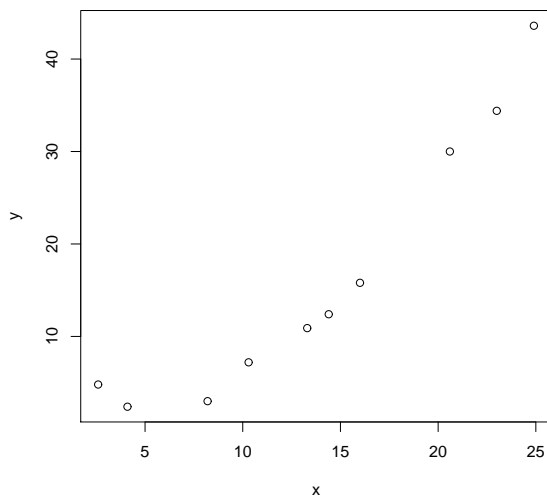
Na výpočet lineárnej regresie sa používa funkcia `lm`.

```
model1 = lm( y ~ x, data = data1)
plot(y ~ x, data = data1)
abline(model1)
```

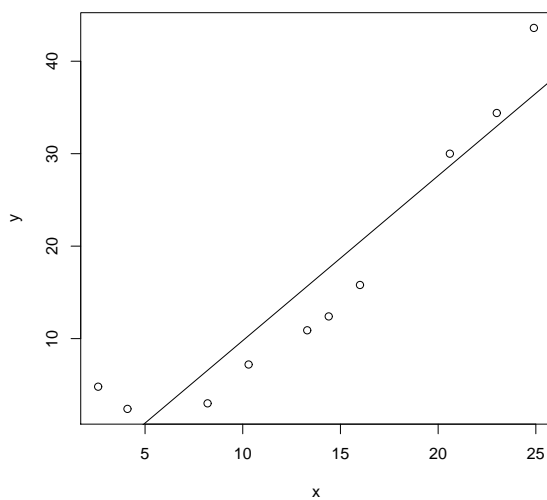
Na obrázku 10.5 je zobrazený výsledok lineárnej regresie.

Zmena pre nájdenie najlepšej aproximácie pomocou kvadratického polynómu je triviálna

```
model2 = lm ( y ~ x + I(x^2), data = data1)
plot(y ~ x, data = data1)
lines(data1$x, predict(model2))
```



Obr. 10.4. Načítanie dát v R



Obr. 10.5. Lineárna regresia v R

10.2.4 Výpoet PCA

Nech je daná matica M , v ktorej sú dáta uložené po riadkoch. Na rozklad matice podľa singulárnych hodnôt sa používa funkcia `svd`. Avšak PCA môžeme spraviť priamo, bez použitia tejto funkcie jedným z nasledujúcich príkazov

```
pcaData = prcomp(M)
pcaData = prcomp(M, scale = TRUE)
```

Prvý z príkazov sa používa, keď rozptyl jednotlivých premenných je približne rovnaký. Ak tento predpoklad nie je splnený, odporúča sa použiť druhý príkaz, ktorý zodpovedá hľadaniu vlastných vektorov korelačnej matice. Výsledkom je objekt (štruktúra) s nasledujúcimi zložkami:

- `pcaData$sdev` rozptyly hlavných komponentov (tj. druhé odmocniny vlastných čísel kovariančnej (v druhom príkaze korelačnej) matice)
- `pcaData$rotation` matica, ktorej stĺpce sú vlastné vektory
- `pcaData$x` dáta v transformovaných súradniciach
- `pcaData$center` stredné hodnoty premenných
- `pcaData$scale` rozpyly premenných

Táto štruktúra sa dá názorne vykresliť pomocou funkcie `biplot(pcaData)`. Krátky sumár sa dá zobrazíť zavolaním `summary(pcaData)`.

10.2.5 Autokorelácia $MA(1)$ procesu

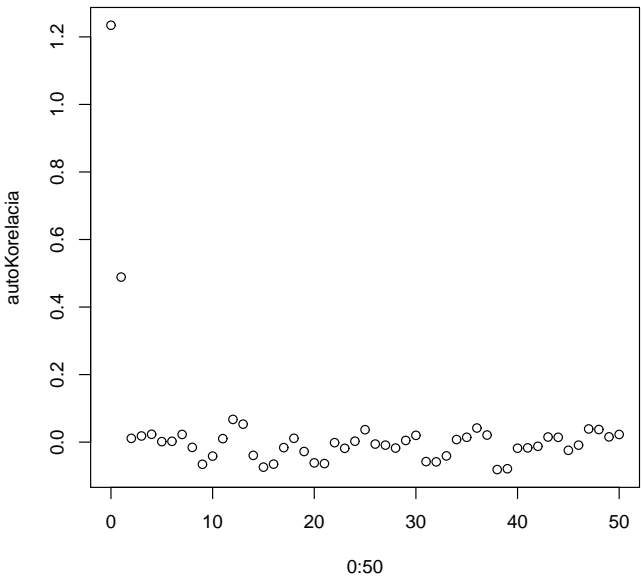
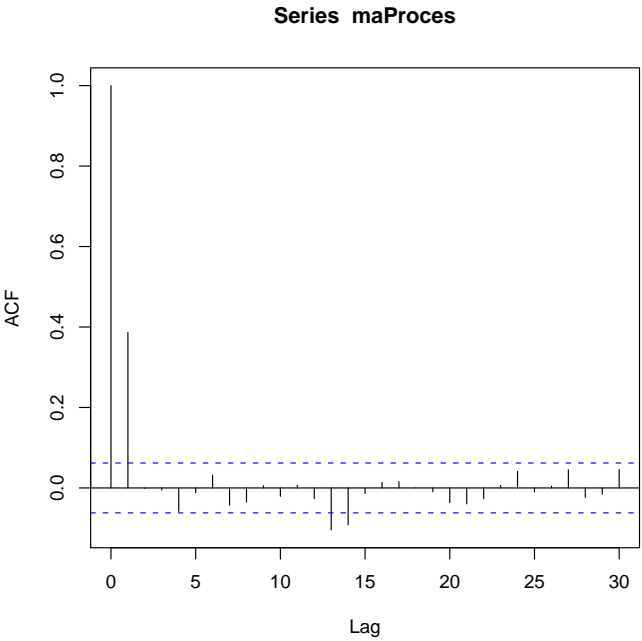
Pri tejto úlohe sa ukáže silná podpora pre pravdepodobnostné a štatistické výpočty. Táto úloha sa dá vykonať v R dvoma riadkami.

```
maProces = arima.sim (list (ma = 0.5) , 1001)
plot (acf (maProces))
```

Na porovnanie s Matlabom a Mathematicou si ukážme, ako sa táto simulácia dá urobiť bez použitia funkcií `arima.sim` a `acf`.

```
bielySum = rnorm(1001);
maProces = bielySum[-1001] + 0.5 * bielySum[-1];
autoKorelacia = rep(0,51)
for (i in 0:50){
  autoKorelacia[i + 1] = cov(maProces[1:(1000 - i)] ,
                             maProces[(i + 1):1000]);
}
plot(0:50, autoKorelacia)
```

Na nasledujúcom obrázku sú grafy autokorelačných funkcií vypočítané aj prvým aj druhým prístupom.

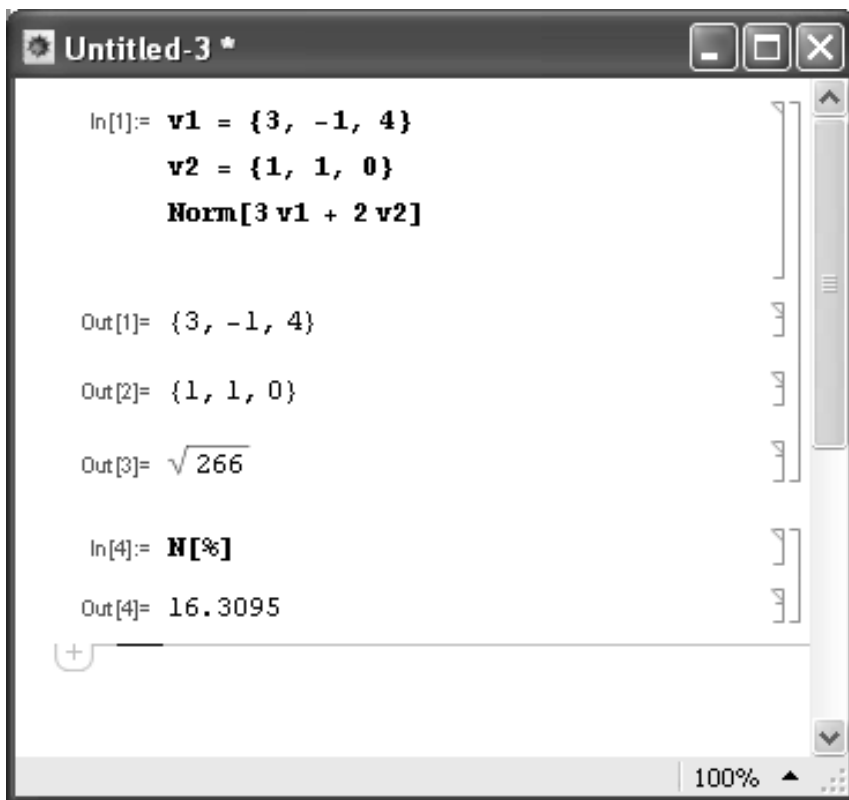


10.3 Mathematica

Mathematica je komerčný softvér určený predovšetkým pre výskumníkov. Vydáva ho firma Wolfram Research, ktorá stojí aj za portálom Wolfram Alpha. Tento portál umožňuje aj laikom zadať netriviálne otázky, a v pozadí softvér Mathematica vypočíta odpoveď.

10.3.1 Základná práca s vektormi

Skúsme vypočítať veľkosť lineárnej kombinácie $3v_1 + 2v_2$ vektorov $v_1 = (3, -1, 4)$ a $v_2 = (1, 1, 0)$. Výsledok týchto príkazov možno vidieť na nasledujúcom obrázku



Podobne ako Matlab, Mathematica vypisuje všetky medzivýsledky, ktoré sa nekončia bodkočiarkou. Odlišnosťou softvéru Mathematica je, že dokáže počítvať presné výsledky, a tak občas treba použiť funkciu `N[]`, aby sme ich skonvertovali na reálne číslo s desatinnou čiarkou.

10.3.2 Riešenie sústavy rovníc

Ak má sústava riešenie, môžeme použiť funkciu `LinearSolve[]`.

```
LinearSolve[ {{1,0,1},
              {0,1,1},
              {1,1,0}},
            {3,4,5}]
```

Ak sústava nemá riešenie, treba použiť funkciu `LeastSquares`.

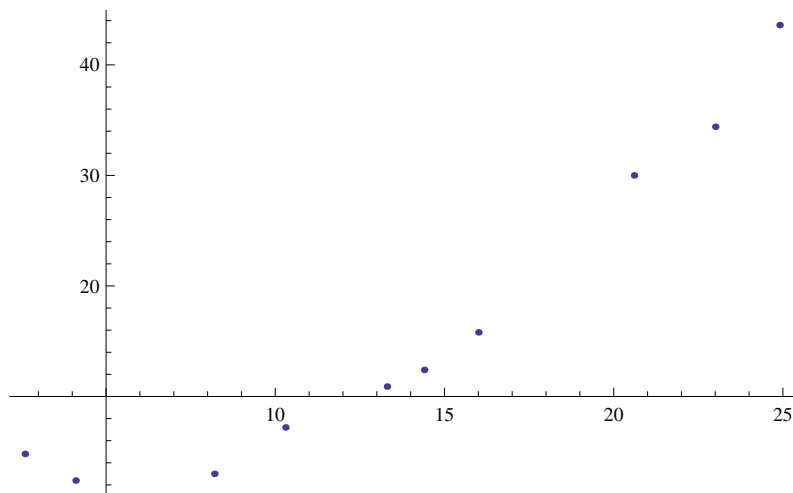
```
LeastSquares[ {{1,0,1},
               {0,1,1},
               {1,1,0},
               {1,1,1}},
             {3,4,5,13}]
```

10.3.3 Metóda najmeních tvorcov

Na načítanie dát môžeme použiť príkaz `Import`, ktorý načíta dáta do premennej `data1`, nasledovaný funkciou `Drop`, aby sme odstránili prvý riadok, ktorý obsahuje názvy stĺpcov.

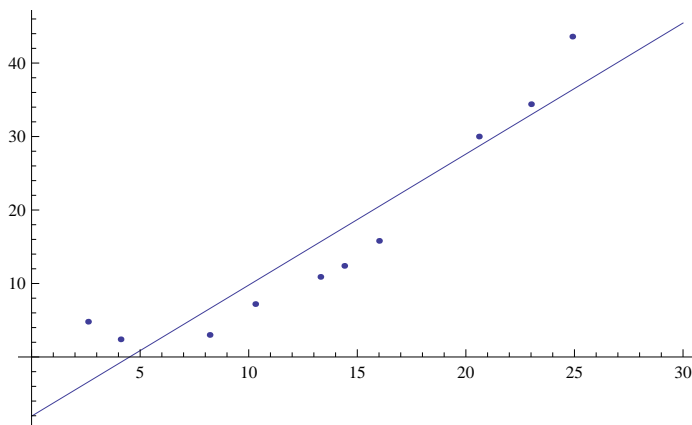
```
data1 = Import["data.txt", "CSV"]
data2 = Drop[data1, 1]
ListPlot[data2]
```

Výsledkom je nasledujúci obrázok.



Na výpočet regresie môžeme použiť funkciu `LinearModelFit`, ktorá nájde najlepšiu aproximáciu polynómom daného stupňa.

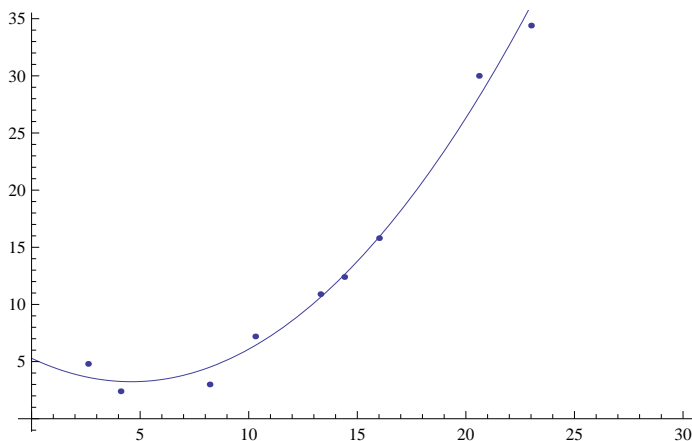
```
model = LinearModelFit[data2, x, x]
Show[ListPlot[data2], Plot[model["BestFit"], {x, 0, 30}]]
```



Tu je zobrazený výsledok lineárnej regresie.

Aby sme našli najlepšiu parabolu, treba použiť funkciu **Fit**.

```
parabola = Fit[data2, {1, x, x^2}, x]
Show[ListPlot[data2], Plot[parabola, {x, 0, 30}]]
```



10.3.4 Výpoet PCA

Nech je daná matica M , v ktorej sú dáta uložené po riadkoch. Na rozklad matice podľa singulárnych hodnôt sa používa funkcia **SingularValueDecomposition**. Avšak PCA môžeme spraviť bez použitia tejto funkcie jedným z nasledujúcich príkazov:

```
pcaData = PrincipalComponents(M, Method -> Covariance)
pcaData = PrincipalComponents(M, Method -> Correlation)
```

Prvý z príkazov sa používa, keď rozptyl jednotlivých premenných je približne rovnaký. Ak tento predpoklad nie je splnený, odporúča sa použiť druhý príkaz, ktorý zodpovedá hľadaniu vlastných vektorov korelačnej matice.

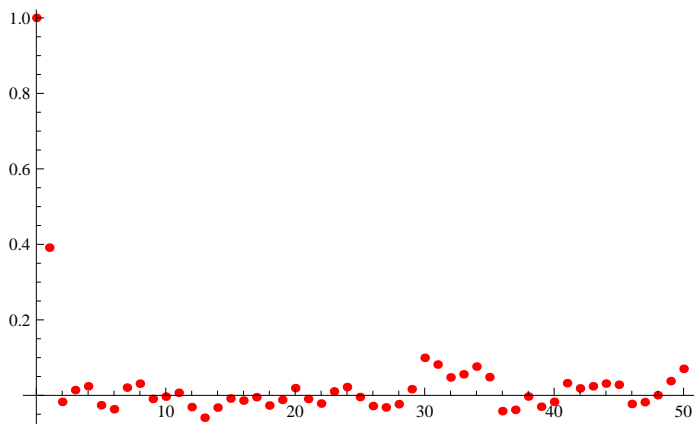
Výsledkom je zoznam vektorov zoradený klesajúco podľa dôležitosti.

10.3.5 Autokorelácia $MA(1)$ procesu

Vygenerujme si 1001 výberov z bieleho šumu, ktorý vygenerujeme pomocou Gaussovho rozdelenia $N(0, 1)$.

```
bielySum = RandomVariate[NormalDistribution[], 1001];
maProces = bielySum[[Range[1, 1000]]]
          + 0.5*bielySum[[Range[2, 1001]]];
autoKorelacia = Table[{i, 0}, {i, 0, 50}];
For[i = 0, i <= 50, i++,
  autoKorelacia[[i + 1, 2]] =
    Correlation[maProces[[Range[1, 1000 - i]]],
      maProces[[Range[i + 1, 1000]]]];
ListPlot[autoKorelacia, PlotRange -> Full,
  PlotStyle -> Directive[Red, PointSize[Medium]]]
```

Výsledkom je graf prvých 51 autokorelácií danej realizácie procesu.



Literatúra

- [1] AXLER, S.: *Linear Algebra Done Right*, Second Edition 1991, Springer-Verlag, New-York, Berlin, Heidelberg, ISBN 0-387-98259-0
- [2] BOROŇ, J.: *Kvalita reči pre IP telefóniu pri rozklade do rôznych báz, diplomová práca*, Fakulta riadenia a informatiky, Žilinská univerzita 2007
- [3] BROCKWELL P.J., DAVIS R. A.: *Introduction to Time Series and Forecasting*, 2nd ed. (Springer texts in statistics) 2002
- [4] BROCKWELL P.J., DAVIS R. A.: *Time series: Theory and Methods*, Springer texts in statistics, 1987
- [5] COOLEY J.W., TUKEY J.W.: *An algorithm for machine calculation of complex Fourier series*, *Math. Comp.* 19 (1965), 297–301
- [6] ČÁKY, P., KLIMO, M., MIHÁLIK, I., MLADŠÍK, R.: *Text-to-speech for Slovak language*, Text, speech and dialogue, 2004, Brno, Czech Republic, Springer 2004, ISBN 3-540-23049-1
- [7] FRIGO M., JOHNSON S. G.: *The Fastest Fourier Transform in the West*, MIT technical report MIT-LCS-TR-728, 1997
- [8] HAMMING R.W.: *Error detecting and error correcting codes*, *Syst. Tech. J.*, 29, (1950), 147–160
- [9] HASSANIEH H., INDYK P., KATABI D., PRICE E.: Nearly optimal sparse Fourier transform, <http://arxiv.org/abs/1201.2501>
- [10] KLARNER, D.A.: *The Mathematical Gardner*, Wadsworth International Belmont, California 1981
- [11] KLIMO, M., BACHRATÁ, K.: *Mathematical education reform for ICT Students*, ICETA 2008
- [12] KUBA, M.: *Úvod do problematiky rozpoznávania reči*, TKR - QUO VADIS?, Žilina 2001 , ISBN 80-7100-826-5
- [13] KUBA, M.: *Rozpoznávanie rečových signálov z malého počtu akustických pozorovaní pomocou skrytých Markovových modelov*, (dizertačná práca), školiteľ: Martin Vaculík. – Žilinská univerzita 2006

- [14] RADER, C.: *Discrete Fourier transforms when the number of data samples is prime*, *IEEE Proc.* **56**, (1968), 1107–1108
- [15] MIHÁLIK I., ŠKVAREK O., ČÁKY P., JANDÁK V.: *Spracovanie textu pre rečový syntetizátor a jazyk SSML*, INFOTRANS 2007, Vydavateľstvo: Univerzita Pardubice 2007, Pardubice, Česká republika. ISBN 978-80-7194-989-3
- [16] KLIMO, M., KOVÁČIKOVÁ, T.: *Voice over Packet*, module No 12 of the linebreak eEDUSER, No /02/B/F/PP/142256 Košice: Elfa 2005, ISBN 80-89066-99-2
- [17] KOVÁČIKOVÁ T., SEGEČ P.: *IP telephony for an interactive communication within e-Learning*, In: ICETA 2003: Košice, Slovak Republic: Košice: Elfa 2003
- [18] KRŠÁK, E.: *Elektronický podpis*, Zborník prednášok na medzinárodnú konferenciu Systémová integrácia 2001, SSSI Žilina 2001, ISBN 80-7100-880-X
- [19] PAPOULIS, A.: *Probability, Random Variables, and Stochastic Processes*, Polytechnic Institute of New York, McGraw-Hill, Inc., New York, Toronto London 1991
- [20] PRATCHETT T., STEWART I., COHEN J.: *Věda Na Zeměploše*, Talpress, ISBN: 80-7197-243-6, EAN: 9788071972433, The Science of Discworld 2004
- [21] PALÚCH, S., STANKOVIANSKA, I.: *Algebra a jej inžinierske aplikácie*, EDIS 2008, Žilina
- [22] WALTER, J.: *Stochastické modely v ekonómii*, SNTL Praha 1970
- [23] WIDROW, B., MC COOL, J.M., LARIMORE, M.G., JOHNSON, C.R., jr.: *Stationary and nonstationary learning characteristics of the LMS adaptive filter*, *Proceedings of the IEEE*, 64, 8, (1976), 1151–1162
- [24] WINOGRAD S.: *On computing the discrete Fourier transform*, *Mathematics of Computation*, 32, 141, (1978), 175–199
- [25] ZIMMERMANN, J.: *Spektrografická a škálografická analýza akustického rečového signálu*, Náuka, Prešov, 2002, ISBN 80-89038-22-0
- [26] ZLATOŠ, P.: *Lineárna algebra a geometria Lineárna algebra a geometria, Cesta z troch rozmerov s presahmi do príbuzných odborov*, Marenčin PT 2011, ISBN 978-80-8114-111-9

Register

- AR(p) proces, 203
 - AR(1) proces, 202–203
- autokorelačná funkcia, 194
- autokovariančná funkcia, 194
- báza
 - Walsh-Hadamardova, 119
- báza vektorového priestoru, 23
 - jednotková, 25
 - ortogonálna, 31
 - zmena bázy, 27
- biely šum, 199
- chybový proces, 48
 - energia, 49
 - veľkosť, 48
- cyklický kód, 182
- diskrétna Fourierova transformácia, 94,
104
 - spätná, 104
- distribučná funkcia, 146
- Durbin-Levinsonov algoritmus, 206–208
- extrapolácia procesu, 83
- filtrácia procesu, 80
- generujúci polynóm, 185
- Gramova-Schmidtova ortogonalizácia, 38–
42
- Hammingova vzdialenosť, 35, 179
- harmonická báza, 94
- impulzná charakteristika systému, 160
- interpolácia procesu, 77
- kódy
 - blokové, 182
 - nesystematické, 182
- konvolúcia, 165
- korelačný koeficient, 147
- kovariancia, 147
- lineárna kombinácia, 20
- lineárna regresia, 50–71, 187–193
 - výpočet v Mathematice, 222
 - výpočet v Matlabe, 213
 - výpočet v R, 217
- lineárny prediktor, 195
- lineárny systém, 160
- MA(q) proces, 201
 - MA(1) proces, 199–202
 - simulácia v Mathematice, 224
 - simulácia v Matlabe, 215
 - simulácia v R, 219
- Mathematica, 221–224
- matica patriaca báze, 26
- Matlab, 212
- metóda
 - kľzavých súčtov, 80
 - najmenších štvorcov, 47–48
- ML metóda, 188
- modulácia, 161
- odľahlá hodnota, 192
- PCA, 136–137
- priemet vektora, 37
- proces
 - deterministický, 147
 - energia, 48
 - náhodný, 147

- realizácia, 147
- Pytagorova veta, 34, 49, 150
- R, 216–219
- súradnice vektora, 26
- sústava lineárnych rovníc
 - riešenie v Mathematice, 221
 - riešenie v Matlabe, 212
 - riešenie v R, 217
- skalár, 14
- skalárny súčin, 31, 97
- spektrum
 - amplitúdové, 104
 - fázové, 104
- spektrum procesu, 28, 104
- stacionárny proces, 194
 - centrovaný, 194
- stredná hodnota, 146
- transformácie časovo invariantné, 164
- vektor, 14
 - syndrómový, 175
 - veľkosť vektora, 33
- vektorový priestor, 13–17
 - dimenzia, 24
 - Euklidov, 31
 - podpriestor vektorového priestoru,
 - 25
 - príklady, 16–18
- vektory
 - bázové, bázy, 24
 - generujúce podpriestor, 25
 - kolmé, 31
 - lineárne nezávislé, 22
 - lineárne závislé, 21
 - vzdialenosť vektorov, 34
- vierohodnosť, 188
- Yule-Walkerove rovnice, 205