ROZHODOVÁNÍ A ŘÍZENÍ

v dopravě

Část I. Základy

Obsah

Předmluva

Předkládaná zpráva je prvním dílem řady zpráv pojednávajících o využití výsledků bayesovské statistiky pro automatické řízení v dopravě. Tato první zpráva je zamýšlena jako úvod pro ty, kteří se s touto oblastí setkávají poprvé. Podává motivaci pro modelování a identifikaci dopravních systémů, ukazuje základní úlohy identifikace a optimálního řízení pro deterministické modely s poruchou a dává přehled přes základní pravděpodobnostní pojmy, zejména ty, které se využívají v bayesovské statistice. Předpokládají se znalosti matematiky na úrovni standardního základního kurzu na technické vysoké škole.

Jako pokračování bude následovat pojednání o pravděpodobnostních modelech v bayesovské statistice včetně základů řešení bayesovské identifikace a optimálního řízení pro tyto pravděpodobnostní modely.

Další pokračování se bude zabývat speciálními modely, jako jsou modely využívající spline-aproximaci hladkých funkcí nebo směsové modely, a jejich identifikací.

Tyto tři zprávy budou představovat základní část, na ni pak budou navazovat pojednání o konkrétních dopravních problémech a jejich řešení, založené na bayesovském přístupu k identifikaci a řízení.

Zpráva je určena především studentům přicházejícím po prvé do styku s problematikou automatického řízení. Má jim pomoci překonat základní bariéru, kterou vytváří jednak poněkud rozšířený matematický aparát užívaný v této oblasti, ale zejména jeho použití, které se může běžnému studentu může zdát příliš umělé.

Komentář k textu:

Poznámky jsou tištěny odlišným typem písma a mají za úkol průběžně komentovat vykládaný text. Jsou odlišeny proto, aby bylo dobře možno je při dalším čtení přeskakovat, a tak zlepšit souvislost hlavní linie výkladu.

Odstavce označené znakem "piky" vyjadřují důležité poznatky. V zájmu snadnějšího zapamatování jsou často vyjádřeny "lidovou mluvou".

Cvičení mají za úkol upozornit na zajímavé problémy, nebo dávají čtenáři příležitost ověřit si porozumění vykládané problematice.

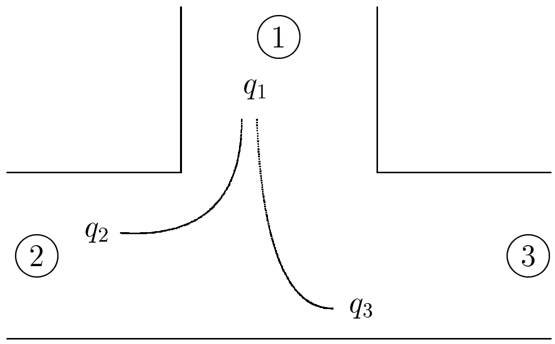
Kapitola 1

Matematické modely v dopravě

Jako motivaci pro naše pojednání o Bayesovském odhadování a řízení v dopravě uvedeme jednoduchý příklad. Jeho úkolem je volně nastínit problematiku a ilustrovat základní pojmy, se kterými se dále budeme setkávat.

1.1 Příklad dopravní úlohy

Uvažujme křižovatku tvaru T podle obrázku.



Automobily do ní přijíždějí po silnici 1 a odbočují buď doprava do silnice 2, nebo doleva do silnice 3. Ve vyznačených místech měříme *intenzitu provozu* q, tj. počet vozidel, která daným místem projedou za jednu minutu. Naším cílem je identifikovat podíl α vozidel odbočujících doprava, nebo jinak řečeno, určit pravděpodobnost α, že auto náhodně přijíždějící po silnici 1 odbočí doprava.

Měření provádíme každou celou minutu a časové okamžiky měření označíme přirozenými čísly 1, 2,..., t. Získáme tak data (posloupnost naměřených hodnot)

$$q_{i,\tau} \quad \text{pro} \quad i = 1, 2, 3 \;\text{ a }\; \tau = 1, 2, ...t. \tag{1.1}$$

Poznámka: Stačilo by měřit libovolné 2 veličiny, protože třetí je jimi jednoznačně určená. Platí totiž $q_{1,t} = q_{2,t} + q_{3,t}$, pro všechna t.

Nabízí se několik řešení se stejným výsledkem ale různou obecností.

1.2 Způsoby řešení úlohy

I. Odhad pomocí statistické definice pravděpodobnosti

Náhodným pokusem je průjezd vozidla křižovatkou. Pokus má dva možné výsledky: odbočení doprava nebo doleva. Sledovaný jev je odbočení doprava. Jeho pravděpodobnost α je při velkém počtu pokusů dána podílem počtu příznivých výsledků (počtu odbočení doprava) a počtu pokusů (celkového počtu vozidel, která křižovatkou projela). Protože počet automobilů projíždějících křižovatkou měříme průběžně, můžeme poměr vozidel odbočujících doprava vyjádřit jako zlomek

$$\alpha = \frac{\sum_{\tau=1}^{t} q_{2,\tau}}{\sum_{\tau=1}^{t} (q_{2,\tau} + q_{3,\tau})}. \tag{1.2}$$

Budeme-li tento výpočet opakovat pro dostatečně velký počet časových okamžiků, budou se hodnoty α lišit navzájem jen nepatrně. V daném okamžiku

♣ můžeme vzorec (1.2) považovat za dobrý odhad skutečného poměru vozidel odbočujících na křižovatce doprava.

II. Rekurzivní odhad pomocí statistické definice pravděpodobnosti

Abychom si nemuseli pamatovat celé posloupnosti starých dat, zavedeme statistiky

$$n_i(\tau) = n_i(\tau-1) + q_{i,\tau}, \quad i = 1,2,3; \quad \tau = 1,2,...,t\,, \tag{1.3}$$

s počáteční hodnotou $n_i(0) = 0,\ i = 1,2,3$.

Pro takto zavedené statistiky platí

$$n_i(t) = \sum_{\tau=1}^{t} q_{i,\tau} + n_i(0) \tag{1.4}$$

a pro odhad α dostaneme

$$\alpha = \frac{n_2(t)}{n_2(t) + n_3(t)}. \tag{1.5}$$

♣ Vzorec (1.5) představuje stejný odhad jako (1.2), ale zapsaný pomocí statistik.

Poznámka: Odhady je možno z průběžně napočítávaných statistik počítat také průběžně.

Poznámka: Máme-li určitý (tzv. apriorní) odhad hodnoty parametru α, můžeme jej do statistik zabudovat tak, že nezvolíme jejich počáteční hodnoty nulové, ale tak, aby vyjadřovaly naše představy o tomto parametru:

- *poměr $\frac{n_2(0)}{n_2(0)+n_3(0)}$ vyjadřuje náš počáteční odhad hodnoty parametru α,*
- *velikost $n_2(0)$ a $n_3(0)$ vyjadřuje naši důvěru v tento odhad vzhledem k několika prvním naměřeným datům.*

`Cvičení:` Ověřte tvrzení předchozí poznámky.

Provedli jsme jedno měření a obdrželi data $q_(1,1) = 2,\ q_(2,1) = 8$. Počáteční hodnoty volíme
1. $n_2(0) = 1,\ n_3(0) = 1$;
2. $n_2(0) = 100,\ n_3(0) = 100$.

Ověřte, že velké hodnoty počátečních podmínek mají tendenci se více prosadit proti prvním naměřeným datům.

III. Odhad pomocí deterministického modelu

Podíl α aut z proudu q_1 odbočí doprava a vytvoří proud q_2. Platí tedy:

$$q_{2,\tau} = \alpha q_{1,\tau}, \qquad \tau = 1, 2, ...t\,. \tag{1.6}$$

Je zřejmé, že takový model je idealizace. Ve skutečnosti se bude poměr aut odbočujících doprava v každém okamžiku měřená vždy trochu lišit, a lze hovořit pouze o jeho "ideální" hodnotě. Tento nedostatek deterministického modelu lze řešit zavedením poruchové veličiny, která popisuje okamžité odchylky od průměrných (nebo "ideálních") veličin. Můžeme postupovat dvojím způsobem.

`▷ Deterministický model s poruchou v parametru`

Uvažujeme poruchu, která působí na parametr α a způsobuje, že tento parametr je proměnný, tj. kolísá kolem své "ideální" hodnoty. Pro jednoduchost budeme v tomto případě uvažovat konstantní příjezd vozidel, tedy $q_{1,\tau} = q_1$, pro všechna τ. Model má potom tvar

$$q_{2,\tau} = \alpha_\tau q_1. \tag{1.7}$$

Odhadování parametrů provedeme takto:
Pro jednotlivé časové okamžiky můžeme psát rovnice

$$\begin{aligned} q_{2,1} &= \alpha_1 q_1, \\ q_{2,2} &= \alpha_2 q_1, \\ ... &\quad ..., \\ q_{2,t} &= \alpha_t q_1. \end{aligned} \tag{1.8}$$

Sečteme levé i pravé strany těchto rovnic, a dostaneme

$$\sum_{\tau=1}^{t} q_{2,\tau} = \sum_{\tau=1}^{t} \alpha_\tau q_1. \tag{1.9}$$

Pomocí (1.4) a rovnosti $t\,q_1 = n_1(t) = n_2(t) + n_3(t)$ můžeme rovnost (1.9) psát ve tvaru

$$n_2(t) = \sum_{\tau=1}^{t} \frac{1}{t} \alpha_\tau t\, q_1 = \frac{1}{t} \sum_{\tau=1}^{t} \alpha_\tau (n_2(t) + n_3(t)).$$

Odtud a s použitím (1.5) plyne

$$\alpha = \frac{n_2(t)}{n_2(t) + n_3(t)} = \frac{1}{t} \sum_{\tau=1}^{t} \alpha_\tau = \bar{\alpha}. \tag{1.10}$$

Vidíme, že:

♣ Odhad (1.5) má význam průměrné hodnoty časově proměnného parametru α_τ.

Poznámka: Popsaný postup vede na optimální odhad jen v nejjednodušších případech. Zaručuje totiž jen nulový průměr odchylek v rovnicích, nikoli nejmenší odchylky. Použité kriterium musí zaručit, aby se sčítané chyby nemohly navzájem kompenzovat. Proto se používá součet kvadrátů nebo absolutních hodnot odchylek.

Poznámka: Přístup používající proměnné parametry modelu není obvyklý, protože práce s takovým modelem je značně obtížná (vede na řešení diferenčních rovnic s proměnnými koeficienty).

▷ `Deterministický model s poruchou v datech`

Uvažujeme poruchu v datech. Model (1.6) má potom tvar

$$q_{2,\tau} = \alpha q_{1,\tau} + e_\tau, \tag{1.11}$$

kde

e_τ je porucha způsobující, že některá auta místo odbočení doprava (jak by v "ideálním" případě měla) odbočí doleva,

α je parametr, který je v tomto případě konstantní.

Odhad parametrů může být následující: Pro jednotlivé časové okamžiky dostáváme rovnice

$$\begin{aligned} q_{2,1} &= \alpha q_{1,1} + e_1 \\ q_{2,2} &= \alpha q_{1,2} + e_2 \\ \dots & \quad \dots \\ q_{2,t} &= \alpha q_{1,t} + e_t. \end{aligned} \tag{1.12}$$

Opět sečteme levé a pravé strany rovnic, využijeme (1.4) a dostáváme

$$n_2(t) = \alpha n_1(t) + \sum_{\tau=1}^{t} e_\tau.$$

Vzorec (1.10), který jsme dostali v předchozích případech

$$\alpha = \frac{n_2(t)}{n_2(t) + n_3(t)} \tag{1.13}$$

dostaneme v případě, že platí $\sum_{\tau=1}^{t} e_\tau = 0$. Můžeme tedy říci:

Poznámka: Řekli jsme, že budou-li mít rovnice (1.12) v průměru nulovou chybu, bude vzorec pro odhad koeficientu α dán vztahem (1.13), podobně jako dříve. Tato úvaha je namístě, jestliže předpokládáme, že skutečně nějaká porucha e_τ, $\tau = 1, 2, ..., t$ existuje a vstupuje do zmíněných rovnic.

Běžnější situace však je, když máme pouze naměřená data $q_{2,\tau}$ a $q_{1,\tau}$, $\tau = 1, 2, ..., t$ a porucha e_τ je pouze fiktivní. Potom lze předchozí předpoklad o poruše formulovat jako požadavek: Určete parametr α v rovnicích (1.12) aby pro naměřená data měla porucha v průměru nulovou chybu. *Tento požadavek je určitým kritériem, podle něhož neznámý parametr určujeme. Výsledkem je již odvozený vzorec (**??**).*

♣ Vzorec (1.13) je opět stejný jako v (1.5) a dává odhad, pro nějž mají rovnice (1.12) v průměru nulovou poruchu.

Cvičení: Ověřte, že stejný výsledek dostaneme i při použití daleko kvalitnějšího kriteria, kterým je minimum součtu kvadrátů poruch v jednotlivých rovnicích, tedy

$$\min_{\alpha} \sum_{i=1}^{t} e_i^2.$$

Návod: v rovnicích vyjádřete odchylky a zapište kriterium – součet kvadrátů odchylek. Jeho minimum vypočtěte tak, že derivaci kriteria podle α položíte rovnu nule. Odtud vyjádřete α.

IV. Řešení pomocí pravděpodobnostního modelu

Pravděpodobnost, že náhodně vybrané auto odbočí doprava je α, doleva $1-\alpha$. V časovém okamžiku τ přijede za jednu minutu $q_{1,\tau} = q_{2,\tau} + q_{3,\tau}$ aut. Pravděpodobnost, že jich $q_{2,\tau}$ odbočí doprava a $q_{1,\tau} - q_{2,\tau} = q_{3,\tau}$ doleva, je dána binomickým rozdělením

$$P(q_{2,\tau}, q_{3,\tau}|\alpha) = \binom{q_{2,\tau} + q_{3,\tau}}{q_{2,\tau}} \alpha^{q_{2,\tau}} (1-\alpha)^{q_{3,\tau}}. \tag{1.14}$$

Tento vztah představuje pravděpodobnostní model naší křižovatky v časovém okamžiku τ.

▷ **Věrohodnostní funkce**

$L_t(\alpha)$, která v sobě obsahuje informaci o hledaném parametru získanou z naměřených dat, je dána součinem modelů (1.14) v jednotlivých časových okamžicích, tj.

$$L_t(\alpha) = \prod_{\tau=1}^{t} P(q_{2,\tau}, q_{3,\tau}|\alpha) = k.\alpha^{\sum_{\tau=1}^{t} q_{2,\tau}} (1-\alpha)^{\sum_{\tau=1}^{t} q_{3,\tau}},$$

kde k je součin kombinačních čísel, který nezávisí na α. Použijeme-li označení (1.4), můžeme věrohodnostní funkci (likelihood) zapsat ve tvaru

$$L_t(\alpha) = k \cdot \alpha^{n_2(t)-1} (1-\alpha)^{n_3(t)-1}, \tag{1.15}$$

kde časový vývoj statistik $n_i(t)$, $i = 2, 3$, je dán vztahem (1.3) s počáteční podmínkou $n_i(0) = -1$.

Poznámka: Tato počáteční podmínka je prozatím zavedena formálně z výpočetních důvodů. Možnost zavedení počátečních podmínek (tzv. apriorní informace) je však velikou předností bayesovského pravděpodobnostního přístupu – viz druhou poznámku ke vzorci (1.3).

Pro další výpočty potřebujeme integrál z této funkce přes celý její definiční obor, tj. přes $\alpha \in [0; 1]$. Ten je

$$\int_0^1 L_t(\alpha) \mathrm{d}\alpha = k \int_0^1 \alpha^{n_2(t)-1} (1-\alpha)^{n_3(t)-1} \mathrm{d}\alpha = k\, B[n_2(t), n_3(t)],$$

kde beta funkce $B[.,.]$ je definována takto (viz Dodatek)

$$B[r, s] = \int_0^1 \alpha^{r-1} (1-\alpha)^{s-1} \mathrm{d}\alpha$$

a má pro nás velmi důležitou vlastnost

$$B[r+1,s] = \frac{r}{r+s} B[r,s].$$

▷ `Maximálně věrohodný odhad parametru` α

Za odhad se volí nejpravděpodobnější hodnota parametru α, tj. maximalizuje se hustota pravděpodobnosti α, nebo věrohodnostní funkce. Hledá se tedy α, pro které je splněna rovnost

$$\frac{\mathrm{d}L_t(\alpha)}{\mathrm{d}\alpha} = 0$$

kde L_t je věrohodnostní funkce z (1.15). Tato podmínka je splněna pro

$$\alpha = \frac{n_2(t)}{n_2(t) + n_3(t)}\ , \tag{1.16}$$

♣ což je opět stejný odhad jako v (1.5), ale statistiky $n_i(t)$ mají počáteční hodnotu $n_i(0) = -1$.

`Cvičení:` Ověřte odhad anulováním derivace věrohodnostní funkce a výpočtem α.

▷ `Středně kvadratický odhad parametru` α

Zvolíme-li za kriterium střední hodnotu kvadratického výrazu $(\alpha - \hat{\alpha})^2$, pomocí kterého měříme kvalitu odhadu $\hat{\alpha}$ parametru α, dostaneme optimální odhad (jak později ukážeme) jako střední hodnotu α podmíněnou všemi naměřenými daty až do času t (označíme je $q(t)$). Její hustota pravděpodobnosti $f(\alpha|q(t))$ je dána normalizovanou věrohodnostní funkcí (tj. touto funkcí dělenou jejím integrálem přes celý definiční obor)

$$\hat{\alpha} = E[\alpha|q(t)] = \int_0^1 \alpha f(\alpha|q(t))\mathrm{d}\alpha = \frac{\int_0^1 \alpha.\alpha^{n_2(t)-1}(1-\alpha)^{n_3(t)-1}\mathrm{d}\alpha}{\int_0^1 \alpha^{n_2(t)-1}(1-\alpha)^{n_3(t)-1}\mathrm{d}\alpha} = \tag{1.17}$$

$$= \frac{B[n_2(t)+1, n_3(t)]}{B[n_2(t), n_3(t)]} = \frac{n_2(t)}{n_2(t)+n_3(t)}.$$

♣ Odhad je opět stejný jako v (1.5).

`Cvičení:` Pokuste se odvodit tvrzení, že minimalizací kvadratického kriteria dostaneme optimální odhad parametru α jako střední hodnotu.
Návod: Kriterium umocněte a proveďte středování. Uvědomte si, že náhodná veličina je α, zatímco odhad $\hat{\alpha}$ je konstanta. (Úprava je podobná odvození výpočetního vzorce pro rozptyl.)

1.3 Typy úloh identifikace a řízení

V předchozích odstavcích jsme nastínili problematiku modelování a identifikace dopravního systému. Viděli jsme, že na systému měříme určité veličiny – výstupy systému a v některých případech můžeme jiné veličiny sami zadávat – vstupy systému. Systém je popsán modelem. To je matematická funkce, která umožňuje určit výstup (nebo jeho statistické charakteristiky) pomocí parametrů modelu a starších veličin (dat), změřených na systému.

Úkolem identifikace je určit hodnoty parametrů (nebo jejich charakteristiky) ze změřených dat. Úkolem řízení je, pro určitý počet kroků (horizont řízení), určit vstupní veličiny systému tak, aby výstupní veličina měla požadovanou hodnotu.

Pro lepší orientaci nyní uvedeme základní úlohy, kterými se budeme dále zabývat. Jsou to tyto úlohy:

- `Řízení pro známé parametry modelu systému.`
  V tomto ne příliš realistickém případě uvažujeme model, jehož parametry jsou známé. Naším úkolem je určit hodnoty vstupních veličin tak, aby výstupní veličina byla co nejblíže zadané konstantní hodnotě .

- `Identifikace parametrů systému.`
  Zde se snažíme ze změřených dat určit parametry modelu systému tak, aby tento model co nejlépe popisoval výstup systému.

- `Predikce výstupu pro neznámé parametry modelu.`
  Úkolem je nalézt takový popis systému, který by předpovídal budoucí hodnotu výstupu jen na základě starších dat, nikoliv parametrů. Takový popis lze získat spojením modelu systému s identifikací jeho parametrů.

- `Řízení výstupu pro neznámé parametry modelu.`
  Řídící veličina, podobně jako predikce výstupu, nesmí záviset na neznámých parametrech modelu systému. Proto musíme syntézu řízení rovněž spojit s identifikací. Podle způsobu, jakým to provedeme, rozlišujeme důvěřivou a opatrnou strategii řízení.

  - *Důvěřivá strategie řízení*, kdy za neznámé parametry dosazujeme jejich bodové odhady.
  - *Opatrná strategie řízení*, kdy při syntéze uvažujeme popis parametrů pomocí jejich hustot pravděpodobnosti.

Kapitola 2

Úvod do řízení a identifikace

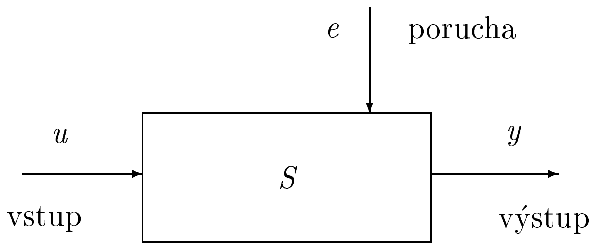
V této kapitole uděláme jen jakýsi náhled na úlohy řízení a identifikace. Pro začátek nebudeme model systému vyjadřovat pomocí hustoty pravděpodobnosti. Budeme uvažovat deterministický popis systému, ve kterém je náhodnost způsobena nějakou poruchou. Algoritmy, které při takto zjednodušeném přístupu odvodíme budou ale dobrým základem algoritmů Bayesovské identifikace a řízení, kterými se budeme zabývat později.

2.1 Model systému

Jak jsme již uvedli,

♣ model systému je představován matematickým vztahem, který určuje budoucí výstup (nebo jeho rozdělení) na základě starších dat změřených na soustavě, parametrů modelu (a rozdělení poruchy).

Veličiny, které v modelu vystupují je možno znázornit pomocí následujícího blokového schématu



Do soustavy S vstupuje signál u – vstup a vystupuje signál y – výstup. Vstup zadáváme my, jako operátoři (často prostřednictvím regulátoru, který jsme určili a který pracuje se staršími vstupy a výstupy – tzv. ve zpětné vazbě). Do soustavy dále vstupuje porucha e. Hodnoty této poruchy nemůžeme měřit, ani je nedokážeme předpovídat. Jen na ni klademe určité předpoklady. Např., že její dlouhodobý průměr je nula, nebo že její hodnoty jsou nezávislé na minulých datech. Vstup i výstup soustavy vzorkujeme, tj. měříme jejich hodnoty v ekvidistantních časových okamžicích. Tyto okamžiky označujeme celými čísly.

Měřená data jsou tedy posloupnostmi hodnot jednotlivých signálů, jejich indexy označují okamžiky měření. V časovém okamžiku τ tedy uvažujeme posloupnosti y_τ, u_τ, e_τ, kde $\tau = 1, 2, ..., t, ..., N$. Indexem t většinou značíme "současný" okamžik a veličiny s tímto indexem současná data. Data pro $\tau = 1, 2, ..., t-1$ nazýváme "stará", nebo již naměřená data. Data s indexem vyšším než t nazýváme "budoucí" data. Celkový počet časových okamžiků, pro které budeme soustavu identifikovat (a řídit), jsme označili N.

Regresní model řádu ν

Tento model je nejčastěji používaným popisem systému. Určuje současný výstup soustavy jako lineární kombinaci současného vstupu a starších vstupů a výstupů. Jako složka náhodnosti se přičítá porucha. Model můžeme zapsat pro jednotlivé členy datových posloupností

$$y_t = b_0 u_t + \sum_{\tau=1}^{\nu} (a_\tau y_{t-\tau} + b_\tau u_{t-\tau}) + c + e_t, \tag{2.1}$$

nebo ve vektorovém tvaru

$$y_t = \varphi_t^T \theta + e_t, \tag{2.2}$$

kde

$t = 1, 2, ..., N$,

ν je hloubka paměti, tzv. *řád modelu*,

$a_1, a_2, ..., a_\nu, b_1, b_2, ..., b_\nu, c$ jsou parametry modelu (konstanty),

y_τ je vzorkovaný výstup (posloupnost čísel),

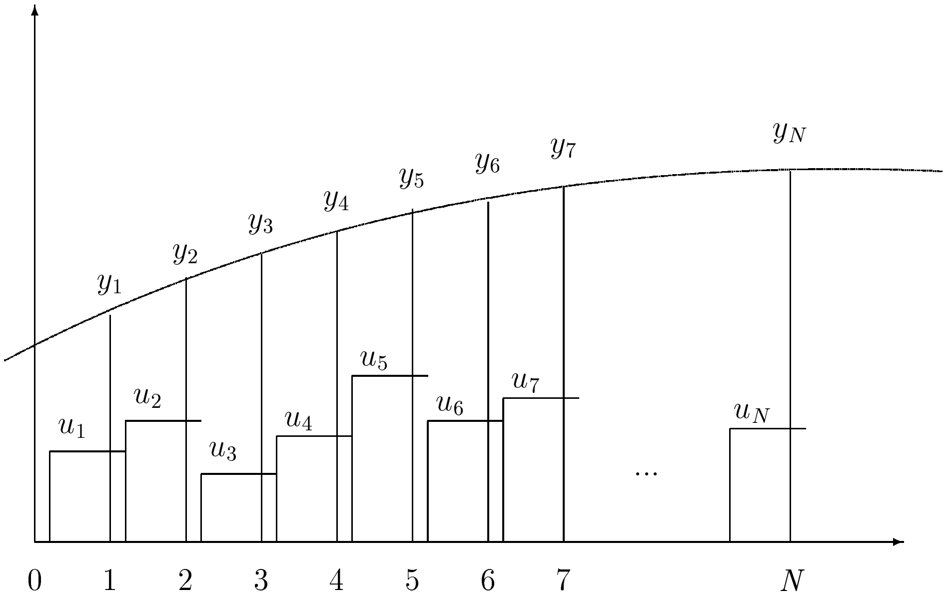
u_τ je vstup (posloupnost čísel, definující po částech konstantní akční veličinu),

e_τ je náhodná porucha (posloupnost čísel – realizací náhodné veličiny s nulovou střední hodnotou a rozptylem σ^2)

$\theta^T = [b_0, a_1, b_1, .., a_n, b_\nu, c]$ je vektor parametrů modelu a

$\varphi_t^T = [u_t, y_{t-1}, u_{t-1}, .., y_{t-\nu}, u_{t-\nu}, 1]$ je regresní vektor.

Posloupnosti vzorkovaného vstupu a výstupu jsou znázorněny na následujícím obrázku.



Příklad [Regresní model 1. řádu bez šumu]
Zvolíme-li $\nu = 1,\ b_0 = 1,\ a_1 = 0.5,\ b_1 = 0$ a poruchu $e_t = 0$, pro všechna t, dostaneme následující tvar regresního modelu

$$y_t = 0.5 y_{t-1} + u_t.$$

Tento model popisuje deterministický, dynamický systém. Pro $u_t = 0$, pro všechna t, se model nazývá autonomní, nebo neřízený, a popisuje doznívání počátečních podmínek y_0. Pro $u_t \neq 0$ model popisuje řízený systém.

Např. pro počáteční podmínku $y_0 = 10$ a řízení: $u_1 = 1,\ u_\tau = 0$ pro $\tau > 1$ je odezva tohoto modelu naznačena v tabulce.

τ	u_τ	výpočet y	y_τ
0	–	–	10
1	1	$0.5{\cdot}10 + 1$	6
2	0	$0.5{\cdot}6 + 0$	3
3	0	$0.5{\cdot}3 + 0$	1.5
4	0	$0.5{\cdot}1.5 + 0$	0.75

Poznámka: Pro deterministický systém je možno při známých vstupech u_t celou odezvu předem spočítat a řídit bez zpětné vazby. Stochastický systém je třeba řídit se zpětnou vazbou.

Příklad [Regresní model 1. řádu se šumem]
Tento model je strukturou shodný s předchozím, ale navíc jsme přidali poruchu, která způsobuje narušení jinak ideálního průběhu výstupu, generovaného deterministickým systémem. Model má tvar

$$y_t = 0.5 y_{t-1} + u_t + e_t,$$

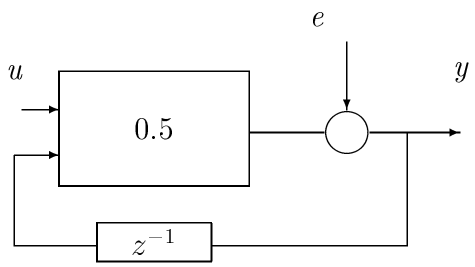
kde

e_t je porucha (šum), o které předpokládáme, že nijak nesouvisí s minulými daty, měřenými na soustavě.

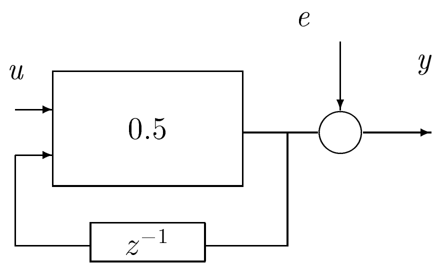
Model popisuje stochastický dynamický systém. Jeho odezva y pro daný vstupní signál u a jednu vybranou realizaci poruchy e je následující

τ	u_τ	e_τ	výpočet y	y_τ
0	–	–	–	10
1	1	0.1	$0.5{\cdot}10 + 1 + 0.1$	6.1
2	0	-0.3	$0.5{\cdot}6.1 + 0 - 0.3$	2.75
3	0	0.2	$0.5{\cdot}2.75 + 0 + 0.2$	1.575
..	..	..	...	..

Poznámka: Do každé další hodnoty výstupu se započítává i šum. Zapojení které uvažujeme, je schématicky znázorněno na levém obrázku.



ANO



NE

2.2 Základní úloha řízení

Cílem je sestavit plán akčních zásahů $u_\tau < \infty$ pro předem zvolený počet kroků řízení N (horizont řízení), který minimalizuje kriterium

$$\mathcal{J} = \sum_{\tau=1}^{N} y_\tau^2$$

je-li soustava popsána deterministickým regresním modelem, tj modelem (2.1) s nulovým šumem e, známe-li potřebné počáteční podmínky $y_0, y_{-1}, .., u_0, u_{-1}, \ldots$

PŘÍKLAD 2.1 Řízení s deterministickým modelem 1. řádu

Úlohu budeme demonstrovat na příkladě, kde horizont $N = 3$, kriterium

$$\mathcal{J} = \sum_{\tau=1}^{3} y_\tau^2,$$

a deterministický regresní model je 1. řádu

$$y_t = a y_{t-1} + b u_t. \tag{2.3}$$

Potřebnou počáteční podmínkou je y_0.

Jednorázová minimalizace kriteria

Protože se jedná o deterministickou soustavu se známými počátečními podmínkami, lze všechny hodnoty výstupu figurující v kriteriu pro známé hodnoty vstupů vyjádřit (predikovat) pomocí počátečních podmínek a řízení

$$\begin{aligned} y_1 &= ay_0 + bu_1, \\ y_2 &= ay_1 + bu_2 = a(ay_0 + bu_1) + bu_2 = a^2 y_0 + abu_1 + bu_2, \\ y_3 &= ay_2 + bu_3 = a(a^2 + abu_1 + bu_2) + bu_3 = a^3 y_0 + a^2 bu_1 + abu_2 + bu_3. \end{aligned}$$

Dosadíme-li tyto predikované hodnoty do kriteria, dostaneme funkci neznámých řízení u_1, u_2, u_3 (počáteční výstup y_0 a parametry a, b předpokládáme známé). Trojice řízení, která kriterium minimalizuje, představuje řízení optimální.

Poznámka: Jak jsme již zmínili, pro deterministický systém je možno spočítat hodnoty optimálního řízení pro všechny časové okamžiky najednou. Není třeba čekat na postupně generované výstupy a zavádět zpětnou vazbu. Minimalizaci ale můžeme provádět také postupně. Vzhledem k návaznosti na dynamické systémy a přirozeně vyžadovanou kauzalitu regulátoru budeme minimalizovat "odzadu" proti směru času.

Protože chceme získat řízení kauzální, je třeba pomocí predikce eliminovat budoucí výstupy, které by způsobovaly nekauzalitu akčních zásahů. Není třeba eliminovat všechny výstupy na horizontu řízení. Vypočtený akční zásah může záviset na minulých, a tedy již známých výstupech. V čase $t = 3$ generujeme vstup u_3, kterým řídíme výstup y_3. Tento výstup tedy pro určení u_3 nesmíme použít. V kriteriu za něj dosadíme z rovnice modelu. Dostaneme

$$\mathcal{J} = y_1^2 \ + \ y_2^2 \ + \ (ay_2 + bu_3)^2.$$

Tím jsme se zbavili y_3, které není pro výpočet u_3 k dispozici a dostali jsme kriterium jako funkci u_3 a starších dat. Jeho minimalizací podle u_3 dostaneme

$$u_3 = -\frac{a}{b} y_2$$

a obdobně v časech $\tau = 2$ a $\tau = 1$ obdržíme u_2 a u_1

$$u_\tau = -\frac{a}{b} y_{\tau-1} \quad \text{pro} \quad \tau = 2, 1.$$

Poznámka: Řízení je kauzální, jak jsme to požadovali a zajistili při jeho odvození. Pro výpočet u_τ se používá starší $y_{\tau-1}$, které je v okamžiku generování akční veličiny již známé.

Poznámka: Postup minimalizace podle klesajícího času zde není nutný. Nezávislost na směru minimalizace je dána tím, že v každém kroku jsme schopni minimalizovat bezezbytku příslušnou část kriteria. Žádný zbytek nepřechází do následujícího kroku sekvenční minimalizace. Např. na u_3 závisí jen poslední člen kriteria. Volbou optimálního řízení $u_3 = -\frac{a}{b} y_2$ tuto část kriteria zcela anulujeme. Ostatní členy kriteria zůstávají nedotčeny. Pouze jejich počet se postupně zmenšuje.

Poznámka: Uvedené kriterium se nazývá "minimum variance" (tj. minimální rozptyl). Výpočet optimálního řízení podle tohoto kriteria je velmi jednoduchý, avšak praktická kvalita řízení může *být pro některé modely špatná. Řídící zákon může generovat nekonečně velké hodnoty řídící veličiny a řízení tak může být prakticky nerealizovatelné.*

Postupná minimalizace kriteria

Abychom předešli nedostatkům, které sebou nese kriterium "minimum variance", budeme uvažovat kriterium, které penalizuje nejenom výstupy ale také vstupy. Velikost penalizace vstupů volí prostřednictvím proměnné $\omega > 0$ uživatel. Pro náš příklad má toto kriterium následující tvar

$$\mathcal{J} = \sum_{\tau=1}^{3} (y_\tau^2 + \omega u_\tau^2).$$

Dosadíme-li z rovnice modelu (2.3) pro čas $t = 3$, lze toto kriterium rozepsat následovně

$$\mathcal{J} = y_1^2 + \omega u_1^2 \ + \ y_2^2 + \omega u_2^2 \ + \ (ay_2 + bu_3)^2 + \omega u_3^2.$$

V tomto případě se nám nepodaří volbou řízení jednotlivé členy kriteria postupně zcela anulovat. Minimalizaci začneme podle u_3 a od posledního členu, který jediný na u_3 závisí. Zbytek po jeho minimalizaci bude funkcí předchozích vstupů a výstupů a bude třeba jej sloučit s ostatními členy kriteria a zahrnout jej do další minimalizace. Protože v tomto případě se nám volbou optimálního řízení nepodaří jednotlivé členy kriteria zcela anulovat, a protože tyto členy jsou po minimalizaci vždy funkcemi předchozích dat, je zde postup minimalizace "odzadu" jediný možný.

`▷ Minimalizace 3. členu`

Minimalizaci provedeme tak, že 3. člen kriteria $\mathcal{J}_3 = (ay_2 + bu_3)^2 + \omega u_3^2$ doplníme na úplný čtverec v proměnné u_3. Kvadrát můžeme anulovat volbou u_3, zbytek po doplnění se připojí k ostatním členům.

$$\mathcal{J}_3 = (ay_2 + bu_3)^2 + \omega u_3^2 = u_3^2(b^2 + \omega) + 2aby_2u_3 + a^2y_2^2 =$$

$$= (b^2 + \omega)\left[u_3^2 + 2u_3\frac{ab}{b^2+\omega}y_2 + \left(\frac{ab}{b^2+\omega}\right)^2 y_2^2 - \left(\frac{ab}{b^2+\omega}\right)^2 y_2^2\right] + a^2y_2^2 =$$

$$= (b^2 + \omega)\left[u_3 + \frac{ab}{b^2-\omega}y_2\right]^2 + a^2y_2^2 - \frac{a^2b^2}{b^2+\omega}y_2^2.$$

Označíme-li

$$r_3 = \frac{ab}{b^2+\omega} \quad \text{a} \quad s_3 = \frac{a^2\omega}{b^2+\omega} = \frac{a\omega}{b}r_3,$$

můžeme optimální akční zásah v čase 3 a zbytek po minimalizaci třetího členu psát takto

$$u_3 = -r_3y_2 \quad \text{a} \quad \min_{u_3} \mathcal{J}_3 = s_3y_2^2.$$

Kriterium po minimalizaci podle u_3 je

$$\mathcal{J} = [y_1^2 + \omega u_1^2] + [y_2^2 + \omega u_2^2] + s_3y_2^2 = [y_1^2 + \omega u_1^2] + (1 + s_3)y_2^2 + \omega u_2^2.$$

Akční zásah u_2 neovlivní první člen.

`▷ Minimalizace 2. členu a zbytku`

$$\mathcal{J}_2 = (1+s_3)y_2^2 + \omega u_2^2 = (1+s_3)(a^2y_1^2 + 2aby_1u_2 + b^2u_2^2) + \omega u_2^2 =$$

$$= [(1+s_3)b^2+\omega]\left[u_2^2 + 2u_2\frac{(1+s_3ab)}{(1+s_3)b^2+\omega}y_1 + \left(\frac{(1+s_3)ab}{(1+s_3b^2)+\omega}\right)^2 y_1^2 - \right.$$

$$\left. - \left(\frac{(1+s_3)ab}{(1+s_3b^2)+\omega}\right)^2 y_1^2\right] + (1+s_3)a^2y_1^2 =$$

$$= [(1+s_3)b^2+\omega]\left[u_2 + \frac{(1+s_3)ab}{(1+s_3)b^2+\omega}y_1\right]^2 + \frac{(1+s_3a^2\omega)}{(1+s_3)b^2+\omega}y_1^2.$$

Označíme-li

$$r_2 = \frac{(1+s_3)ab}{(1+s_3)b^2+\omega} \quad \text{a} \quad s_2 = \frac{(1+s_3)a^2\omega}{(1+s_3)b^2+\omega} = \frac{a\omega}{b}r_2,$$

můžeme optimální akční zásah v čase $t = 2$ a zbytek po minimalizaci druhého členu psát ve tvaru

$$u_2 = -r_2y_1 \quad \text{a} \quad \min_{u_2} \mathcal{J}_2 = s_2y_1^2.$$

Kriterium po minimalizaci podle u_2 je

$$\mathcal{J} = [y_1^2 + \omega u_1^2] + s_2y_1^2 = (1+s_2)y_1^2 + \omega u_1^2.$$

▷ `Minimalizace 1. členu a zbytku`

$$\mathcal{J}_1 = (1+s_2)y_1^2 + \omega u_1^2.$$

Obdobnými úpravami jako v předchozích krocích s označením

$$r_1 = \frac{(1+s_2)ab}{(1+s_2)b^2+\omega} \quad \text{a} \quad s_1 = \frac{a\omega}{b}r_1$$

dojdeme opět k výsledku

$$u_1 = -r_1y_0 \quad \text{a} \quad \min_{u_1} \mathcal{J}_1 = s_1y_0^2.$$

Výsledky mají formálně stejný tvar, pouze časový index klesá. Celou postupnou minimalizaci lze zapsat v následujícím algoritmu.

Algoritmus

Řízení

Generování řídícího zákona

$s_{N+1} = 0$

for $t = N$ **downto** 1 **do**

$r_t = \frac{(1+s_{t+1})ab}{(1+s_{t+1})b^2+\omega}$

$s_t = \frac{a\omega}{b} r_t$

end for t

$J_0 = s_1$

Generování řídící veličiny

y_0 - apriorní data

for $t = 1$ **to** N **do**

$u_t = -r_t y_{t-1}$

end for t

Poznámka: Kriterium "minimum variance" dostaneme, položíme-li $\omega = 0$. Potom bude

$$r_i = \frac{a}{b}, \quad u_i = -\frac{a}{b} y_{i-1} \ \text{ a } \ s_i = 0, \quad i = 3, 2, 1$$

což odpovídá výsledkům, které jsme již pro tento případ odvodili.

`Cvičení:` Uvedený algoritmus naprogramujte v jazyce MATLAB a testujte jeho funkci pro různé penalizace vstupu ω při různých počátečních podmínkách.
Návod: program (včetně žádané hodnoty výstupu, tj. průběhu, který má výstup sledovat) je uveden v Dodatku (4.4).

2.3 Základní úloha identifikace

Pro návrh řízení potřebujeme znát koeficienty modelu řízené soustavy. Ty ale ve většině praktických případů známy nejsou. Proto je musíme odhadovat z dat měřených na soustavě. Tento proces nazýváme identifikací. Základem identifikace je minimalizace kriteria

$$\mathcal{J} = \sum_{\tau=1}^{N} e_\tau^2 \to \min, \tag{2.4}$$

kde e_τ, $\tau = 1, 2, ..N$, interpretujeme jako chyby se kterými jsou splněny rovnice definované modelem (2.1) po dosazení naměřených dat a odhadnutých koeficientů. Chyby, pro které je dosaženo minima kriteria, se nazývají *rezidua* a odpovídající koeficienty modelu jsou *bodové odhady parametrů soustavy.*

Základní úlohu budeme demonstrovat na jednoduchém příkladě

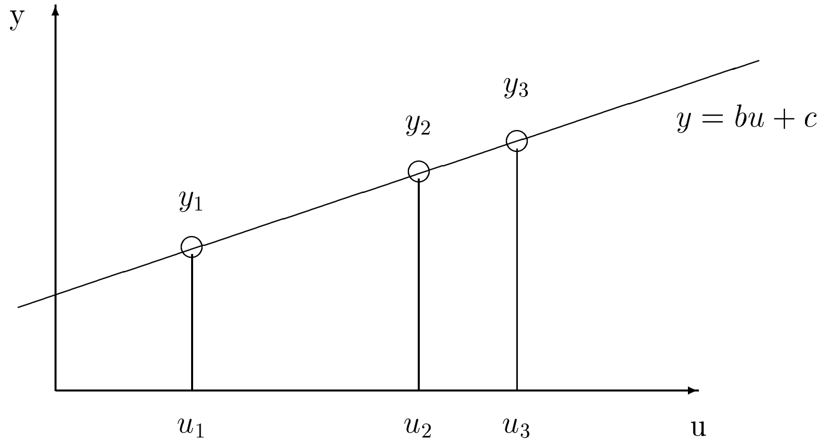
PŘÍKLAD 2.2 Identifikace statického regresního modelu (skalární vyjádření)

▷ `Formulace úlohy pro statický model`

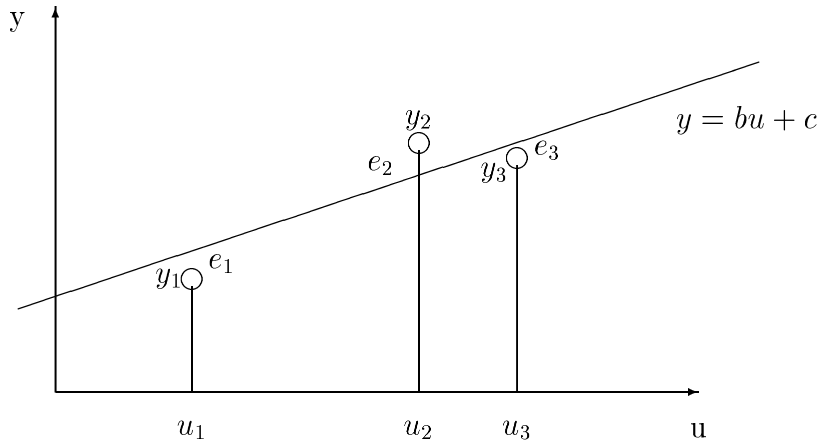
Uvažujme regresní model, kde výstup y_t v daném okamžiku t závisí pouze na veličinách u_t a e_t ve stejném čase časovém okamžiku. Takový model se nazývá statický a je zadán rovnicí

$$y_t = bu_t + c + e_t. \tag{2.5}$$

Tento model pro $e_t = 0$, pro všechna t, popisuje lineární závislost mezi vstupem u a výstupem y. Dvojice $[u_t, y_t]$ leží na přímce se směrnicí b.



Bude-li $e_t \neq 0$ bude tato lineární závislost "narušena šumem", jak ukazuje následující obrázek.



Skutečné hodnoty výstupu y_t nebudou ležet na přímce, ale budou odchýleny vertikálně o hodnoty šumu e_t. Naším úkolem je určit odhady parametrů přímky tak, aby součet kvadrátů těchto odchylek byl minimální. Takové odhady parametrů budeme nazývat *odhady regresních koeficientů* a minimální odchylky nazveme *rezidua*.

K odvození odhadů pro takto jednoduchou úlohu je výhodné použít operátor průměrování, který je pro posloupnost $x = (x_1, x_2, ..., x_N)$ definován takto (viz Dodatek (4.2))

$$\overline{x} = \frac{1}{N} \sum_{i=1}^{N} x_i.$$

▷ `Výpočet regresních koeficientů`

Pro model (2.5) a kriterium (2.4) při použití operátoru průměrování (4.2) máme určit regresní koeficienty modelu a, b tak, abychom minimalizovali zvolené kriterium.

Abychom pro výpočet regresních koeficientů mohli využít operátor průměrování, budeme formálně minimalizovat kriterium

$$\frac{1}{N}\mathcal{J} = \frac{1}{N}\sum_{\tau=1}^{N} e_\tau^2 = \overline{e^2}.$$

V příkladě budeme pro přehlednost uvažovat horizont $N = 3$ (tj. odhady parametrů počítáme ze tří naměřených datových dvojic $[y_t, u_t]$, $t = 1, 2, 3$).

Z rovnice modelu vyjádříme $e = y - bu - c$ a dosadíme do kriteria

$$\frac{1}{N}\mathcal{J} = \overline{e^2} = \overline{(y - bu - c)^2} = \overline{y^2 + b^2u^2 + c^2 - 2byu - 2cy + 2bcu} =$$

$$= \overline{y^2} + b^2\overline{u^2} + c^2 + 2b\overline{yu} - 2c\overline{y} + 2bc\overline{u}.$$

Tento výraz budeme doplňovat na čtverec v proměnných b a c. Začneme proměnnou c

$$\frac{1}{N}\mathcal{J} = c^2 - 2c(\overline{y} - b\overline{u}) + (\overline{y} - b\overline{u})^2 - (\overline{y} - b\overline{u})^2 + \overline{y^2} + b^2\overline{u^2} - 2b\overline{yu} =$$

$$= [c - (\overline{y} - b\overline{u})]^2 + \overline{y^2} - \overline{y}^2 - 2b(\overline{yu} - \overline{y}.\overline{u}) + b^2(\overline{u^2} - \overline{u}^2).$$

Protože na c závisí jenom člen v hranatých závorkách, jehož druhá mocnina je nezáporná, dostaneme minimum kriteria jeho anulováním, tedy volbou $c = \hat{c} = \overline{y} - b\overline{u}$, pro libovolné b. Zbytek kriteria po této minimalizaci bude

$$\min_c \frac{1}{N}\mathcal{J} = S_y^2 - 2bS_{yu} + b^2S_u^2$$

kde S_y^2, S_{yu} a S_u^2 jsou druhé výběrové momenty proměnných y a u definované vztahy

$$\begin{aligned} S_y^2 &= \overline{(y-\overline{y})^2} &&= \overline{y^2} - \overline{y}^2, \\ S_{yu} &= \overline{(y-\overline{y})(u-\overline{u})} &&= \overline{yu} - \overline{y}.\overline{u}, \\ S_u^2 &= \overline{(u-\overline{u})^2} &&= \overline{u^2} - \overline{u}^2. \end{aligned} \tag{2.6}$$

Zbytek kriteria dále doplníme na čtverec v proměnné b a budeme minimalizovat

$$\min_c \frac{1}{N}\mathcal{J} = S_u^2\left[b^2 - 2b\frac{S_{yu}}{S_u^2} + \left(\frac{S_{yu}}{S_u^2}\right)^2 - \left(\frac{S_{yu}}{S_u^2}\right)^2\right] + S_y^2$$

$$= S_u^2\left[b - \frac{S_{yu}}{S_u^2}\right]^2 + S_y^2 - \frac{(S_{yu})^2}{S_u^2}.$$

Minimum kriteria je dosaženo pro $b = \hat{b} = \frac{S_{yu}}{S_u^2}$. Minimální hodnota kriteria je

$$\min_{b,c} \mathcal{J} = S_y^2 - \frac{(S_{yu})^2}{S_u^2} \tag{2.7}$$

s odhady regresních koeficientů

$$\hat{b} = \frac{S_{yu}}{S_u^2}, \qquad \hat{c} = \overline{y} - \hat{b}\overline{u}, \tag{2.8}$$

kde výběrové druhé momenty jsou definovány v (2.6.)

`Cvičení:` Uvedený algoritmus naprogramujte v jazyce MATLAB a vyzkoušejte jej pro různé disperze šumů a různé posloupnosti vstupů.
Návod: Nejdříve vypočtěte druhé výběrové momenty, potom určete odhady regresních koeficientů. Odhadnutou přímku vykreslete a porovnejte se skutečnou. (Pro kontrolu je program uveden v Dodatku 4.5).

Uvedené odvození je názorné, ale nehodí se pro modely vyšších řádů. Nyní uvedeme pro stejný model vhodnější způsob odvození pomocí maticového počtu.

PŘÍKLAD 2.3 Identifikace statického regresního modelu (maticové vyjádření)

Statický model z předchozího příkladu vyjádříme pomocí vektorů

$$y_\tau = \begin{bmatrix} u_\tau & 1 \end{bmatrix} \begin{bmatrix} b \\ c \end{bmatrix} + e_\tau = \varphi_\tau^T \theta + e_\tau.$$

Pro $\tau = 1, 2, 3$ můžeme sestavit maticový model

$$Y = \Phi\theta + E, \tag{2.9}$$

kde

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}, \quad \Phi = \begin{bmatrix} u_1 & 1 \\ u_2 & 1 \\ u_3 & 1 \end{bmatrix} = \begin{bmatrix} \varphi_1^T \\ \varphi_2^T \\ \varphi_3^T \end{bmatrix}, \quad E = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}, \quad \theta = \begin{bmatrix} b \\ c \end{bmatrix}$$

Kriterium je

$$\frac{1}{3}\mathcal{J} = \frac{1}{3}\sum_{i=1}^{3} e_i^2 = \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \frac{1}{3}E^T E = \frac{1}{3}(Y - \Phi\theta)^T (Y - \Phi\theta) =$$

$$= \frac{1}{3}(Y^T Y - Y^T \Phi\theta - \theta^T \Phi^T Y + \theta^T \Phi^T \Phi\theta).$$

Opět doplníme na čtverec a dostaneme

$$\frac{1}{3}\mathcal{J} = \frac{1}{3}\{\theta^T(\Phi^T\Phi)\theta - \theta^T(\Phi^T\Phi)(\Phi^T\Phi)^{-1}\Phi^T Y - Y^T\Phi(\Phi^T\Phi)^{-1}(\Phi^T\Phi)\theta +$$

$$+[(\Phi^T\Phi)^{-1}\Phi^T Y]^T(\Phi^T\Phi)[(\Phi^T\Phi)^{-1}\Phi^T Y] - Y^T\Phi(\Phi^T\Phi)^{-1}\Phi^T Y + Y^T Y\} =$$

$$= \frac{1}{3}\{[\theta - (\Phi^T\Phi)^{-1}\Phi^T Y]^T \ (\Phi^T\Phi) \ [\theta - (\Phi^T\Phi)^{-1}\Phi^T Y] +$$

$$+ Y^T Y - Y^T\Phi(\Phi^T\Phi)^{-1}\Phi^T Y\}.$$

Minima kriteria dosáhneme anulováním hranaté závorky, tj. volbou

$$\theta = \hat{\theta} = (\Phi^T\Phi)^{-1}\Phi^T Y. \tag{2.10}$$

Minimální hodnota kriteria je

$$\min_{\theta} \mathcal{J} = Y^T Y - Y^T \Phi (\Phi^T \Phi)^{-1} \Phi^T Y. \tag{2.11}$$

▷ `Struktura matic`

Podívejme se nyní blíže na strukturu jednotlivých matic obsažených v předchozích výsledcích a porovnejme je s výsledky předchozího příkladu pro skalární odvození.

Základní prvky

$$\frac{1}{3} Y^T Y = \overline{y^2}, \qquad \frac{1}{3} \Phi^T Y = \overline{\varphi y} = \begin{bmatrix} \overline{yu} \\ \overline{y} \end{bmatrix},$$

$$\frac{1}{3} \Phi^T \Phi = \overline{\varphi \varphi^T} = \begin{bmatrix} \overline{u^2} & \overline{u} \\ \overline{u} & 1 \end{bmatrix}, \qquad (\frac{1}{3} \Phi^T \Phi)^{-1} = \frac{1}{S_u^2} \begin{bmatrix} 1 & -\overline{u} \\ -\overline{u} & \overline{u^2} \end{bmatrix}.$$

Odhady parametrů

$$\hat{\theta} = \frac{1}{S_u^2} \begin{bmatrix} 1 & -\overline{u} \\ -\overline{u} & \overline{u^2} \end{bmatrix} \begin{bmatrix} \overline{yu} \\ \overline{y} \end{bmatrix} = \frac{1}{S_u^2} \begin{bmatrix} \overline{yu} - \overline{y}\,\overline{u} \\ -\overline{u}\,\overline{yu} + \overline{y}\overline{u^2} \quad (+\overline{y}\overline{u}^2 - \overline{y}\overline{u}^2) = \end{bmatrix}$$

$$= \frac{1}{S_u^2} \begin{bmatrix} S_{yu} \\ \overline{y} S_u^2 - \overline{u} S_{yu} \end{bmatrix} = \begin{bmatrix} \frac{S_{yu}}{S_u^2} \\ \overline{y} - \left(\frac{S_{yu}}{S_u^2} \right) \overline{u} \end{bmatrix} = \begin{bmatrix} \hat{b} \\ \hat{c} \end{bmatrix}.$$

Zbytek kriteria

$$\min_{\theta} \mathcal{J} = \overline{y^2} - [\overline{yu}, \quad \overline{y} \begin{bmatrix} \hat{b} \\ \overline{y} - \hat{b}\overline{u} \end{bmatrix} = \overline{y^2} - \hat{b}\overline{yu} + \overline{y}^2 - \hat{b}\overline{y}\,\overline{u} = S_u^2 - \hat{b} S_{yu}.$$

Výsledky samozřejmě souhlasí s těmi, které jsme odvodili při použití skalárního modelu.

`Cvičení:` Algoritmus identifikace naprogramujte v jazyce MATLAB a porovnejte jeho výsledky s výsledky předchozího algoritmu, odvozeného ve skalární symbolice.
Návod: Ze simulovaných dat sestavte podle (2.9) vektor Y a datovou matici Φ. Z nich podle (2.10) zkonstruujte odhad parametrů $\hat{\theta}$.

V předchozím příkladě jsme použili maticovou symboliku odvození identifikace pro statický regresní model. Nyní použijeme tuto metodu k odvození identifikace pro dynamický model s jedním vstupem a jedním výstupem a obecným řádem ν.

PŘÍKLAD 2.4 Identifikace regresního modelu obecného řádu ν

▷ `Model`

Uvažujeme regresní model, popsaný v okamžiku t vztahem

$$y_t = \varphi_t^T \theta + e_t, \quad t = 1, 2, ...N, \tag{2.12}$$

kde

$$\varphi_t^T = [u_t, y_{t-1}, u_{t-1}, ..y_{t-\nu}, u_{t-\nu}, 1],$$

$$\theta^T = [b_0, a_1, b_1, ..a_\nu, b_\nu, c].$$

▷ `Maticový model`

Pro data naměřená do daného (současného) okamžiku t, tj. v okamžicích $\tau = 1, 2, ..t$, lze předchozí model zapsat maticově

$$Y = \Phi\theta + E, \tag{2.13}$$

kde

$$\Phi = \begin{bmatrix} u_1 & y_0 & u_0 & .. & y_{-\nu+1} & u_{-\nu+1} & 1 \\ u_2 & y_1 & u_1 & .. & y_{-\nu+2} & u_{-\nu+2} & 1 \\ .. & .. & .. & .. & .. & .. & .. \\ u_N & y_{N-1} & u_{N-1} & .. & y_{N-\nu} & u_{N-\nu} & 1 \end{bmatrix} = \begin{bmatrix} \varphi_1^T \\ \varphi_2^T \\ .. \\ \varphi_N^T \end{bmatrix}.$$

Matice Φ má N řádků (počet naměřených dat) a $2\nu + 1$ sloupců (délka regresoru).

▷ `Kriterium`

$$\mathcal{J} = \frac{1}{N}(Y - \Phi\theta)^T(Y - \Phi\theta) = \frac{1}{N}\sum_{i=1}^{N}(y_i - \varphi_i^T\theta)^T(y_i - \varphi_i^T\theta) = \tag{2.14}$$

$$= \frac{1}{N}\sum_{i=1}^{N}\left([y_i, \quad \varphi_i^T] \begin{bmatrix} -1 \\ \theta \end{bmatrix}\right)^T \left([y_i \quad \varphi_i^T] \begin{bmatrix} -1 \\ \theta \end{bmatrix}\right) =$$

$$= [-1 \quad \theta] \left(\frac{1}{N}\sum_{i=1}^{N}\begin{bmatrix} y_i \\ \varphi_i \end{bmatrix} [y_i, \quad \varphi_i^T]\right)\begin{bmatrix} -1 \\ \theta \end{bmatrix}$$

Součet v kulaté závorce posledního výrazu, který označíme $V(t)$, se nazývá *rozšířená informační matice*

$$V(N) = \frac{1}{N}\sum_{i=1}^{N}\begin{bmatrix} y_i \\ \varphi_i \end{bmatrix} [y_i, \quad \varphi_i^T] = \begin{bmatrix} \overline{y^2} & \overline{y\varphi^T} \\ \overline{y\varphi} & \overline{\varphi\varphi^T} \end{bmatrix}. \tag{2.15}$$

Tato matice obsahuje veškerou informaci o identifikovaných parametrech, kterou jsme získali z naměřených dat. Její hodnotu je možno vypočítat buď jednorázově, jak to ukazuje vztah (2.15), nebo rekurentně pro $t = 1, 2, .., N$

$$V(t) = \frac{1}{t}\left\{(t-1)\cdot V(t-1) + \begin{bmatrix} y_t \\ \varphi_t \end{bmatrix} [y_t, \quad \varphi_t^T]\right\} \tag{2.16}$$

se startem $V(0) = 0$, nebo s apriorní informací vyjádřenou počáteční hodnotou $V(0) = V_0$, kde matice V_0 je symetrická a pozitivně definitní.

Rozšířená informační matice má blokovou strukturu a pro její jednotlivé bloky zavedeme značení

$$V(t) = \begin{bmatrix} \overline{y^2} & \overline{y\varphi^T} \\ \overline{y\varphi} & \overline{\varphi\varphi^T} \end{bmatrix} = \begin{bmatrix} V_y(t) & V_{y\varphi}^T(t) \\ V_{y\varphi}(t) & V_\varphi(t) \end{bmatrix}.$$

Obdobným způsobem jako v předchozím případě, tj. doplněním na čtverec, lze odvodit následující algoritmus identifikace.

Algoritmus

Identifikace

Konstrukce informační matice

`Máme:` $V(t-1)$ – stará informační matice.

`Změříme:` y_t, u_t – nová data.

`Sestavíme:` $\varphi_t^T = [u_t, y_{t-1}, \varphi_{t-1}(1:(2n-1)), 1]$

– nový regresor (stará data posuneme o 2 kroky doprava, poslední tři vypustíme, na první dvě pozice umístíme u_t a y_{t-1} a na poslední dáme jedničku).

`Přepočteme:` $V(t) = V(t-1) + \begin{bmatrix} y_t \\ \varphi_t \end{bmatrix} [y_t, \ \ \varphi_t^T]$

– nová informační matice.

Výpočet odhadů koeficientů

`Rozdělíme:`

$$V(t) = \begin{bmatrix} V_y(t) & V_{y\varphi}^T(t) \\ V_{y\varphi}(t) & V_\varphi(t) \end{bmatrix}. \tag{2.17}$$

`Odhad parametrů:`

$$\hat{\theta}_t = V_\varphi^{-1}(t) V_{y\varphi}(t). \tag{2.18}$$

`Minimální hodnota kriteria:`

$$\Lambda_t = V_y(t) - V_{y\varphi}^T(t) V_\varphi^{-1}(t) V_{y\varphi}(t) = V_y(t) - \hat{\theta}^T V_{y\varphi}(t). \tag{2.19}$$

`Cvičení:` Algoritmus identifikace naprogramujte v jazyce MATLAB a ověřte jeho funkci pro data, simulovaná různými modely (tj. různého řádu, s různě velikým šumem a různým vstupním signálem). Porovnejte simulované a identifikované parametry pro případ shody struktury simulovaného a identifikovaného modelu. Rozmyslete si, co se při identifikaci děje při neshodě struktury (např. neshodě řádů) těchto modelů.
Návod: Program pro případné experimenty je uveden v Dodatku 4.7.

▷ `Odvození řešení`
Z (2.14) plyne (pro přehlednost vynecháme časový index)

$$\mathcal{J} = [-1, \ \ \theta] \begin{bmatrix} V_y & V_{y\varphi}^T \\ V_{y\varphi} & V_\varphi \end{bmatrix} \begin{bmatrix} -1 \\ \theta \end{bmatrix} = V_y - 2V_{y\varphi}^T \theta + \theta^T V_\varphi \theta$$

Tento výraz doplníme na čtverec v proměnné θ (viz Dodatek 4.3) a dostaneme

$$\mathcal{J} = \theta^T V_\varphi \theta - 2\theta^T V_\varphi V_\varphi^{-1} V_{y\varphi} + (V_\varphi^{-1} V_{y\varphi})^T V_\varphi (V_\varphi^{-1} V_{y\varphi}) + V_y - V_{y\varphi}^T V_\varphi^{-1} V_{y\varphi} =$$
$$= (\theta - V_\varphi^{-1} V_{y\varphi})^T V_\varphi (\theta - V_\varphi^{-1} V_{y\varphi}) + V_y - V_{y\varphi}^T V_\varphi^{-1} V_{y\varphi},$$

kde $V_\varphi^T = V_\varphi$, neboť V_φ je symetrická matice.

Anulováním první části tohoto výrazu dostáváme odhady parametrů, zbytek je minimální hodnota kriteria.

Poznámka: Numericky lepšího řešení je možno dosáhnout pomocí metod faktorizace. Informační matici faktorizujeme pomocí elementárních rotací na součin horní a dolní trojúhelníkové matice

$$V(t) = G^T(t)G(t),$$

kde matici G rozdělíme ve shodě s dělením matice V

$$G = \begin{bmatrix} G_y & 0 \\ G_{y\varphi} & G_\varphi \end{bmatrix}$$

tak, že G_y je číslo, $G_{y\varphi}$ je zbytek prvního sloupce a G_φ je dolní trojúhelníková matice. Kriterium pak můžeme psát takto

$$\mathcal{J} = [-1, \quad \theta] \begin{bmatrix} G_y & G_{y\varphi}^T \\ 0 & G_\varphi^T \end{bmatrix} \begin{bmatrix} G_y & 0 \\ G_{y\varphi} & G_\varphi \end{bmatrix} \begin{bmatrix} -1 \\ \theta \end{bmatrix} =$$

$$= [-G_y, \quad -G_{y\varphi}^T + \theta^T G_\varphi^T] \begin{bmatrix} -G_y \\ -G_{y\varphi} + \theta G_\varphi \end{bmatrix} = G_y^2 + \| G_\varphi \theta - G_{y\varphi} \|^2,$$

kde $\| \, . \, \|$ značí kvadratickou normu použitou pro výpočet kriteria. Odhad regresních koeficientů je dán anulováním druhé části výrazu

$$\hat{\theta}_t = G_{\varphi,t}^{-1} \, G_{y\varphi,t}.$$

Výhodou tohoto vyjádření je, že se invertuje trojúhelníková matice, což je podstatně méně pracné než pro matici plnou. V praxi se tento postup užívá nikoliv pro informační matici, ale pro její inverzi, což je kovarianční matice parametrů. Toto odvození je ale podstatně složitější.

Kapitola 3

Pravděpodobnosti

V této kapitole zopakujeme ze základů pravděpodobnosti to, co budeme dále potřebovat pro vytváření pravděpodobnostních modelů a pro práci s nimi. Nebudeme zde systematicky procházet všechny definice pravděpodobnosti, v případě potřeby je nutné znalosti doplnit z některé základní učebnice.

Dále se budeme setkávat se dvěma typy náhodných veličin. Jednak to jsou signály v konkrétním časovém okamžiku (např. y_t, resp. u_t jsou vstup, resp. výstup v čase t). Tyto náhodné veličiny jsou bližší klasickému pojetí pravděpodobnosti. Jejich naměřené hodnoty tvoří datový výběr. Náhodnost v hodnotách měřených dat je možno možno interpretovat jako vliv určité poruchy, např. jako chyby měření.

Další typ náhodné veličiny budou představovat neznámé parametry. Tady se již jedná o čistě Bayesovskou interpretaci náhodnosti, která je způsobena neznalostí. Dobře víme, že parametry jsou konstanty. Jejich správná hodnota je nám však neznámá, a my je popíšeme, podle Bayesova přístupu, tak, že specifikujeme všechny jejich možné hodnoty a jejich pravděpodobnostní rozdělení. Tím dostávají rovněž charakter náhodné veličiny.

V dalším výkladu budeme mít na mysli především signály jako reprezentanty náhodných veličin. Ty jsou totiž pro osvětlení základních pojmů přirozenější.

3.1 Základní pojmy

Náhodný pokus spočívá ve změření určitého signálu v daném časovém okamžiku. Např. v čase t změříme výstup a obdržíme hodnotu y_t. Tato naměřená hodnota je výsledek pokusu – elementární náhodný jev.

Náhodná veličina (nv) X je proměnná, jejíž hodnoty jsou přiřazeny výsledkům náhodného pokusu. Protože výsledky našeho pokusu jsou číselné povahy, můžeme hodnoty náhodné veličiny s těmito výsledky přímo ztotožnit. Hodnoty náhodné veličiny jsou tedy přímo naměřené hodnoty signálů.

♣ Zavedení náhodné veličiny spočívá v "očíslování" výsledků náhodného pokusu.

Poznámka: Náhodná veličina je určena náhodným pokusem. Přiřazuje výsledkům náhodného pokusu čísla. Zavádí se proto, aby i v případě, že výsledky nejsou numerické povahy, bylo s *nimi možno počítat - např. určit průměrnou hodnotu výsledků. Jsou-li výsledky pokusu čísla, můžeme je s hodnotami náhodné veličiny ztotožnit.*

Rozdělení náhodné veličiny Náhodná veličina svým oborem hodnot vymezuje množinu potenciálně možných výsledků. Sama neříká nic o tom, jak jsou jednotlivé výsledky pravděpodobné – tj. nepopisuje náhodný pokus, pouze vymezuje "ring, ve kterém se bude hrát". Plným popisem náhodného pokusu je definice náhodné veličiny a jejího rozdělení. Toto rozdělení je dáno určením pravděpodobností všech jednotlivých hodnot diskrétní náhodné veličiny, nebo určením pravděpodobností všech možných intervalů hodnot spojité náhodné veličiny. V našem případě budeme rozdělení diskrétní náhodné veličiny zadávat pomocí pravděpodobnostní funkce, spojité náhodné veličiny pomocí hustoty pravděpodobnosti. Pro oba typy rozdělení budeme používat název hustota pravděpodobnosti a zkracovat jej hp. Definici hp uvedeme dále.

♣ Náhodná veličina a její rozdělení jsou úplným popisem náhodného pokusu.

Poznámka: Podle toho, zda je množina hodnot náhodné veličiny spočetná (konečná) nebo nespočetná, dělíme náhodnou veličinu na diskrétní *nebo* spojitou. *U* diskrétní nv *má smysl mluvit o pravděpodobnostech jednotlivých hodnot, pro* spojitou nv *jsou pravděpodobnosti jednotlivých hodnot vždy nulové, a mluvíme proto o pravděpodobnostech intervalů.*

Poznámka: Budeme-li dále hovořit o "pravděpodobnosti hodnoty x náhodné veličiny X", máme na mysli "pravděpodobnost, s níž náhodná veličina X nabývá hodnotu x". Samotná hodnota náhodné veličiny je reálné číslo, které samozřejmě žádnou pravděpodobnost nemá.

Příklady

Náhodné pokusy:

hod mincí – výsledek: rub, líc;
odbočení auta v křižovatce – výsledek: doprava, rovně, doleva;
doba čekání na dopravní prostředek – výsledek: nezáporná reálná čísla.

Přiřazení hodnot náhodné veličiny výsledkům pokusu:

rub $\to$ *0, líc* $\to$ *1*;
doprava $\to$ *1, rovně* $\to$ *0, doleva* $\to$ *-1*;
doba čekání $\to$ *nezáporné reálné číslo.*

Distribuční funkce $F_X(x)$ náhodné veličiny X je reálná funkce definovaná na celé reálné ose společně pro diskrétní i spojitou nv vztahem

$$F_X(x) = P(X \leq x)$$

a udává pravděpodobnost, že nv X nabude hodnot menších nebo rovných číslu x, tedy pravděpodobnost všech intervalů typu $(-\infty; x)$, kde x je libovolná hodnota náhodné veličiny X.

Hustota pravděpodobnosti (**pravděpodobnostní funkce**) nv X je funkce $f(x)$, pro kterou platí

$$F_X(x) = \begin{cases} \int_{-\infty}^{x} f(\xi)\mathrm{d}\xi & \text{(spojitá nv)}, \\ \sum_{i \le x} f(i) & \text{(diskrétní nv)}. \end{cases}$$

Poznámka: Použijeme-li definici distribuční funkce a definici hp, můžeme pravděpodobnost intervalu (a, b) vyjádřit takto

$$P(X \in (a,b)) = P(X < b) - P(X < a) = F(b) - F(a) = \int_{-\infty}^{b} f(x)dx - \int_{-\infty}^{a} f(x)dx = \int_{a}^{b} f(x)dx.$$

Vidíme tedy, že platí:

♣ Pravděpodobnost intervalu $(a,\, b)$ hodnot spojité náhodné veličiny je dána integrálem $\int_a^b f(x)\mathrm{d}x$ z hp přes tento interval.

Poznámka: Obdobné tvrzení platí i pro pravděpodobnost množiny hodnot diskrétní náhodné veličiny:

♣ Pravděpodobnost množiny hodnot $\{x_i\}$, $i \in I$ diskrétní náhodné veličiny je dána součtem pravděpodobností $\sum_{i \in I} f(x_i)$ těchto hodnot.

Sdružená distribuční funkce $F_{XY}(x, y)$ pro dvě nv X a Y je definována vztahem

$$F_{XY}(x, y) = P(X \le x, Y \le y) = P([X \le x] \wedge [Y \le y]),$$

ve kterém požadujeme splnění obou nerovností současně.

Sdružená hustota pravděpodobnosti náhodných veličin X a Y je funkce $f(x, y)$ dvou reálných proměnných, pro kterou platí

$$F_{XY}(x, y) = \begin{cases} \int_{-\infty}^{x} \int_{-\infty}^{y} f(\xi, \eta)\mathrm{d}\xi\mathrm{d}\eta & \text{(spojitá nv)}, \\ \sum_{i \le x} \sum_{j \le y} f(i, j) & \text{(diskrétní nv)}. \end{cases}$$

♣ Pravděpodobnost, že náhodná veličina padne do dané množiny, je dána integrálem z hp přes tuto množinu.

Marginální hustota pravděpodobnosti $f(x)$ ze sdružené hp $f(x, y)$ je

$$f(x) = \begin{cases} \int_{-\infty}^{\infty} f(x, y)\mathrm{d}y & \text{(spojitá nv)}, \\ \sum_{x \in Y} f(x, y) & \text{(diskrétní nv)}. \end{cases}$$

♣ Je-li hp kopeček písku, pak marginální hp dostaneme rovnoměrným nahrnutím kopečku do jedné z os soustavy souřadnic.

Podmíněná hustota pravděpodobnosti $f(x|y)$ je funkce jedné reálné proměnné x s parametrem y, definovaná vztahem

$$f(x|y) = \frac{f(x,y)}{f(y)}.$$

Podmíněná hp je úměrná sdružené. Jmenovatel v předchozí definici (který nezávisí na proměnné x) představuje jen normování na integrál (součet) roven jedné.

♣ Je-li hp kopeček písku, pak podmíněnou hp dostaneme jako řez povrchu kopečku rovinou kolmou k podstavě a rovnoběžnou s jednou osou souřadnic, protínající druhou osu v bodě specifikovaném v podmínce hp.

3.2 Diskrétní rozdělení

Na příkladech ukážeme základní nástroje a jejich vnitřní strukturu pro popis diskrétních náhodných veličin.

Jednorozměrná nv

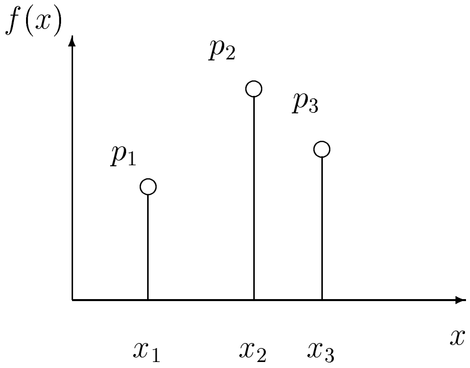
▷ `Zadání`

Diskrétní nv má za obor hodnot konečnou (spočetnou) množinu nezáporných reálných čísel. Její rozdělení (hp) je proto dáno posloupností hodnot a posloupností odpovídajících pravděpodobností. Hodnoty pravděpodobností lze zadat tabulkou, grafem nebo jako posloupnost předpisem. Uvedeme příklady jednotlivých zadání.

Tabulka:

x	x_1	x_2	x_3
$f(x)$	p_1	p_2	p_3

Graf:



Předpis: např. alternativní rozdělení

$$f(x) = p^x(1-p)^{1-x}, \quad x \in \{0, 1\}.$$

▷ `Význam`
Diskrétní hp $f_X(x)$ udává pravděpodobnosti jednotlivých hodnot diskrétní náhodné veličiny X

$$f(x_i) = f_X(x_i) = P(X = x_i) = p_i$$

▷ `Platí`

$$\sum_i p_i = 1$$

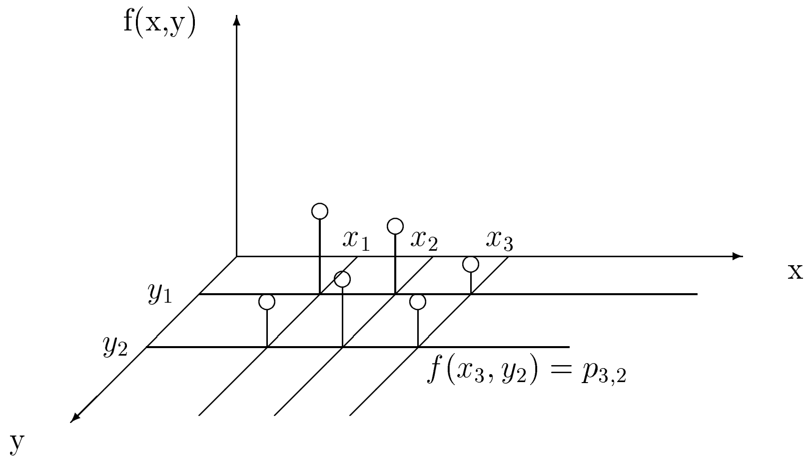
Dvojrozměrná nv

▷ `Zadání`
Dvojrozměrná diskrétní hp je funkce dvou proměnných, které každé uspořádané dvojici (x, y) přiřazuje pravděpodobnost, že při současné realizaci dvou nv X, Y nabude nv X hodnotu x a nv Y hodnotu y. Tato funkce je reprezentována maticí a lze ji opět vyjádřit v tabulce, grafem nebo předpisem.

Tabulka:

$(x,\, y)$	y_1	y_2
x_1	$p_{1,1}$	$p_{1,2}$
x_2	$p_{2,1}$	$p_{2,2}$
x_3	$p_{3,1}$	$p_{3,2}$

Graf:



Předpis: např. rozdělení beta

$$f(x,y) = \frac{1}{B(x,y)} \alpha^x (1-\alpha)^y, \quad x, y \in N^+, \quad \alpha \in \langle 0, 1 \rangle.$$

▷ `Význam`

Je popsán přímo zadáním

$$f(x,y) = P([X = x] \wedge [Y = y]).$$

▷ `Platí`

$$\sum_i \sum_j p_{i,j} = 1.$$

▷ `Marginální pravděpodobnosti`

Jsou pravděpodobnosti hodnot jedné nv, bez ohledu na hodnoty druhé nv. Určí se jako součty ve sloupcích nebo řádcích matice, reprezentující sdruženou hp.

$$f(y_j) = \sum_i p_{i,j}, \quad \forall j \quad \text{nebo} \quad f(x_i) = \sum_j p_{i,j}, \quad \forall i.$$

▷ `Podmíněné pravděpodobnosti`

Jsou pravděpodobnosti hodnot jedné nv, jestliže je pevně zvolená hodnota druhé nv. Určí se jako normalizované řezy sdruženou hp, tedy jako sloupce nebo řádky příslušné matice

$$f(x_i|Y = y_j) = \frac{p_{i,j}}{\sum_i p_{i,j}}, \quad \forall i \quad \text{nebo} \quad f(y_j|X = x_i) = \frac{p_{i,j}}{\sum_j p_{i,j}}, \quad \forall j.$$

3.3 Spojité rozdělení

Podobně jako pro diskrétní nv ukážeme strukturu základních nástrojů také pro příklady spojitých nv.

Jednorozměrná nv

▷ `Zadání`
Spojitá nv nabývá reálných hodnot. Její rozdělení je většinou zadáno pomocí hp, což je reálná funkce jedné reálné proměnné. Tu je možno zadat funkčním předpisem nebo grafem.

Předpis:

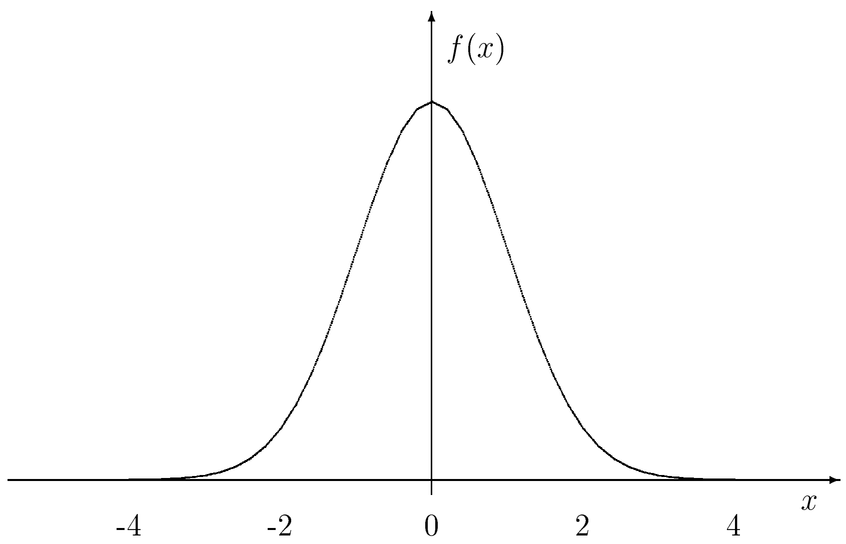
například normální rozdělení

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, x \in X,$$

nebo: exponenciální rozdělení

$$f(x) = \frac{1}{\delta} \mathrm{e}^{-\frac{x}{\delta}}, x \geq 0.$$

Graf: např. pro normální rozdělení



▷ `Význam`
Pomocí hp lze určit pravděpodobnost, že při realizaci náhodné veličiny bude její hodnota ležet v intervalu např. $(a,\ b)$

$$P(a \leq X \leq b) = \int_a^b f(x) \mathrm{d}x.$$

Speciálně, pro $a = -\infty$ dostaneme pravděpodobnost, která se rovná hodnotě distribuční funkce v bodě b

$$P(-\infty \leq X \leq b) = \int_{-\infty}^b f(x) \mathrm{d}x = F(b).$$

▷ Platí

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1.$$

Dvojrozměrná nv

▷ Zadání

Hp je v tomto případě reálná funkce dvou reálných proměnných a pro její zadání se nejspíše hodí funkční předpis.

Předpis: např. dvourozměrné rovnoměrné rozdělení

$$f(x,y) = \frac{1}{(b-a)(d-c)}\chi_{(a,b)\times(c,d)}(x,y),$$

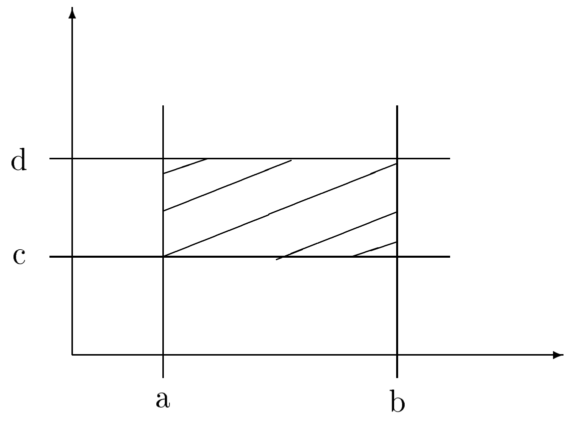
kde funkce χ_M je indikátor množiny M, tj.

$$\chi_M(a) = \begin{cases} 1 & \text{pro} \quad a \in M, \\ 0 & \text{pro} \quad a \notin M. \end{cases}$$

▷ Význam

Pomocí hp lze jednoduše určit pravděpodobnost obdélníkové oblasti $(a,b) \times (c,d)$

$$P([a \leq X \leq b] \wedge [c \leq Y \leq d]) = \int_a^b \int_c^d f(x,y)\mathrm{d}x\mathrm{d}y.$$



Obecně lze pomocí hp určit pravděpodobnost libovolné množiny $M \subset R^2$

$$P([x,y] \in M) = \iint_M f(x,y)\mathrm{d}x\mathrm{d}y.$$

▷ Platí

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1.$$

▷ `Marginální a podmíněná hp`
Pro tyto hp platí integrální definiční vztahy z úvodu kapitoly. Často jsme ale v situaci, kdy známe hp marginální $f(y)$ a podmíněnou $f(x|y)$ a potřebujeme vypočítat "druhou" marginální $f(x)$ a "opačnou" podmíněnou $f(y|x)$ hp. Potom v obecném vzorci použijeme řetězové pravidlo $f(x,y) = f(x|y)f(x)$ a dostaneme

Marginální hp		Podmíněná hp
$f(x) = \int f(x|y)f(y)\mathrm{d}y$	a	$f(y|x) = \dfrac{f(x|y)f(y)}{\int f(x|y)f(y)\mathrm{d}y}$

Další možnost, jak ze sdružené hp dostat marginální a podmíněnou hp, je rozložit sdruženou hp na součin dvou členů s jednotkovými integrály tak, aby první člen závisel jen na jedné proměnné. Tento člen potom představuje marginální hp, zbývající člen je podmíněná hp. V případě normálního rozdělení, kdy v exponentu je kvadratická forma, lze pro rozklad použít doplnění na čtverec. Tuto situaci ilustrujeme na následujícím příkladě.

`Příklad` [`Dvojrozměrné normální rozdělení`]

Sdružená hustota pravděpodobnosti

Hustotu pravděpodobnosti je zvykem psát ve dvojím různém vyjádření: skalárně a vektorově. Protože normální rozdělení je velmi časté a oba zápisy hp se podle okolností užívají, napíšeme obě verze.

▷ `Vektorový zápis`

$$f(z) = \frac{1}{2\pi\sqrt{|R|}}\exp\left\{-\frac{1}{2}(z-\mu)^T R^{-1}(z-\mu)\right\}$$

kde

$$z = \begin{bmatrix} x \\ y \end{bmatrix}, \qquad \mu = E[z] = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

$$R = D[z] = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix},$$

$|R| = \sigma_1^2\sigma_2^2 - (\sigma_{12})^2$ je determinant kovarianční matice R,

σ_1^2, σ_2^2 jsou rozptyly X a Y: $\sigma_1^2 = E[(X-\mu_1)^2]$, $\sigma_2^2 = E[(Y-\mu_2)^2]$

σ_{12} je kovariance nv X a Y: $\sigma_{12} = E[(X-\mu_1)(Y-\mu_2)]$

▷ `Skalární zápis`

$$f(z)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\},$$

kde

$\rho=\frac{\sigma_{12}}{\sigma_1\sigma_2}$ je korelační koeficient.

Návod k odvození:

$$\begin{bmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{12} & \sigma_2^2\end{bmatrix}^{-1}=\frac{1}{1-\rho^2}\begin{bmatrix}\frac{1}{\sigma_1^2} & -\frac{\rho}{\sigma_1\sigma_2}\\ -\frac{\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2}\end{bmatrix}$$

`Cvičení:` Ověřit inverzi a souvislost obou vyjádření hp.

Marginální a podmíněná hustota pravděpodobnosti
Sdružené rozdělení:

$$f(x,y)=\mathcal{N}((\mu_1,\mu_2),R),\quad R=\begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix}.$$

Marginální rozdělení:

$$f(x)=\mathcal{N}(\mu_1,\sigma_1^2),\qquad f(y)=\mathcal{N}(\mu_2,\sigma_2^2).$$

Podmíněné rozdělení:

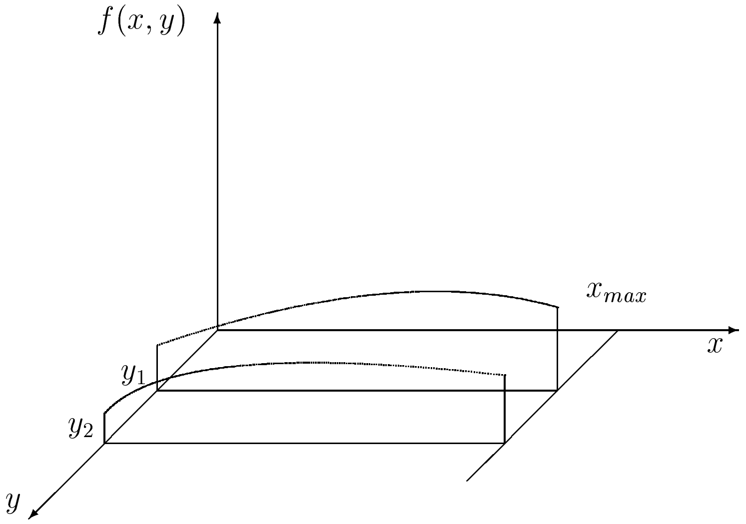
$$f(x|y)=\mathcal{N}\left(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2),\ \sigma_1^2(1-\rho^2)\right),$$

$$f(y|x)=\mathcal{N}\left(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(y-\mu_1),\ \sigma_2^2(1-\rho^2)\right).$$

`Cvičení:` Napsat hustoty a ověřit jejich tvar.
Návod: doplnění na čtverec, nejdříve zkusit pro $\mu=0$.

3.4 Sdružené kombinované rozdělení

Sdružené rozdělení dvou náhodných veličin, kde jedna nv je diskrétní a druhá spojitá, nazýváme kombinovaným rozdělením. Hp je potom funkce dvou proměnných, z nichž jedna je diskrétní a druhá spojitá. Příklad takové hp s první nv diskrétní a druhou spojitou je uveden na následujícím obrázku.



Opět musí platit základní podmínka o normování pravděpodobnosti. Ta má pro hp na obrázku následující tvar:

$$\sum_{y \in \{y_1, y_2\}} \int_0^{x_1} f(x,y) \mathrm{d}x = 1.$$

3.5 Operace s hustotami pravděpodobností

Uvedeme ty operace s hp, které budou dále základním pracovním nástrojem při odvození algoritmu bayesovské identifikace a řízení. Proto se jimi budeme zabývat poněkud podrobněji, než bývá v pravděpodobnosti zvykem. Budeme se snažit nejen formulovat jejich matematický smysl, ale také demonstrovat jejich praktickou interpretaci. Tu ukážeme na velmi jednoduchých reálných situacích. Nejprve poukážeme na jednotlivé pojmy, které se k takovým operacím vztahují.

Na následující situaci budeme tyto pojmy demonstrovat:

Čekám návštěvu. Přijde buď Tonda, nebo Franta (někdo určitě přijde, ale nepřijdou současně), přitom obě možnosti jsou stejně pravděpodobné. Tonda mi přinese koláč s pravděpodobností 0.4, Franta s pravděpodobností 0.3.

Otázka A S jakou pravděpodobností dostanu koláč od Tondy?

Otázka B S jakou pravděpodobností dostanu koláč?

Otázka C Když jsem po zazvonění otevřel, byl za dveřmi koláč (kamarád se schoval). S jakou pravděpodobností to byl Tonda?

Řešení A Kdybych věděl určitě, že přijde Tonda, byla by míra mého očekávání koláče od Tondy rovna pravděpodobnosti 0.4. Protože je ale pravděpodobnost Tondova příchodu jen polovina (0.5), bude i míra mého očekávání poloviční, a tedy hledaná pravděpodobnost je $0.4 \cdot 0.5 = 0.2$.

Řešení B Na rozdíl od předchozího případu mě nyní zajímá, jaká je možnost, že dostanu koláč vůbec, od Tondy nebo od Franty. Protože přijde právě jeden z kamarádů, bude moje míra důvěry v koláč dána mírou důvěry v koláč od Tondy plus mírou důvěry v koláč od Franty. Hledaná pravděpodobnost je tedy $0.4\,0.5 + 0.3 \cdot 0.5 = 0.35$.

Řešení C Moje očekávání koláče od Tondy a od Franty je v poměru $0.4 : 0.3$. Protože jsem koláč již dostal, musí být celková míra koláče rovna jedné. Proto musí platit, že pravděpodobnost, že to byl Tonda, kdo koláč donesl je $0.4/(0.4 + 0.3)$. Pravděpodobnost Frantova koláče je $0.3/(0.4 + 0.3)$. Toto řešení vyhovuje, neboť poměr pravděpodobností je správný a jejich součet je jedna.

♣ Bayesovství chápe pravděpodobnost jevu jako míru důvěry, která je splnění jevu přidělena (z celkového množství jednotkové důvěry). Přiděluje se ze dvou zdrojů:
1. z prvotní (apriorní) znalosti zkoumaného problému a
2. z dat, naměřených při experimentu.

V našem příkladě pravděpodobnosti příchodu jednotlivých kamarádů, i pravděpodobnosti koláče od jednotlivých kamarádů vyjadřují naši apriorní znalost. Zjištění koláče za dveřmi představuje naměření dat.

Prvá a druhá úloha je jen kombinací apriorní znalosti, třetí úloha již využívá informaci v naměřených datech (koláč za dveřmi).

Nyní ukážeme pravděpodobnostní interpretaci jednotlivých pojmů.

Definujeme *jevy*

$N \equiv$ přijde Návštěvník se jménem ..., $N \in \{t = \text{Tonda}, f = \text{Franta}\}$
$D \equiv$ dostanu Dárek ..., $D \in \{k = \text{Koláč}, n = \text{Nic}\}$

pravděpodobnosti jevů, které jsme zavedli jsou

$$P(N = t) = P(t) = 0.5, \qquad P(N = f) = P(f) = 0.5.$$

Pravděpodobnosti hodnot jevu D zatím neznáme, nebyly zadány. O pravděpodobnostech koláče se však v zadání hovoří. Přijde-li Tonda, přinese koláč s pravděpodobností 0.4 atd. To jsou pravděpodobnosti, vázané na splnění určité podmínky. Tato podmínka je formulována za pomoci jevu N. Jedná se o podmíněné pravděpodobnosti, a ty se vztahují na podmíněné jevy – jestliže přijde Tonda, dárek bude...

Definujeme *podmíněné jevy*

$D|N \equiv$ Dárek, za podmínky, že známe návštěvníka $D|N \in \{k = \text{koláč}, n = \text{nic}\}$
Např. $D = k|N = t$ znamená: dárek bude koláč, jestliže vím, že přišel Tonda.

pravděpodobnosti podmíněných jevů jsou zadány

$$P(D = k|N = t) = P(k|t) = 0.4, \qquad P(D = k|N = f) = P(k|f) = 0.3$$

$$P(D = n|N = t) = P(n|t) = 0.6, \qquad P(D = n|N = f) = P(n|f) = 0.7$$

Poznámka: V předchozích zápisech jsme použili symbolický (zkrácený) zápis pravděpodobnosti $P(x)$ místo $P(X = x)$. Protože reálná čísla, kterým samotné x je, nemají žádnou pravděpodobnost, má tento zápis smysl jen v daném kontextu, tj. je-li definována náhodná veličina X a víme-li, že se o ní v danou chvíli jedná. Podobný zápis je běžný i pro hp a my jej budeme v dalších kapitolách používat.

V otázkách a při jejich řešení se vyskytují další pojmy. Je to především výrok "bude koláč od Tondy". Pozor! Nezaměnit s výrokem "bude koláč, jestliže vím, že přišel Tonda". V tomto druhém výroku víme, že Tonda přišel a předmětem zkoumání je jen to, zda přinesl koláč nebo nic. V prvním výroku nevíme ani jedno ani druhé a zajímáme se o to, zda obě tyto skutečnosti nastaly současně. Výrok "bude koláč od Tondy" je jedním z možných výsledků *sdruženého jevu* (D, N) a jeho pravděpodobnost je *sdružená pravděpodobnost*. Výrok "bude koláč, jestliže vím, že přišel Tonda" je výrokem podmíněným.

Dalším výrokem je výrok "bude koláč", a je nám jedno od koho. Tento výrok je výsledkem *marginálního jevu* D (vzhledem k předchozímu sdruženému jevu (D, N)). Z tohoto hlediska je marginální jev také jev N.

Nakonec se objevuje výrok "návštěva je Tonda, jestliže jsem dostal koláč". To je opět výsledek podmíněného jevu, ale opačně podmíněného, než ten, jehož pravděpodobnosti jsou zadány v příkladě. Tento jev označíme $N|D$ a týká se pouze jména návštěvníka, když výsledek jevu dárek je znám.

Jednotlivé otázky a jimi formulované úlohy jsou interpretacemi základních vztahů Bayesovské statistiky.

A. Řetězové pravidlo

$$P(D, N) = P(D|N)P(N)$$

tj. pravděpodobnost, že bude dárek od návštěvníka = pravděpodobnost, že bude dárek, když přijde určitý návštěvník × pravděpodobnost, že přijde určitý návštěvník

Zápis pro konkrétní hodnoty je např.:

$$P(k, t) = P(k|t)P(t)$$

tj. pravděpodobnost, že bude koláč od Tondy = pravděpodobnost, že bude koláč, když přijde Tonda × pravděpodobnost, že přijde Tonda.

B. Úplná pravděpodobnost

$$P(D) = \sum_N P(D|N)P(N)$$

nebo podrobněji

$$P(D) = \sum_{n \in N} P(D|N = n)P(N = n)$$

znamená: pravděpodobnost dárku = pravděpodobnost dárku, přijde-li určitý návštěvník × pravděpodobnost určitého návštěvníka a to sečteno přes všechny možné návštěvníky.

Pro konkrétní hodnoty je např.:

$$P(k) = P(k|t)P(t) + P(k|f)P(f),$$

který znamená: pravděpodobnost koláče = pravděpodobnost koláče, přijde-li Tonda × pravděpodobnost Tondy + pravděpodobnost koláče, přijde-li Franta × pravděpodobnost Franty.

Poznámka: Předchozí výpočet je vlastně výpočtem marginální pravděpodobnosti $P(D) = \sum_N P(D,N)$. *Pokud sdruženou pravděpodobnost ještě rozložíme podle řetězového pravidla, nazývá se tento vztah pravidlem úplné pravděpodobnosti.*

C. Bayesův vzorec

$$P(N|D) = \frac{P(D|N)P(N)}{\sum_N P(D|N)P(N)}.$$

Tento vzorec říká jaká je pravděpodobnost, že přichází určitý návštěvník, když víme, jaký byl dárek. Pro výpočet se používá opačná podmíněná pravděpodobnost, tedy pravděpodobnosti dárků za podmínky návštěvníků a pravděpodobnosti příchodu jednotlivých návštěvníků.

Pro určité hodnoty tento vztah může být

$$P(t|k) = \frac{P(k|t)P(t)}{P(k|t)P(t) + P(k|f)P(f)}.$$

Poznámka: Jmenovatel Bayesova vzorce je vzorcem pro úplnou pravděpodobnost. Tedy jmenovatel je vlastně marginální pravděpodobnost. Odvození Bayesova vzorce v tomto tvaru je zvláště jednoduché. Stačí dvojím způsobem rozložit sdruženou pravděpodobnost

$$P(D|N)P(N) = P(D,N) = P(N,D) = P(N|D)P(D),$$

a vyjádřit požadovanou podmíněnou pravděpodobnost.

`Příklad` [`Operace s hp`]

Na vybraném místě silniční komunikace byla sledována rychlost projíždějících vozidel. Bylo zjištěno, že 30% projíždějících vozidel má cizí SPZ a zbytek naší. Z cizích vozidel překročilo povolenou rychlost 20% vozidel, z našich 25%. Určete:

A pravděpodobnost, že právě projíždějící vozidlo je cizí a překračuje povolenou rychlost.

B pravděpodobnost, že právě projíždějící vozidlo překračuje povolenou rychlost.

C Zjistili jsme, že právě projíždějící vozidlo překračuje povolenou rychlost. Jaká je pravděpodobnost, že je to cizí vozidlo.

Označíme *jevy* a jejich *hodnoty*:

$Z \equiv$ vozidlo má Značku ..., $Z \in \{c = \text{cizí}, n = \text{naší}\}$

$R \equiv$ vozidlo jede Rychlostí ..., $R \in \{v = \text{Vyšší}, s = \text{Správnou}\}$

pravděpodobnosti hodnot jevu Z jsou

$$P(c) = 0.3, \qquad P(n) = 0.7.$$

Podmíněný jev je $R|Z$ a *pravděpodobnosti* jeho hodnot jsou

$$P(v|c) = 0.2, \qquad P(v|n) = 0.25,$$

$$P(s|c) = 0.8, \qquad P(s|n) = 0.75.$$

Řešení A Ptáme se na sdruženou pravděpodobnost $(Z = c, R = v)$, která je dána řetězovým pravidlem

$$P(Z,R) = P(R|Z)P(Z).$$

Po dosazení hodnot a jejich pravděpodobností dostaneme

$$P(c,v) = P(v|c)P(c) = 0.2 \cdot 0.3 = 0.06.$$

Řešení B Hledáme marginální pravděpodobnost pomocí vzorce pro úplnou pravděpodobnost

$$P(Z,R) = \sum_Z P(R|Z)P(Z).$$

Po dosazení

$$P(v) = P(v|c)P(c) + P(v|n)P(n) = 0.2 \cdot 0.3 + 0.25 \cdot 0.7 = 0.235.$$

Řešení C Máme určit obráceně podmíněnou pravděpodobnost a k tomu použijeme Bayesův vzorec

$$P(Z|R) = \frac{P(R|Z)P(Z)}{\sum_Z P(R|Z)P(Z)}.$$

Dosadíme a dostaneme

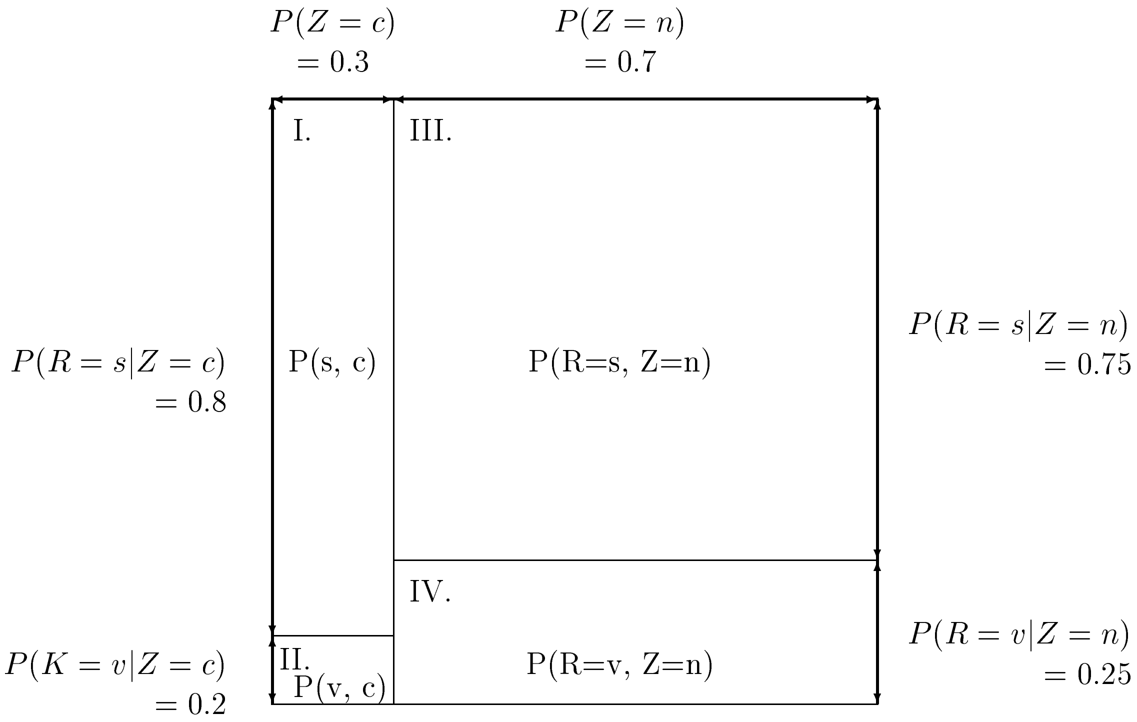
$$P(c|v) = \frac{P(v|c)P(c)}{P(v|c)P(c) + P(v|n)P(n)} = \frac{0.2 \cdot 0.3 + 0.25 \cdot 0.7}{0.2 \cdot 0.3 + 0.25 \cdot 0.7 + 0.25 \cdot 0.7} \doteq 0.255.$$

Strukturu předchozích operací budeme ještě ilustrovat pomocí grafických schémat. Použijeme dvojí vyjádření, vhodné pro demonstraci podmíněné pravděpodobnosti a operací, s ní spojených. Jedná se o znázornění pravděpodobností pomocí obdélníkového schematu a pomocí pravděpodobnostního stromu.

Obdélníkové schema

Při tomto zobrazení vycházíme z *řetězového pravidla*, které říká, že sdružená pravděpodobnost je dána jako součin podmíněné a marginální pravděpodobnosti. Zobrazíme-li tedy podmíněnou a marginální pravděpodobnost jako odvěsny obdélníku, bude jeho plocha představovat sdruženou pravděpodobnost. Základní prostor pro sdruženou pravděpodobnost bude čtverec o straně jedna. V tomto základním čtverci lze zobrazovat jednotlivé typy pravděpodobností a demonstrovat operace mezi nimi.

V našem příkladě budeme na vodorovné straně zobrazovat pravděpodobnost jevu Z - značka a na svislých stranách pravděpodobnosti podmíněného jevu $R|Z$ - rychlost podmíněná značkou. Dostaneme následující schema



Základní čtverec se rozdělil na čtyři obdélníky I. až IV., které mají význam sdružených pravděpodobností.

I.	cizí auto, které nepřekročilo rychlost
II.	cizí auto, které překročilo rychlost
III.	naše auto, které nepřekročilo rychlost
IV.	naše auto, které překročilo rychlost

Z těchto pravděpodobností můžeme vystavět libovolnou pravděpodobnost, týkající se jevů Z a R. Na tomto místě bych rád citoval větu svého školitele a předního odborníka v oblasti bayesovské teorie řízení u nás Ing. V. Peterky, DrSc:

♣ Známe-li (tj. umíme-li zkonstruovat) sdruženou pravděpodobnost, známe všechno.

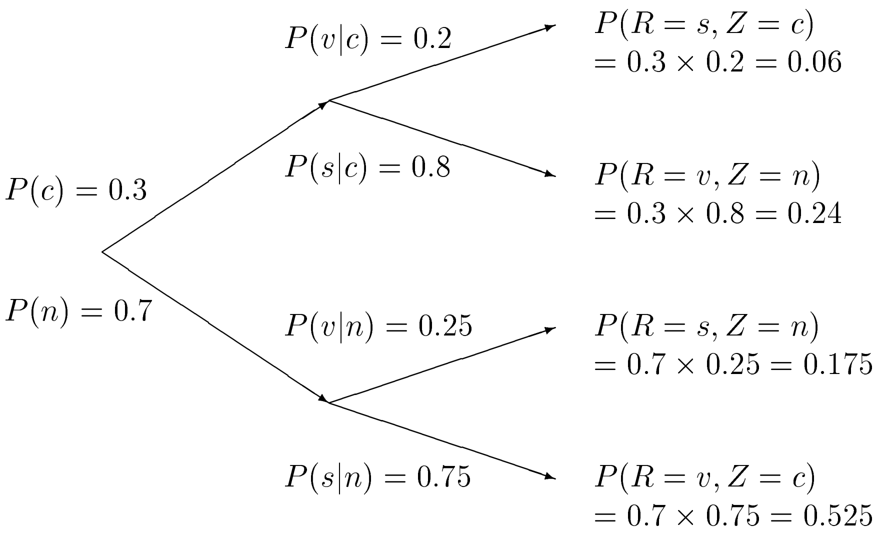
Marginální pravděpodobnost např. pro jev $R = v$ je dána součtem všech obdélníků, v nichž se vyskytuje jako odvěsna pravděpodobnost hodnoty v. Jsou to obdélníky I. a III. Vyjádříme-li jejich plochy jako součiny odvěsen a plochy sečteme, dostáváme vzorec pro *úplnou pravděpodobnost*.

Obrácená *podmíněná pravděpodobnost* např. jevu $Z|R = c|v$ se odvodí takto: Protože jev $R = v$ nastal, musí mít pravděpodobnost 1. To znamená, že musíme definovat nový základní čtverec pro jevy $Z|R = v \in \{c|v,\, n|v\}$ a přitom poměr pravděpodobností hodnot $c|v$ a $n|v$ by měl být dán poměrem příslušných sdružených pravděpodobností jevů (c, v) a (n, v). Tedy $P(c|v) : P(n|v) = P(c, v) : P(n, v) = 0.06 : 0.175$. Tento poměr je dán obsahy horních obdélníků. Stačí tedy jejich obsahy normovat tak, aby součet byl jedna a obdržíme příslušné podmíněné pravděpodobnosti. Ta, kterou zjišťujeme, je dána levým obdélníkem. Její normování provedeme tak, že plochu tohoto obdélníku dělíme součtem ploch obou horních obdélníků. Tomu, při vyjádření ploch součinem stran obdélníků, odpovídá Bayesův vztah.

Pravděpodobnostní strom

Jiným nástrojem nejen pro vyjádření ale i pro výpočet sdružených pravděpodobností pomocí podmíněných a marginálních, je pravděpodobnostní strom. Je také založen na řetězovém pravidlu a je dokonce vhodný i pro jeho vícenásobné použití. Strom začíná v jednom bodě a postupně se větví podle hodnot jednotlivých jevů. Hodnoty předchozích jevů při větvení se přitom dostávají do podmínek následujících jevů.

V našem příkladě bude strom velmi jednoduchý. První větvení je podle hodnot jevu Z na větev $Z = c$ a $Z = n$. K větvím zároveň připisujeme pravděpodobnosti těchto jevů. Na konci obou větví dojde k dalšímu větvení, podle hodnot jevu R. Protože ale konci horní větve přísluší $Z = c$, budou další větve z tohoto místa příslušet podmíněným jevům $R|(Z = c)$. Podobně větvení ze spodního bodu bude reprezentovat podmíněné jevy $R|(Z = n)$. Na konci dostáváme 4 body – konce větvení. Každému bodu přiřadíme pravděpodobnost, rovnu součinu jednotlivých pravděpodobností od začátku větvení. Podle řetězového pravidla konce větvení odpovídají jednotlivým hodnotám sdruženého jevu (Z, R). A podobně jako v předchozím případě: známe-li sdružené pravděpodobnosti, známe všechno.



Kapitola 4

Dodatek

4.1 Speciální funkce

Uvedeme definice a nejdůležitější vlastnosti některých speciálních funkcí, které budeme při práci s hp potřebovat.

Funkce Gamma

Definice:
pro z reálné (komplexní)

$$\Gamma(z) = \int_0^\infty \mathrm{e}^{-z} t^{z-1} \mathrm{d}t \tag{4.1}$$

Vlastnosti:
pro $z = n \in N$ platí

$$\Gamma(n+1) = n!$$

Funkce Beta

Definice:

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1} \mathrm{d}t$$

Vlastnosti:

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

$$B(x+1,y) = \frac{x}{x+y} B(x,y)$$

Odvození pro $x = m \in N$ a $y = n \in N$

$$B(m,n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)} = \frac{(m-1)!(n-1)!}{(m+n-1)!}$$

$$B(m+1,n) = \frac{(m)!(n-1)!}{(m+n)!} = \frac{m(m-1)!(n-1)!}{(m+n)(m+n-1)!} = \frac{m}{m+n} B(m,n)$$

4.2 Operátor průměrování a jeho vlastnosti

Definice operátoru průměrování

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i. \tag{4.2}$$

Pro tento operátor lze snadno odvodit následující pravidla: (x, y jsou posloupnosti, a je konstanta)

Průměr násobku

$$\overline{ax} = a\overline{x}.$$

Odvození: $\overline{ax} = \frac{1}{n}\sum_{i=1}^{N} ax_i = a\frac{1}{N}\sum_{i=1}^{N} x_i = a\overline{x}.$

Průměr součtu

$$\overline{x+y} = \overline{x} + \overline{y}.$$

Odvození: $\overline{x+y} = \frac{1}{N}\sum_{i=1}^{N}(x_i + y_i) = \frac{1}{N}\sum_{i=1}^{N} x_i + \frac{1}{N}\sum_{i=1}^{N} y_i = \overline{x} + \overline{y}.$

Průměr z konstanty

$$\overline{a} = a.$$

Odvození: $\overline{a} = \frac{1}{N}\sum_{i=1}^{N} a = \frac{1}{N}Na = a.$

Průměr z průměru

$$\overline{\overline{x}} = \overline{x}.$$

Odvození: stejné jako předchozí, protože $\overline{x}$ je konstanta.

Průměr součinu a kvadrátu

$$\overline{xy} \neq \overline{x}{\cdot}\overline{y}, \qquad \overline{x^2} \neq \overline{x}^2.$$

Průměr kvadrátů odchylek od průměru

$$\overline{(x-\overline{x})^2} = \overline{x^2} - \overline{x}^2.$$

Odvození: $\overline{(x-\overline{x})^2} = \overline{x^2 - 2x\overline{x} + \overline{x}^2} = \overline{x^2} - \overline{2x\overline{x}} + \overline{\overline{x}^2} = \overline{x^2} - \overline{x}^2.$

Průměr součinů odchylek od průměrů

$$\overline{(x-\overline{x})(y-\overline{y})} = \overline{xy} - \overline{x} \cdot \overline{y}.$$

Odvození: obdobné jako předchozí.

`Příklad` [`Střední hodnota modelu pomocí operátoru průměrování`]
Pro $N = 3$ a model $y_\tau = bu_\tau + c + e_\tau$ platí

$$\overline{y} = \frac{1}{3}\sum_{\tau=1}^{3} y_i = \frac{1}{3}\sum_{\tau=1}^{3}(bu_i + c + e_i) = b\frac{1}{3}\sum_{\tau=1}^{3} u_i + \frac{1}{3}\sum_{\tau=1}^{3} c + \frac{1}{3}\sum_{\tau=1}^{3} e_i = b\overline{u} + c + \overline{e}$$

nebo totéž pomocí operátoru průměrování

$$\overline{y} = \overline{bu + c + e} = \overline{bu} + \overline{c} + \overline{e} = b\overline{u} + c + \overline{e}.$$

4.3 Doplnění na čtverec

Tato úprava se používá velmi často a k různým účelům. Nejčastěji pro hledání extrému kvadratické funkce a pro integraci normální hp. Může být použita jak pro skalární, tak i vektorové funkce.

Pro skalární koeficienty
Je dána kvadratická forma $ax^2 - 2bx + c$. Její doplnění na čtverec v proměnné x je následující

$$ax^2 - 2bx + c = a\left[x^2 - 2\frac{b}{a}x + \left(\frac{b}{a}\right)^2 - \left(\frac{b}{a}\right)^2\right] + c = a\left(x - \frac{b}{a}\right)^2 + c - \frac{b^2}{a}$$

Pro maticové koeficienty
Nechť A je čtvercová, symetrická pozitivně definitní matice stupně n, B, x jsou n-rozměrné sloupcové vektory a C je skalár. Potom platí následující úprava kvadratické formy na úplný čtverec

$$x^T Ax - x^T B - B^T x + C = (x - A^{-1}B)^T A(x - A^{-1}B) + C - B^T A^{-1} B$$

$\triangleright$ `Odvození`
Toto doplnění na čtverec je častou úpravou. Proto naznačíme jeho postup.

1. Vezmeme kvadratickou část $x^T Ax$. Matice A je jádro kvadrátu, a musí být umístěna uvnitř každého členu.

2. Vezmeme lineární člen, kde vlevo stojí x^T a vnutíme doprostřed matici A

$$x^T B = x^T A A^{-1} B,$$

Tím jsme správně upravili první tři členy. Dostáváme

$$x^T Ax - x^T A A^{-1} B - B^T A^{-1} Ax + ...$$

3. Část upraveného lineárního členu na pravé straně je $A^{-1}B$. Z ní sestavíme požadovaný čtvrtý člen lineární formy tak, aby tyto čtyři členy bylo možno doplnit na čtverec. Požadovaný čtvrtý člen je

$$(A^{-1}B)^T A(A^{-1}B) = B^T A^{-1} B$$

Protože jsme ale požadovaný čtvrtý člen přidali, musíme jej zase odečíst. K němu se ještě přidá původní čtvrtý člen C a vytvoří skalární člen, který zůstává po minimalizaci (anulování kvadrátu) jako zbytek.

4.4 Algoritmus optimálního řízení soustavy 1. řádu

```
clc, clear, clear all
n=30;            % horizont
a=.9; b=.2;      % parametry
om=.1;           % penalizace u
y(1)=10;         % počáteční y
r=zeros(1,n+1); s=zeros(1,n+1); s(n+1)=0; % poč. s

% konstrukce řídícího zákona (výpočet r, s)
for t=n:-1:1
  s1=1+s(t+1);
  r(t)=s1*a*b/(s1*b*b+om);
  s(t)=a*om/b*r(t);
end
w=0; % žádaná hodnota výstupu (na tu řídíme)

% simulace s řízením soustavy
for t=1:n
  if t==10, w=20; end
  if t==20, w=-15; end
  u(t+1)=-r(t)*(y(t)-w);  % regulátor
  y(t+1)=a*y(t)+b*u(t+1); % soustava
end

% zobrazení výsledků
t1=1:(t+1); plot(t1,u,'o',t1,y)
```

4.5 Algoritmus identifikace – základní vzorce

```
% Identifikace metodou nejmenších čtverců
%         (základní vzorce)
% -------------------------------------
clc, clear, clear all
randn('seed',17);
n=30;              % horizont identifikace
as=.9; bs=.2;      % parametry simulace
yy=zeros(n,1); uu=zeros(n,1);

% simulace a měření dat
for t=1:n
  e=1*randn;       % šum s amplitudou 1
  u=t;             % vstup (běží v čase)
  y=as*u+bs+e;     % simulovaný model
  yy(t)=y;         % data y
  uu(t)=u;         % data u
end

% výpočet parametrů a zbytku kriteria
mu=sum(uu)/n;      % střední hodnota u
my=sum(yy)/n;      % střední hodnota y
du=uu-mu;          % odchylka u
dy=yy-my;          % odchylka y
s2u=du'*du/n;      % rozptyl u
s2y=dy'*dy/n;      % rozptyl y
syu=dy'*du/n;      % kovariance y,u

% regresní koeficienty
a=syu/s2u;         % směrnice reg. přímky
b=my-a*mu;         % absolutní člen reg. přímky
la=s2y-a*syu;      % zbytek kriteria (rozptyl šumu)

% zobrazení výsledků
a,b,la             % tisk
yr=a*uu+b;         % reg. přímka
figure(1)
plot(uu,yy,'o',uu,yr) % graf dat a reg. přímky
```

4.6 Algoritmus identifikace soustavy 1. řádu

```
% Identifikace metodou nejmenších čtverců
%   skalární regresní model 1. řádu
% -------------------------------------
clc, clear, clear all
randn('seed',17);
```

```
n=30;             % horizont
a=.9; b=.2;       % parametry
yg=zeros(1,n); v=zeros(3);

for t=1:n
  % simulace
  e=1*randn;      % šum
  u=t;            % vstup (běží v čase)
  y=a*u+b+e;      % model
  yg(t)=y;        % y pro graf

  % update informační matice
  v=v+[y u 1]'*[y u 1];
  vy =v(1,1);     % rozdělení
  vyz=v(2:3,1);   %      informační
  vz =v(2:3,2:3); %      matice
end

% výpočet parametrů a zbytku kriteria
v=v/n;            % normování inf. mat. (není třeba)
th=inv(vz)*vyz;   % parametry
la=(vy-th'*vyz);  % zbytek kriteria

% zobrazení výsledků
th,la,v % tisk
tt=1:n;
yr=th(1)*tt+th(2); % reg. přímka
figure(1)
plot(tt,yg,'o',tt,yr)
```

4.7 Algoritmus identifikace soustavy řádu m

```
% Identifikace metodou nejmenších čtverců
%   skalární regresní model řádu m
% ---------------------------------------
clc, clear, clear all
randn('seed',17);

% zadání simulace
n=300;                      % horizont identifikace
m=2;                        % řád modelu
%  y(t-1) y(t-2)    u(t) u(t-1) u(t-2)  1   % regresní vektor (pro 2. řád)
a=[ .9    -.1];   b=[.5   -.3    .1];  c=1; % parametry simulace
% sestavení vektoru parametrů
th=b(1);
for i=1:m
  th=[th; a(i); b(i+1)];
```

```
end
th=[th; c];
% deklarace pomocných vektorů a matic
yg=zeros(1,n);              % vektror vystupů pro graf
mv=2*m+2; fi=zeros(mv,1);   % regresní vektor
mv1=mv+1; v=1e-8*eye(mv1);  % rozšířená informační matice

% simulační smyčka + sběr dat
for t=1:n
  % simulace
  e=.1*randn;               % generování šumu
  u=.2*randn;               % generování vstupu
  fi=[u; fi(1:2*m); 1];     % přepočet regresního vektoru
  y=th'*fi+e;               % simulace výstupu

  % sběr dat
  v=v+[y; fi]*[y; fi]';     % přepočet inf. matice

  fi=[y; fi(1:2*m); 1];     % přepočet regresního vektoru
  yg(t)=y;                  % y pro graf
end

% rozdělení informační matice
vy =v(1,1);
vyz=v(2:mv1,1);
vz =v(2:mv1,2:mv1);

% výpočet parametrů a zbytku kriteria
v=v/n;                      % normování inf. mat. (není třeba)
theta=inv(vz)*vyz;          % odhad parametrů
lambda=(vy-theta'*vyz);     % odhad rozptylu šumu

% zobrazení výsledků
figure(1), plot(yg)
[th, theta] % tisk
```

Literatura

[1] J. Anděl. *Matematická statistika.* SNTL, Praha, 1978.

[2] J. Hátle, J. Likeš. *Základy počtu pravděpodobnosti a matematické statistiky.* SNTL, Praha, 1974.

[3] I. Nagy, M. Kárný, J. Novovičová, and M. Valečková. Mixture-model identification in traffic control problems. Technical Report 1973, ÚTIA, Prague, 1999.

[4] M. Kárný, I. Nagy. Dynamical bayesian decision-making: part I. Technical Report 1971, ÚTIA, Prague, 1999.

[5] J. Novovičová. Pravděpodobnost a matematická statistika. Skripta FD ČVUT, Praha, 1999.

[6] V. Peterka. Bayesian approach to system identification. In P. Eykhoff, editor, *Trends and Progress in System Identification*, pages 239–304. Pergamon Press, Oxford, 1981.

[7] R. C. Rao. *Lineární metody statistické indukce a jejich aplikace.* ACADEMIA, Praha, 1978.