

Dynamické systémy

Filip Trojan

1. července 1999

Obsah

| | | |
|----------|---|-----------|
| 1 | ÚVOD | 5 |
| 2 | SPOJITÉ DYNAMICKÉ SYSTÉMY | 7 |
| 2.1 | Teorie | 7 |
| 2.2 | Ekonomické aplikace | 9 |
| 2.2.1 | Walrasův model tržní rovnováhy | 9 |
| 2.2.2 | Model IS-LM | 10 |
| 2.2.3 | Neoklasický model růstu | 12 |
| 2.2.4 | Lotkův-Volterrův model lovce a oběti | 14 |
| 3 | DISKRÉTNÍ DYNAMICKÉ SYSTÉMY | 15 |
| 3.1 | Teorie | 15 |
| 3.2 | Ekonomické aplikace | 19 |
| 3.2.1 | Samuelsonův hospodářský cyklus (1939) | 19 |
| 3.2.2 | Hickův příspěvek k teorii obchodních cyklů (1950) | 22 |
| 4 | SIMULACE DYNAMICKÝCH SYSTÉMŮ | 23 |
| 4.1 | Teorie | 23 |
| 4.2 | Příklady | 24 |
| 4.2.1 | Pindyck-Rubinfeldův třírovnicový model | 24 |
| 5 | OPTIMÁLNÍ ODHAD STAVŮ SYSTÉMU | 29 |
| 5.1 | Optimální estimátor | 29 |
| 5.2 | Odhad stavu dynamického systému | 39 |
| 5.3 | Kalmanův filtr | 41 |
| 5.4 | Smoothing | 45 |
| 5.5 | Rozšířený Kalmanův filtr | 46 |
| 5.6 | Odhad nelineárních systémů metodou Monte Carlo | 48 |
| 5.7 | Shrnutí | 51 |
| 5.8 | Příklady k procvičení | 51 |
| 6 | IDENTIFIKACE DYNAMICKÝCH SYSTÉMŮ | 53 |
| 6.1 | Úvod | 53 |
| 6.2 | Klasický přístup | 53 |
| 6.2.1 | Metoda nejmenších čtverců (LS) | 54 |
| 6.2.2 | Metoda maximální věrohodnosti | 56 |
| 6.3 | Bayesův přístup | 56 |

| | | |
|----------|--|-----------|
| 7 | TEORIE CHAOSU | 61 |
| A | Přehled maticové algebry | 63 |
| A.1 | Blokové matice | 63 |
| A.2 | Maticový kalkulus | 64 |
| B | Pravděpodobnost | 65 |
| B.1 | Základní pojmy | 65 |
| B.2 | Číselné charakteristiky (momenty) | 65 |
| B.3 | Normální rozdělení | 66 |
| B.4 | Podmíněná hustota pravděpodobnosti | 66 |
| B.5 | Podmíněné číselné charakteristiky | 67 |
| B.6 | Lineární regrese | 67 |

Kapitola 1

ÚVOD

Slovo *system* je jedním z nejčastěji používaných pojmů ve všech odvětvích vědy. Používá se pro označení pojmů abstraktních i konkrétních. Teorie systémů, která zaznamenala nebyvalý rozvoj zejména v posledních padesáti letech, zkoumá objekty i jevy ve vzájemných souvislostech.

Pojem systém je skutečně používán v nejrůznějších souvislostech. V tomto kurzu budeme pod pojmem dynamický systém chápat určitý časově neměnný vztah mezi okamžitými a minulými nebo budoucími hodnotami daných veličin. Přesná definice podle V. V. Němyckého zní takto:

Nechť S je separabilní metrický prostor a $T \subseteq \mathbb{R}$. Dynamický systém je pak množina transformací $\Phi : T \times T \times S \rightarrow S$, které splňují:

- a) Pro všechna $t_0, t_1, t_2 \in T, x \in S$ platí $\Phi(t_2, t_0, x) = \Phi(t_2, t_1, \Phi(t_1, t_0, x))$
- b) Pro všechna $t, t_0 \in T$ a $x \in S$ a pro všechny posloupnosti $\{t_n\}, \{x_n\}$, pro které platí $t_n \rightarrow t$ a $x_n \rightarrow x$ platí $\Phi(t_n, t_0, x_n) \rightarrow \Phi(t, t_0, x)$.

Množinu S nazveme stavový prostor a libovolný bod $x \in S$ pak nazýváme stavem systému. Obraz $\Phi(t, t_0, x)$ představuje stav systému v čase t , byl-li systém v čase t_0 ve stavu x .

Dynamické systémy se dělí podle několika kritérií:

Podle typu množiny T Diskrétní systémy pro spočetnou množinu T (většinou $T = \mathbb{Z}$) a spojité pro nespočetnou množinu T (většinou $T = \mathbb{R}$).

Podle typu stavové veličiny Deterministické systémy, jestliže S je množina reálných vektorů dimenze n a stochastické systémy, jestliže S je množina náhodných vektorů dimenze n .

Podle vazby systému na okolí Rozlišujeme systémy neřízené (uzavřené), jestliže systém nemá žádné spojení s okolím, tj. množina transformací má jediný prvek, a systémy řízené (otevřené), kdy transformace F závisí na stavu okolí $u \in \mathbb{R}^m$, takzvaný řídicí vektor.

Podle časové závislosti vlastností systému Máme časově neměnné systémy (časově invariantní nebo též autonomní systémy), které splňují $\Phi(t, t_0, x) = \Phi(t + \Delta t, t_0 + \Delta t, x)$, a časově proměnlivé systémy, jejichž vlastnosti se s časem mění.

Podle funkčních závislostí mezi veličinami Dělíme systémy na *lineární*, jestliže vnější popis systému je možný pomocí lineárních rovnic a *nelineární*, kde lineární popis je neadekvátní.

Aplikace dynamických systémů je možno nalézt v těchto oborech:

- elektrické systémy
- mechanické systémy
- hydraulické systémy
- tepelné systémy
- elektromechanické systémy
- biologické systémy
- ekonomické systémy

Každý z uvedených oborů volí k dynamickým systémům poněkud odlišný přístup vzhledem k odlišnosti příslušných vědních oborů. Snad nejvíce jsou propracovány systémy elektrické, výsledky však nejsou vždy přenositelné. V tomto textu se budeme zabývat výhradně systémy ekonomickými, které jsou v mnoha ohledech komplikovanější a abstraktnější než klasické systémy technické. Přesto jejich výzkum zaznamenal v posledních desetiletích bouřlivého rozvoje a jejich aplikace se stále častěji objevují jako nástroje ekonomického rozhodování. Tato skutečnost ukazuje dynamické systémy jako disciplínu, která je hodna studia.

Kapitola 2

SPOJITÉ DYNAMICKÉ SYSTÉMY

2.1 Teorie

V této kapitole se omezíme pouze na deterministické systémy, takže stavový prostor $S = \mathbb{R}^n$. Spojité dynamické systémy mohou být popsány mnoha různými způsoby:

- diferenciální rovnicí
- operátorovým přenosem
- frekvenčním přenosem
- frekvenční charakteristikou
- impulsní odezvou
- přechodovou charakteristikou
- rozložením nul a pólů přenosu

Všechny uvedené charakteristiky jsou vzájemně převoditelné (viz. např. [?]). Věnujme se nyní reprezentaci pomocí diferenciální rovnice. Dynamický systém je obecně popsán systémem obyčejných diferenciálních rovnic m -tého řádu ve tvaru

$$x^{(m)} = f(x^{(m-1)}, \dots, \dot{x}, x) \quad (2.1)$$

kde $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ je hladká funkce.

Při splnění tzv. Cauchyho počáteční podmínky

$$\begin{aligned} x(0) &= \xi_0 \\ \dot{x}(0) &= \xi_1 \\ &\dots \\ x^{(m-1)}(0) &= \xi_{m-1} \end{aligned} \quad (2.2)$$

je řešením systému jednoznačně určena funkce $x : \mathbb{R} \rightarrow \mathbb{R}^n$.

Veta 2.1. Existuje soustava ODR 1. řádu taková, že systém m -tého řádu (2.1) je s touto soustavou ekvivalentní, tj. existuje bijekce mezi množinami všech řešení obou soustav.

Důkaz. Položme $y = (x, \dot{x}, \dots, x^{(m-1)}) = (y_0, y_1, \dots, y_{m-1})$. Zřejmě platí $\dot{y} = (\dot{x}, \ddot{x}, \dots, x^{(m)})$. Přitom

$$\begin{aligned} \dot{x} &= y_1 \\ \ddot{x} &= y_2 \\ &\dots \\ x^{(m-1)} &= y_{m-1} \\ x^{(m)} &= f(y_0, y_1, \dots, y_{m-1}) \end{aligned} \tag{2.3}$$

Počáteční podmínky můžeme přepsat jako

$$\begin{aligned} y_0(0) &= \xi_0 \\ y_1(0) &= \xi_1 \\ &\dots \\ y_{m-1}(0) &= \xi_{m-1} \end{aligned} \tag{2.4}$$

což můžeme zkráceně napsat jako

$$\begin{aligned} \dot{y} &= F(y) \\ y(0) &= \xi \end{aligned} \tag{2.5}$$

což je systém diferenciálních rovnic prvního řádu s Cauchyho počáteční podmínkou. Vztah mezi x a y je bijekce, takže oba systémy jsou ekvivalentní. \square

Nadále se tedy budeme zabývat systémy prvního řádu

$$\dot{x} = f(x) \tag{2.6}$$

kde $\dot{x} = \left(\frac{dx_1}{dt}, \dots, \frac{dx_n}{dt} \right)^T$ a f je funkce $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$.

Příklad 2.2. Uvažujme skalární systém druhého řádu ve tvaru

$$\ddot{y} = ay + by + c$$

s Cauchyho počáteční podmínkou

$$\begin{aligned} y(0) &= y_0 \\ \dot{y}(0) &= y_1 \end{aligned}$$

Položme $x = (x_1, x_2) = (y, \dot{y})$. Potom platí:

$$\dot{x}_1 = x_2 \tag{2.7a}$$

$$\dot{x}_2 = ax_2 + bx_1 + c \tag{2.7b}$$

s počáteční podmínkou

$$\begin{aligned} x_1(0) &= y_0 \\ x_2(0) &= y_1. \end{aligned}$$

Rovnice (2.7) definují spojitý dynamický systém prvního řádu s Cauchyho počáteční podmínkou.

System diferenciálních rovnic (2.6) má i svou fyzikální interpretaci. Vektorová funkce $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definuje na \mathbb{R}^n rychlostní pole. Vyskytne-li se v bodě x částice o jednotkové hmotnosti, udělí se jí rychlost $f(x)$. Trajektorie částice je pak již jednoznačně určena funkcí f .

Byl-li systém v čase t_0 ve stavu x_0 , můžeme definovat zobrazení Φ_t stavového prostoru do sebe takové, že ke každému stavu $x(0)$ a času t přiřadí stav $x(t)$. Toto zobrazení se nazývá *operátor toku* nebo také *evoluční operátor*. Toto zobrazení je totožné s transformací Φ podle Nemyckého a platí

$$\Phi_t(x) = \Phi(t, t_0, x)$$

Operátor toku je také řešením soustavy (2.6), neboť v každém bodě t je soustava (2.6) splněna a zároveň je možno na základě operátoru toku stanovit ke každému stavu x jeho budoucnost i minulost. Řešení soustavy (2.6) vyhovující počáteční podmínce $x(t_0) = x_0$ lze vyjádřit pomocí evolučního operátoru takto:

$$x(t) = \Phi_{t-t_0}(x_0)$$

Takovou křivku $x(t)$ nazýváme *trajektorie* systému.

Z definice evolučního operátoru vyplývají následující vztahy:

$$\Phi_{s+t}(x) = \Phi_s(\Phi_t(x)) \quad (2.8)$$

$$\Phi_t(\Phi_{-t}(x)) = \Phi_0(x) = x \quad (2.9)$$

$$\Phi_t^{-1} = \Phi_{-t} \quad (2.10)$$

Definice 2.3. Existuje-li $x \in S$ takové, že $\Phi_t(x) = x$ pro všechna t , pak bod x nazveme *kritickým bodem* (pevným bodem) systému. Kritické body hrají při vyšetřování dynamických systémů významnou roli.

Definice 2.4. Soustava parametrických křivek $\Phi_t(x)$ pro každé $x \in S$ tvoří tzv. *fázový portrét* systému.

Zkoumáním velkého množství dynamických systémů se zjistilo, že tvary trajektorií systémů mohou nabývat množství různých podob. Některé z nich konvergují, monotonně nebo periodicky, k nějaké limitě, což může být bod a nebo množina bodů (kružnice). Mluvíme o *stabilním systému*. Některé další trajektorie se po vychýlení z rovnováhy pohybují směrem ven, aby se do ní opět časem vrátily (homoklinický orbit) nebo divergují do nevlastního bodu (pak mluvíme o *nestabilním systému*). Některé trajektorie jsou nejprve přitahovány do pevného bodu aby se, poté co dosáhnou jeho bezprostředního okolí, odchýlily do nekonečna (sedlový pevný bod). Některé trajektorie mají periodický charakter (těm pak říkáme *orbity*). Některé systémy jsou lokálně stabilní, tj. konvergují k limitě pouze v okolí pevného bodu, jiné jsou stabilní globálně, tj. konvergují z kteréhokoliv bodu. Existují dokonce trajektorie, které se v některém bodě rozdělují do několika větví. Takovému jevu říkáme *bifurkace*. Jsou známy systémy, které reagují nespojitým katastrofickým skokem na infinitesimální změnu některého parametru. Takové a ještě mnohé neprozkoumané podoby mohou mít dynamické systémy.

2.2 Ekonomické aplikace

2.2.1 Walrasův model tržní rovnováhy

Walras ukazuje rovnováhu trhu jako výsledek přizpůsobovacího procesu typu aukce: poté, co licitátor vyhlásí cenu p , kupující vyhlásí poptávku $D(p)$ a prodávající vyhlásí nabídku $S(p)$.

Rozhodující pro další postup je přebytek poptávky $E(p) = D(p) - S(p)$. Je-li $E(p) > 0$, proběhne nové kolo při vyšší ceně. Je-li $E(p) < 0$, proběhne nové kolo při nižší ceně. Po dosažení rovnováhy $E(p) = 0$ se provede směna. Přizpůsobovací proces můžeme modelovat rovnicí

$$\dot{p} = kE(p)$$

kde k je kladná konstanta, odrážející rychlost přizpůsobování. V našem příkladě uvažujeme lineární funkce poptávky $D(p) = \alpha + \beta p$ a nabídky $S(p) = \gamma + \delta p$. Počátek přizpůsobovacího procesu je $p(0) = p_0$.

Upravíme stavovou rovnici:

$$\dot{p} = k(\alpha + \beta p - \gamma - \delta p) = k(\beta - \delta)p + k(\alpha - \gamma)$$

Stavová rovnice je lineární diferenciální rovnice prvního řádu s konstantními koeficienty. Řešení hledáme ve tvaru

$$p(t) = a + be^{ct}$$

Dosazením do stavové rovnice a počáteční podmínky vychází

$$a = \frac{\gamma - \alpha}{\beta - \delta} \quad (2.11)$$

$$b = p_0 - \frac{\gamma - \alpha}{\beta - \delta} \quad (2.12)$$

$$c = k(\beta - \delta) \quad (2.13)$$

takže řešením je křivka

$$p(t) = \frac{\gamma - \alpha}{\beta - \delta} + \left(p_0 - \frac{\gamma - \alpha}{\beta - \delta} \right) e^{k(\beta - \delta)t}$$

Systém je stabilní při splnění podmínky $\beta - \delta < 0$. Ale β je sklon křivky nabídky a δ je sklon křivky poptávky. Je-li nabídka rostoucí a poptávka klesající, je podmínka splněna a trh je stabilní a konverguje k rovnovážné ceně

$$p(\infty) = p_e = \frac{\gamma - \alpha}{\beta - \delta}$$

při které je excés poptávky nulový, tzn. $D(p(\infty)) = S(p(\infty))$.

2.2.2 Model IS-LM

Uvažujme jednoduchou uzavřenou ekonomiku, která se chová podle následujících pravidel:

- (i) Produkt Y roste s převisem agregátní poptávky D , tj. $\dot{Y} = h(D - S)$
- (ii) Úroková míra r roste s převisem poptávky po penězích L , tj. $\dot{r} = m(L - \bar{M})$, kde \bar{M} je nabídka peněz určená centrální bankou.
- (iii) Poptávka po penězích je proporcionalní produktu, tj. $L = kY$. Uvažuje se pouze transakční poptávka v duchu Fisherovy kvantitativní teorie peněz.
- (iv) Spotřeba C je proporcionalní produktu, tj. $C = cY$.

- (v) Investice I klesají s rostoucím úrokem, tj. $I = -ar$.
- (vi) Agregátní poptávka je součtem spotřeby a investic, tj. $D = C + I$.
- (vii) Agregátní nabídka S je národní produkt, tj. $S = Y$.
- (viii) $0 < h, m, a, 0 < c < 1$.

Položíme-li mezní sklon k úsporám $s = 1 - c$, můžeme psát

$$\begin{aligned}\dot{Y} &= h(D - Y) = h(C + I - Y) = h(c - 1)Y -ahr = -hsY -ahr \\ \dot{r} &= mkY - m\bar{M}\end{aligned}$$

Derivujeme-li první rovnici a do vzniklého výrazu dosadíme rovnici druhou, dostaneme

$$\ddot{Y} = -hs\dot{Y} -ahr = -hs\dot{Y} -ahmkY +ahm\bar{M}$$

neboli

$$\ddot{Y} +hs\dot{Y} +ahmkY =ahm\bar{M}$$

což je lineární ODR druhého řádu. Napišme její charakteristickou rovnici

$$\lambda^2 +hs\lambda +ahmk = 0$$

Zde se řešení rozpadá na tři části

a) $hs^2 > 4amk$

V tomto případě existují 2 reálná řešení charakteristické rovnice

$$\begin{aligned}\lambda_1 &= \frac{1}{2} \left(-hs + \sqrt{h^2s^2 - 4ahmk} \right) \\ \lambda_2 &= \frac{1}{2} \left(-hs - \sqrt{h^2s^2 - 4ahmk} \right)\end{aligned}$$

kterým odpovídá řešení DR ve tvaru

$$Y(t) = \frac{\bar{M}}{k} + C_1e^{\lambda_1 t} + C_2e^{\lambda_2 t}$$

Toto řešení bude stabilní, jestliže bude platit $\lambda_1 < 0, \lambda_2 < 0$. Ukážeme, že tato podmínka je splněna

Vzhledem k podmínce (viii) platí $ahmk > 0$. Pak také

$$\begin{aligned}-4ahmk < 0 &\Leftrightarrow h^2s^2 - 4ahmk < h^2s^2 \Leftrightarrow \sqrt{h^2s^2 - 4ahmk} < |hs| \Leftrightarrow \\ \sqrt{h^2s^2 - 4ahmk} < hs &\Leftrightarrow \frac{1}{2} \left(-hs + \sqrt{h^2s^2 - 4ahmk} \right) < 0 \Leftrightarrow \lambda_1 < 0\end{aligned}$$

a podobně

$$-\sqrt{h^2s^2 - 4ahmk} < hs \Leftrightarrow \frac{1}{2} \left(-hs - \sqrt{h^2s^2 - 4ahmk} \right) < 0 \Leftrightarrow \lambda_2 < 0$$

Z čehož vyplývá, že systém je stabilní a je přitahován k pevnému bodu \bar{M}/k .

b) $hs^2 = 4amk$

Za této situace má charakteristická rovnice dvojnásobný reálný kořen

$$\lambda_1 = \lambda_2 = \lambda = -\frac{1}{2}hs < 0$$

kterému odpovídá řešení

$$Y(t) = \frac{\bar{M}}{k} + C_1 e^{\lambda t} + C_2 t e^{\lambda t}$$

Toto řešení je stabilní a konverguje k pevnému bodu \bar{M}/k .

c) $hs^2 < 4amk$

V tomto případě existují 2 komplexní kořeny charakteristické rovnice

$$\begin{aligned}\lambda_1 &= \frac{1}{2} \left(-hs + i\sqrt{4ahmk - h^2s^2} \right) \\ \lambda_2 &= \frac{1}{2} \left(-hs - i\sqrt{4ahmk - h^2s^2} \right)\end{aligned}$$

Položíme-li $\alpha = -\frac{1}{2}hs$ a $\beta = \frac{1}{2}\sqrt{4ahmk - h^2s^2}$, pak můžeme řešení napsat ve tvaru

$$Y(t) = \frac{\bar{M}}{k} + e^{\alpha t} (C_1 \cos \beta t + C_2 \sin \beta t) \quad (2.14)$$

Protože $\alpha < 0$, je řešení stabilní a tlumenými sinusovými kmity se přibližuje k pevnému bodu \bar{M}/k . Při vysokém sklonu k úsporám s je toto přibližování rychlejší, stejně tak jako při vysoké hodnotě koeficientu h , který zachycuje schopnost ekonomiky reagovat na zvýšenou poptávku růstem produktu. Periodické výkyvy produktu jsou obrazem hospodářských cyklů s periodou $4\pi/\sqrt{4ahmk - h^2s^2}$.

Reálná existence hospodářských cyklů v ekonomickém životě napovídá, že nejčastěji se vyskytuje případ c). Trajektorie skutečné ekonomiky však nikdy neodpovídá ideální křivce (2.14) vzhledem ke značně zjednodušeným předpokladům modelu, zejména izolaci od okolí a konstantní nabídce peněz.

2.2.3 Neoklasický model růstu

Tento model, vypracovaný Swanem a Solowem (1956) je založen na následujících předpokladech:

- (i) Pracovní síla L roste konstantním tempem n , tj. $\dot{L}/L = n$.
- (ii) Úspory jsou lineárně závislé na produktu, tj. $S = sY$.
- (iii) Všechny úspory jsou investovány do kapitálu K .
- (iv) Investice $I = \dot{K} + \delta K$.
- (v) Produkce probíhá za podmínky konstantních výnosů z rozsahu, tj. $Y = F(K, L) = LF(K/L, 1) \equiv Lf(k)$, kde $k = K/L$ je vybavenost práce kapitálem.
- (vi) $s, \delta \in (0, 1)$.

Spočtěme \dot{k} :

$$\dot{k} = \frac{\dot{K}L - K\dot{L}}{L^2} = \frac{\dot{K}}{L} - \frac{K\dot{L}}{L^2}$$

z toho vyplývá, že

$$\frac{\dot{k}}{k} = \frac{L}{K}\dot{k} = \frac{\dot{K}}{K} - \frac{\dot{L}}{L}$$

dosazením podmínky (i) dostáváme

$$\frac{\dot{k}}{k} = \frac{\dot{K}}{K} - n$$

z podmínek (ii) a (iii) vyplývá, že

$$\dot{K} = sY - \delta K$$

dosazením dostáváme

$$\frac{\dot{k}}{k} = s\frac{Y}{K} - (\delta + n) = s\frac{L}{K}f(k) - (\delta + n) = \frac{s}{k}f(k) - (\delta + n)$$

Vynásobením k máme

$$\dot{k} = sf(k) - (\delta + n)k = sf(k) - \lambda k$$

kde f je rostoucí konkávní funkce, splňující Janadovy podmínky $\lim_{k \rightarrow 0} f'(k) = \infty$ a $\lim_{k \rightarrow \infty} f'(k) = 0$. V případě Cobb-Douglasovy produkční funkce

$$Y = K^\alpha L^{1-\alpha}$$

dostáváme

$$\frac{Y}{L} = K^\alpha L^{-\alpha} = k^\alpha = f(k)$$

a diferenciální rovnice růstu je ve tvaru

$$\dot{k} = -\lambda k + sk^\alpha$$

což je Bernoulliho rovnice. Provedeme substitucí $x = k^{1-\alpha}$.

Pak

$$\begin{aligned} \dot{x} &= (1-\alpha)k^{-\alpha}\dot{k} \implies \dot{k} = \dot{x} \frac{k^\alpha}{1-\alpha} \\ k &= k^{1-\alpha}k^\alpha = xk^\alpha \end{aligned}$$

Dosazením do DR máme

$$\dot{x} \frac{k^\alpha}{1-\alpha} = -\lambda xk^\alpha + sk^\alpha$$

odkud dostáváme nakonec

$$\dot{x} = -(1-\alpha)\lambda x + (1-\alpha)s$$

což je rovnice lineární. Jejím řešením je trajektorie

$$x(t) = \left(x_0 - \frac{s}{\lambda}\right) e^{-(1-\alpha)\lambda t} + \frac{s}{\lambda}$$

nebo ekvivalentně trajektorie

$$k^{1-\alpha} = \left(k_0^{1-\alpha} - \frac{s}{\lambda}\right) e^{-(1-\alpha)\lambda t} + \frac{s}{\lambda}$$

protože platí $1-\alpha > 0$, $\lambda = \delta + n > 0$, je model stabilní s pevným bodem

$$k^* = \left(\frac{s}{\delta + n}\right)^{\frac{1}{1-\alpha}}$$

2.2.4 Lotkuv-Volterrův model lovce a oběti

Tento model má svůj původ v letech po 1. světové válce. Lotka v roce 1925 a Volterra v letech 1931 a 1937 pozorovali výskyty dravých ryb (predators) a jejich kořistí (preys) v Jaderském Moři. Na základě tohoto pozorování zformulovali model, který vysvětluje fluktuace v počtech rybích populací. Tento model je vynikající ilustrací nelineárního dynamického systému v rovině.

Populace kořistí $x(t)$ roste přirozeným tempem a (tj. $\dot{x} = ax$). Jejich počet je snižován o cxy díky přítomnosti dravců $y(t)$, kteří loví v hejnech kořistí.

Současně s tím přirozená populace dravců klesá tempem b (tj. $\dot{y} = -by$), ale za přítomnosti obětí se tento pokles snižuje o dxy . Za předpokladu $a, b, c, d > 0$ pak máme systém

$$\dot{x} = ax - cxy \quad (2.15)$$

$$\dot{y} = dxy - by \quad (2.16)$$

Položíme-li $\dot{x} = 0, \dot{y} = 0$, obdržíme dva kritické body systému. $z = (0, 0)^T$ je tzv. rovnováha po vyhnutí (extinction equilibrium) a $z = (\frac{b}{d}, \frac{a}{c})^T$ je tzv. rovnováha koexistence (coexistence equilibrium). Numerické řešení DR (2.15) je v Matlabu ve funkci `odedemo`.

Kapitola 3

DISKRÉTNÍ DYNAMICKÉ SYSTÉMY

3.1 Teorie

Jako diskretní označujeme ty systémy, jejichž chování je definováno na diskretní časové množině $\mathcal{T} = \{kT : k \in \mathbb{Z}\}$. Proměnná k se nazývá *krok*. Délka kroku T bývá též označována jako *vzorkovací interval* nebo *perioda vzorkování*. Ze spojité funkce $f(t)$ lze získat diskretní funkci $g(k)$ vzorkováním. Přitom platí $f(t) = f(kT) = g(k)$.

Pro vnější popis diskretních systémů se běžně používá diferenční rovnice a operátorový přenos. Diferenční rovnice, popisující neřízený deterministický systém diskretního typu má podobu

$$y(k) = f(y(k-1), \dots, y(k-l_y)) \quad (3.1)$$

kde $y \in \mathbb{R}^p$ je výstup systému. Tato rovnice ukazuje závislost výstupu $y(k)$ na minulých výstupech $y(k-i)$. Řád systému l_y pak udává paměť systému.

Přidáme-li do systému prvek řízení, tj. řídicí vektor $u \in \mathbb{R}^m$, dostaneme rovnici

$$y(k) = f(y(k-1), \dots, y(k-l_y), u(k-1), \dots, u(k-l_u)) \quad (3.2)$$

Takto napsaná rovnice naznačuje, že řízení $u(k-1)$ přímo ovlivňuje výstup $y(k)$. Účinek řízení je tedy zpožděn o jeden krok. Takto se chová převážná většina reálných systémů. Někdy je dokonce zpoždění větší. Pak se v rovnici (3.2) objevují členy až od $u(k-n_u)$.

Sledujeme-li pouze závislost výstupu systému na minulých výstupech a vstupech, říkáme tomu *vnější popis*. Zavedeme-li stav systému jako veličinu, která plně určuje vývoj systému do budoucna, pak se systém dá napsat jako systém prvního řádu. Mluvíme pak o *vnitřním* nebo také *stavovém* popisu systému. Přesněji stav $x(k)$ a řízení $u(k)$ určují jednoznačně stavy $x(k+1), x(k+2), \dots$ a výstupy $y(k), y(k+1), y(k+2), \dots$

Definice 3.1. Dynamický systém popsáný rovnicemi

$$x(t+1) = f(x(t), u(t)) \quad (3.3a)$$

$$y(t) = g(x(t), u(t)) \quad (3.3b)$$

nazveme deterministickým dynamickým systémem ve stavovém tvaru

Rovnici (3.3a) nazýváme *stavovou rovnicí* (rovnice přechodu, state equation, plant equation, transition equation) a rovnice (3.3b) je zvána *rovnicí výstupu* (rovnice měření, observation equation, output equation, measurement equation).

Systémy ve stavovém tvaru hrají v teorii dynamických systémů významnou roli, protože právě pro stavové systémy jsou vypracovány algoritmy optimální estimace, identifikace a optimálního řízení. Proto je vhodné zabývat se otázkou existence a jednoznačnosti stavového popisu systému, který je zadán svým vnějším popisem (3.2). Stavový popis vždy existuje, není však obecně jednoznačně určen. Jeden z možných převodů popisuje následující věta.

Věta 3.2. Položíme-li

$$x(k) = (y(k)^T, \dots, y(k - l_y + 1)^T, u(k - 1)^T, \dots, u(k - l_u + 1)^T)^T$$

pak systém (3.2) lze napsat jako systém ve stavovém tvaru.

Důkaz. Původní rovnici

$$y(k) = f(y(k - 1), \dots, y(k - l_y), u(k - 1), \dots, u(k - l_u))$$

lze substitucí $t + 1 = k$ napsat ve tvaru

$$y(t + 1) = f(y(t), \dots, y(t - l_y + 1), u(t - 1), \dots, u(t - l_u + 1)) = f(x(t))$$

Zřejmě platí

$$x(t + 1) = (y(t + 1)^T, \dots, y(t - l_y + 2)^T, u(t)^T, \dots, u(t - l_u + 2)^T)^T$$

takže první složka vektoru $x(t + 1)$ už svou rovnicí má. Najdeme stavové rovnice pro všechny zbývající složky. V těchto rovnicích mohou na pravých stranách vystupovat pouze složky vektorů $x(t)$ a $u(t)$.

$$\begin{aligned} y(t + 1) &= f(x(t)) \\ y(t) &= A_1 x(t) \\ &\vdots \\ y(t - l_y + 2) &= A_{l_y} x(t) \\ u(t) &= u(t) \\ u(t - l) &= B_1 x(t) \\ &\vdots \\ u(t - l_u + 1) &= B_{l_u - 1} x(t) \end{aligned}$$

kde $A_1, \dots, A_{l_y}, B_1, \dots, B_{l_u - 1}$ jsou permutační matice, vybírající z vektoru $x(t)$ vždy jen příslušnou část. Celkem lze tedy soustavu napsat jako

$$x(t + 1) = F(x(t), u(t))$$

což je popis systému ve stavovém tvaru □

Příklad 3.3. Uvažujme lineární dynamický systém ve tvaru

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + b_1 u_{t-1} + b_2 u_{t-2}$$

Posuneme čas o jeden krok dopředu a máme

$$y_{t+1} = a_1 y_t + a_2 y_{t-1} + b_1 u_t + b_2 u_{t-1}$$

Nyní položíme $x_t = (x_{1t}, x_{2t}, x_{3t})^T = (y_t, y_{t-1}, u_{t-1})^T$. Potom platí

$$x_{t+1} = (y_{t+1}, y_t, u_t)^T$$

a můžeme psát

$$x_{1,t+1} = a_1 x_{1,t} + a_2 x_{2,t} + b_2 x_{3,t} + b_1 u_t$$

$$x_{2,t+1} = x_{1,t}$$

$$x_{3,t+1} = u_t$$

$$y_t = x_{1,t}$$

což je systém ve stavovém tvaru (3.3).

Definice 3.4. Dynamický systém popsáný rovnicemi

$$x(t+1) = f(x(t), u(t), v(t)) \quad (3.4a)$$

$$y(t) = g(x(t), u(t), w(t)) \quad (3.4b)$$

kde $v(t) \in \mathcal{L}n_v$, $w(t) \in \mathcal{L}n_w$ jsou náhodné vektory, splňující podmínku $E v(t) = 0$, $E w(t) = 0$, nazveme stochastickým dynamickým systémem ve stavovém tvaru.

Náhodný vektor $v(t)$ nazýváme *šum procesu* (process noise) a náhodný vektor $w(t)$ nazýváme *šum výstupu* (output noise). V naprosté většině případů mají rovnice (3.4a) a (3.4b) speciální tvar

$$x(t+1) = f(x(t), u(t)) + v(t)$$

$$y(t) = g(x(t), u(t)) + w(t)$$

odtud je motivován požadavek na centrovanost obou šumů. Přechod od deterministického ke stochastickému systému je výrazným kvalitativním skokem, poněvadž původní stavový prostor reálných vektorů \mathbb{R}^{n_x} přešel ve stavový prostor $\mathcal{L}n_x$ náhodných vektorů. Stejně tak výstupy $y(t)$ jsou nyní náhodnými vektory. Odtud plyne, že ke zkoumání takových systémů je nutno použít teorii pravděpodobnosti.

Definice 3.5. Necht' $A \in \mathbb{R}^{n_x \times n_x}$, $B \in \mathbb{R}^{n_x \times n_u}$, $C \in \mathbb{R}^{n_y \times n_x}$, $D \in \mathbb{R}^{n_y \times n_u}$. Soustava

$$x(t+1) = Ax(t) + Bu(t) + v(t) \quad (3.5)$$

$$y(t) = Cx(t) + Du(t) + w(t) \quad (3.6)$$

popisuje *lineární* stochastický dynamický systém ve stavovém tvaru. Šumy $v(t)$, $w(t)$ musí splňovat následující podmínky

- (i) Šumy v_t a w_t jsou normální s nulovou střední hodnotou.
- (ii) $Ev_t w_t^T = 0$
- (iii) $Ev_t v_t^T = Q$
- (iv) $EW_t w_t^T = R$

Zabývejme se nyní stabilitou lineárního stochastického dynamického systému. Obecně je dynamický systém stabilní tehdy, když je pevný bod systému atraktorem neboli přitahujícím pevným bodem. Z libovolného počátečního stavu potom stav systému konverguje k tomuto pevnému bodu. Stochastický systém je stabilní právě tehdy, když k němu příslušný deterministický systém je stabilní. Pracujeme-li navíc se systémem řízeným, pak o konvergenci stavů je možno hovořit pouze tehdy, je-li vstup systému konstantní. Ve svém důsledku pak vlastně zkoumáme stabilitu deterministického neřízeného systému.

Pevný bod lineárního systému s konstantním vstupem u určíme z rovnice $x = Ax + Bu$ odkud dostáváme

$$x_e = (I - A)^{-1}Bu \quad (3.7)$$

Věta 3.6. Stochastický systém (3.5) je stabilní právě tehdy, když všechna vlastní čísla matice A leží v jednotkovém kruhu.

Důkaz. Systém (3.5) je stabilní, právě když je stabilní systém

$$x(t+1) = Ax(t) + Bu \quad (3.8)$$

kde u je známý konstantní vstup systému. Pevný bod tohoto systému je (3.7). Zavedeme-li substituci $y(t) = x(t) - x_e$, pak stavová rovnice přejde v

$$\begin{aligned} x(t+1) - x_e &= Ax(t) + Bu - x_e \Leftrightarrow \\ x(t+1) - x_e &= Ax(t) - Ax_e + Bu + Ax_e - x_e \Leftrightarrow \\ x(t+1) - x_e &= A(x(t) - x_e) + Bu - (I - A)x_e \Leftrightarrow \\ y(t+1) &= Ay(t) + Bu - (I - A)(I - A)^{-1}Bu \Leftrightarrow \\ y(t+1) &= Ay(t) \end{aligned}$$

Má-li $x(t)$ konvergovat k x_e , pak musí $y(t)$ konvergovat k nule. Vyjdeme-li z počáteční podmínky $y(0) = y_0$, pak stav $y(t)$ se dá explicitně spočítat jako

$$y(t) = A^t y_0$$

Je vidět, že má-li $y(n)$ pro libovolné y_0 konvergovat k nule, musí A^n konvergovat k nulové matici.

Pro každou čtvercovou matici A existuje regulární matice P taková, že

$$P^{-1}AP = D$$

kde D je diagonální matice, která má v diagonále vlastní čísla matice A

$$D = \text{diag}(\lambda_1, \dots, \lambda_{n_x})$$

Sloupce matice P pak jsou vlastní vektory matice A . Tato dekompozice je známá jako *Schurova dekompozice* matice A . Matici A pak můžeme napsat jako

$$A = PDP^{-1},$$

matici A^2 zapíšeme jako

$$A^2 = PDP^{-1}PDP^{-1} = PD^2P^{-1}$$

a tak dále až matici A^n zapíšeme jako

$$A^n = PD^nP^{-1}$$

Matice D^n je opět diagonální a obsahuje prvky

$$D^n = \text{diag}(\lambda_1^n, \dots, \lambda_{n_x}^n)$$

Posloupnost A^n konverguje k nule právě tehdy, když posloupnost D^n konverguje k nule a to je právě tehdy, když všechny posloupnosti λ_i^n konvergují k nule. To je splněno, když pro všechna $i \in \{1, \dots, n_x\}$ platí

$$|\lambda_i| < 1 \quad (3.9)$$

Protože vlastní čísla λ_i jsou obecně komplexní, neznamená (3.9) nic jiného, než že všechny vlastní čísla se v komplexní rovině nachází uvnitř jednotkového kruhu. \square

Chování stabilního stochastického systému je významně ovlivněno vlastnostmi stochastické složky. Bude-li šum slabý, bude systém po dostatečně dlouhém čase skutečně vykazovat stabilitu a bude jen nepatrně kolísat kolem pevného bodu. Naopak stabilní systém se silným šumem se může chovat naprosto chaoticky. Čím budou vlastní čísla blíže k nule, tím rychleji systém konverguje k pevnému bodu. Při hodnotách blízkých jedné už se konvergence značně zpomaluje a po překročení jednotkového kruhu se z přitahujícího pevného bodu stane odpuzující pevný bod a systém se stane nestabilním.

3.2 Ekonomické aplikace

3.2.1 Samuelsonův hospodářský cyklus (1939)

Samuelsonův model je postaven na multiplikátoru a akcelátoru. Struktura modelu je popsána těmito rovnicemi:

$$\begin{aligned} C_t &= Y_{t-1} \\ I_t &= v(C_t - C_{t-1}) \\ G_t &= 1 \\ Y_t &= C_t + I_t + G_t \end{aligned}$$

Předpokládá se $0 < c < 1, v > 0$. $1/1 - c$ je multiplikátor a v je akcelátor. Substitucí dostaneme typickou diferenční rovnici 2. řádu

$$Y_t - c(1 + v)Y_{t-1} + cvY_{t-2} = 1$$

Jednoduchým cvičením se dá dokázat, že platí:

- (a) Každý skalární násobek řešení homogenní rovnice je řešením homogenní rovnice.
 (b) Součet každých dvou řešení homogenní rovnice je řešením homogenní rovnice.
 (c) Součet libovolného řešení nehomogenní rovnice s libovolným řešením homogenní rovnice je řešením nehomogenní rovnice.

Z uvedeného vyplývá, že množina všech řešení homogenní rovnice tvoří vektorový podprostor a množina všech řešení nehomogenní rovnice tvoří afinní podprostor v prostoru funkcí. Stačí tedy najít jedno řešení y_0 nehomogenní rovnice (tzv. *partikulární řešení*) a bázi y_1, \dots, y_n všech řešení homogenní rovnice. Libovolné řešení nehomogenní rovnice pak dostaneme jako

$$y = y_0 + a_1 y_1 + \dots + a_n y_n$$

Lehce se ukáže, že funkce

$$y_0(t) = \frac{1}{1-c}$$

je partikulárním řešením naší diferenční rovnice. Hledejme nyní řešení homogenní rovnice ve tvaru $Y_t = \lambda^t$. Pak musí platit

$$\lambda^t - c(1+v)\lambda^{t-1} + cv\lambda^{t-2} = 0$$

Vydělením rovnice výrazem λ^{t-2} dostaneme

$$\lambda^2 - c(1+v)\lambda + cv = 0$$

Z této rovnice dostaneme λ jako kořen kvadratické rovnice. Řešení závisí na znaménku diskriminantu

$$\Delta = c^2(1+v)^2 - 4cv$$

a) $c > 4v/(1+v)^2$

V tomto případě existují 2 reálná řešení, která tvoří bázi řešení homogenní rovnice.

$$\begin{aligned} \lambda_1 &= \frac{1}{2} \left(c(1+v) + \sqrt{c^2(1+v)^2 - 4cv} \right) \\ \lambda_2 &= \frac{1}{2} \left(c(1+v) - \sqrt{c^2(1+v)^2 - 4cv} \right) \end{aligned}$$

kterým odpovídá obecné řešení ve tvaru

$$Y_t = \frac{1}{1-c} + a_1 \lambda_1^t + a_2 \lambda_2^t$$

Toto řešení bude stabilní, jestliže bude platit $|\lambda_1| < 1, |\lambda_2| < 1$.

$$\begin{aligned} |\lambda_1| < 1 \wedge |\lambda_2| < 1 &\Leftrightarrow |\lambda_1 \lambda_2| < 1 \Leftrightarrow \\ &\Leftrightarrow \frac{1}{4} \left| \left(c(1+v) + \sqrt{c^2(1+v)^2 - 4cv} \right) \left(c(1+v) - \sqrt{c^2(1+v)^2 - 4cv} \right) \right| < 1 \Leftrightarrow \\ &\Leftrightarrow \left| \frac{1}{4} 4cv \right| < 1 \Leftrightarrow cv < 1 \end{aligned}$$

b) $c = 4v/(1+v)^2$

Za této situace má kvadratická rovnice dvojnásobný reálný kořen

$$\lambda_1 = \lambda_2 = \lambda = \frac{c(1+v)}{2}$$

kterému odpovídá řešení

$$Y_t = \frac{1}{1-c} + a_1 \lambda^t + a_2 t \lambda^t$$

Toto řešení je stabilní za podmínky $c(1+v) < 2 \Leftrightarrow v < 1$.

c) $c < 4v/(1+v)^2$

V tomto případě existují 2 komplexní kořeny charakteristické rovnice

$$\begin{aligned}\lambda_1 &= \frac{1}{2} \left(c(1+v) + i\sqrt{4cv - c^2(1+v)^2} \right) \\ \lambda_2 &= \frac{1}{2} \left(c(1+v) - i\sqrt{4cv - c^2(1+v)^2} \right)\end{aligned}$$

Báze řešení homogenní rovnice je tvořena funkcemi $u = \lambda_1^t$ a $v = \lambda_2^t$, což jsou komplexní funkce. Ukážeme, že lze najít reálné funkce, které tvoří bázi řešení. Zavedme goniometrické vyjádření komplexních kořenů

$$\begin{aligned}\lambda_1 &= r(\cos \theta + i \sin \theta) \\ \lambda_2 &= r(\cos \theta - i \sin \theta)\end{aligned}$$

Pak platí

$$\begin{aligned}a_1 \lambda_1^t + a_2 \lambda_2^t &= \\ &= a_1 r^t (\cos \theta t + i \sin \theta t) + a_2 r^t (\cos \theta t - i \sin \theta t) = \\ &= r^t ((a_1 + a_2) \cos \theta t + i(a_1 - a_2) \sin \theta t) = \\ &= r^t (b_1 \cos \theta t + b_2 \sin \theta t)\end{aligned}$$

přítom r je absolutní hodnota společná komplexním číslům λ_1, λ_2 , která je rovna

$$\sqrt{\lambda_1 \bar{\lambda}_1} = \sqrt{\lambda_1 \lambda_2} = \sqrt{cv}$$

Systém bude stabilní za podmínky $r < 1 \Leftrightarrow cv < 1$.

Samuelsonův ekonomický systém může tedy vykazovat 9 kvalitativně různých druhů chování, a to v závislosti na aktuální hodnotě dvou parametrů - v a c . Každému typu chování přináší v parametrickém prostoru (v, c) jedna množina. Těchto 9 množin pokrývá celý parametrický prostor, který jsme si hned v úvodu omezili na množinu $(0, \infty) \times (0, 1)$.

1. $c > 4v/(1+v)^2 \wedge cv < 1$: monotonní konvergence - stabilní
2. $c > 4v/(1+v)^2 \wedge cv = 1$: neutrální stabilita
3. $c > 4v/(1+v)^2 \wedge cv > 1$: monotonní divergence - nestabilní
4. $c = 4v/(1+v)^2 \wedge v < 1$: jeden kriticky tlumený kmit - stabilní

5. $c = 4v/(1+v)^2 \wedge v = 1$: lineární divergence - nestabilní
6. $c = 4v/(1+v)^2 \wedge v > 1$: monotonní divergence - nestabilní
7. $c < 4v/(1+v)^2 \wedge cv < 1$: periodická konvergence - stabilní
8. $c < 4v/(1+v)^2 \wedge cv = 1$: konstantní periodická oscilace - stabilní
9. $c < 4v/(1+v)^2 \wedge cv > 1$: periodická divergence - nestabilní

3.2.2 Hickův příspěvek k teorii obchodních cyklů (1950)

K ilustraci předešlé věty použijeme Hicksův model hospodářského cyklu z roku 1950. Tento model vysvětluje hospodářský cyklus pomocí interakce mezi multiplikátorem a akcelerátorem. Model má následující strukturu:

$$\begin{aligned} C_t &= (1-s)Y_{t-1} \\ I_t &= A_0(1+g)^t + v(Y_{t-1} - Y_{t-2}) \\ Y_t &= C_t + I_t \end{aligned}$$

Spotřeba C_t je zpožděná lineární funkce Y_{t-1} . Parametr s splňující podmínku $0 < s < 1$ je mezní sklon k úsporám. Investice I_t mají dvě komponenty: autonomní investice $A_0(1+g)^t$ rostoucí exponenciálně konstantním tempem g a indukované investice $v(Y_{t-1} - Y_{t-2})$, které jsou proporcionalní růstu GNP v minulém období. Konstanta $v > 0$ je akcelerátor. Substitucí dostaneme rovnici

$$Y_t = (1-s+v)Y_{t-1} + vY_{t-2} + A_0(1+g)^t$$

Autonomní investice můžeme chápat jako exogenní veličinu, kterou zde pro jednoduchost ztotožníme s řízením u_t , tj.

$$Y_t = (1-s+v)Y_{t-1} + vY_{t-2} + u_{t-1}$$

Posuneme čas o jeden krok dopředu a máme

$$Y_{t+1} = (1-s+v)Y_t + vY_{t-1} + u_t$$

Nyní položíme $x_t = (x_{1t}, x_{2t})^T = (Y_t, Y_{t-1})^T$. Potom platí

$$x_{t+1} = (Y_{t+1}, Y_t)^T$$

odtud pak

$$\begin{aligned} x_{1,t+1} &= (1-s+v)x_{1,t} + vx_{2,t} + u_t \\ x_{2,t+1} &= x_{1,t} \end{aligned} \tag{3.10}$$

a také

$$Y_t = x_{1,t} \tag{3.11}$$

rovnice (3.10) je stavovou rovnicí a rovnice (3.11) je výstupní rovnicí uvažovaného dynamického systému.

Kapitola 4

SIMULACE DYNAMICKÝCH SYSTÉMŮ

4.1 Teorie

Pojem simulace se úzce váže na použití výpočetní techniky jakožto nástroje pro její realizaci a výstup výsledků. Simulace dynamických systémů se významně liší podle toho, jestli se jedná o systém spojitý nebo diskrétní.

Definice 4.1. Buď dán spojitý systém ve stavovém tvaru (2.6) s počáteční podmínkou. Dále necht' $T = \{t_k : k \in \{1, \dots, N\}, N \in \mathbb{N}\}$ je tzv. *časový horizont*. Simulovat daný systém na horizontu T pak znamená nalézt vzorky trajektorie systému v časech t_k . Jinými slovy hledáme posloupnost $\{x_0, \dots, x_N\}$ takovou, že $x_k = x(t_k)$, kde funkce $x(t)$ je řešením soustavy DR (2.6).

Definice 4.2. Buď dán řízený stochastický diskrétní systém (3.3) s počáteční podmínkou. Dále necht' $T = \{1, \dots, N\}, N \in \mathbb{N}$ je *časový horizont*. Buď $\{u(0), \dots, u(N)\}$ známá posloupnost řídicích vektorů—*trajektorie řízení* a $\{v(0), \dots, v(N)\}, \{w(0), \dots, w(N)\}$ posloupnosti realizací náhodných šumů v, w . Simulovat systém (3.3) na horizontu T pak znamená nalézt trajektorii $\{x(0), \dots, x(N)\}$ stavů x a trajektorii $\{y(0), \dots, y(N)\}$ výstupů y tak, aby byly splněny rovnice (3.3).

Pro simulaci spojitých systémů je potřeba numericky vyřešit soustavu diferenciálních rovnic a řešení vyčíslit v bodech t_k . K tomu se využívá celá řada známých numerických metod, které lze najít např. v přehledu [?]. V systému Matlab jsou tyto procedury standartně k dispozici ve funkcích `ode23`, případně `ode45`.

Simulace diskrétních systémů je nesrovnatelně jednodušší. Vyřešit diferenční rovnici v podstatě znamená dosadit do její pravé strany. V případě stochastického systému je k tomu ještě potřeba vygenerovat realizaci příslušného šumu. Neurčitost stochastického systému tak způsobuje také neurčitost výsledku simulace, tj. trajektorií x a y .

4.2 Příklady

4.2.1 Pindyck-Rubinfeldův třírovnicový model

Tento lineární model použil Pindyck v učebnici [?] při testování problému optimálního řízení. Model obsahuje 4 výstupní veličiny a 2 vstupní (řídící) proměnné. Je zadán svým vnějším popisem

$$C_t = c_0 + c_1 Y_t + c_2 C_{t-1} \quad (4.1a)$$

$$I_t = i_0 + i_1 Y_t + i_2 (Y_{t-1} - Y_{t-2}) + i_3 R_{t-4} \quad (4.1b)$$

$$R_t = r_0 + r_1 Y_t + r_2 (Y_t - Y_{t-1}) + r_3 (M_t - M_{t-1}) + r_4 (R_{t-1} + R_{t-2}) \quad (4.1c)$$

$$Y_t = C_t + I_t + G_t \quad (4.1d)$$

Poslední rovnice je ekonomická identita, která se nepočítá jako modelová rovnice a proto mluvíme o třírovnicovém modelu.

Proměnné vystupující v modelu jsou:

| | | |
|-----|-------------------|-------------------|
| C | soukromá spotřeba | výstupní veličina |
| I | hrubé investice | výstupní veličina |
| R | úroková míra | výstupní veličina |
| Y | národní produkt | výstupní veličina |
| G | vládní spotřeba | řídící veličina |
| M | množství peněz | řídící veličina |

Parametry modelu (4.1) byly odhadnuty metodou nejmenších čtverců (OLS). Jejich hodnoty jsou:

$$\begin{array}{llll} c_0 = -9.4540 & c_1 = 0.05410 & c_2 = 0.9260 & \\ i_0 = -66.1950 & i_1 = 0.16840 & i_2 = 0.2181 & i_3 = -11.2650 \\ r_0 = -0.5561 & r_1 = 0.00051 & r_2 = 0.0135 & r_3 = -0.0853 \quad r_4 = 0.4259 \end{array}$$

Pro každou rovnici byly vypočteny reziduální rozptyly. Jejich hodnoty jsou:

$$\begin{array}{ll} s_1^2 = 11.230^2 & \text{pro rovnici (4.1a)} \\ s_2^2 = 23.864^2 & \text{pro rovnici (4.1b)} \\ s_3^2 = 0.854^2 & \text{pro rovnici (4.1c)} \end{array}$$

Model má dopravní zpoždění 0, což znamená, že vstup u_t přímo ovlivňuje výstup y_t . Lineární soustavu zapíšeme maticově

$$y_t = \alpha_0 y_t + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + \alpha_4 y_{t-4} + \beta_0 u_t \quad (4.2)$$

kde jsme položili

$$\begin{aligned}
 y_t &= \begin{bmatrix} C_t \\ I_t \\ R_t \\ Y_t \end{bmatrix} & u_t &= \begin{bmatrix} M_t - M_{t-1} \\ G_t \\ 1 \end{bmatrix} & \alpha_0 &= \begin{bmatrix} 0 & 0 & 0 & c_1 \\ 0 & 0 & 0 & i_1 \\ 0 & 0 & 0 & r_1 + r_2 \\ 1 & 1 & 0 & 0 \end{bmatrix} \\
 \alpha_1 &= \begin{bmatrix} c_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & i_1 \\ 0 & 0 & r_4 & -r_2 \\ 0 & 0 & 0 & 0 \end{bmatrix} & \alpha_2 &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i_1 \\ 0 & 0 & r_4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & \alpha_3 &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\
 \alpha_4 &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & i_3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & \beta_0 &= \begin{bmatrix} 0 & 0 & c_0 \\ 0 & 0 & i_0 \\ r_3 & 0 & r_0 \\ 0 & 1 & 0 \end{bmatrix}
 \end{aligned}$$

Poslední element vektoru u_t je jednička. Tato úprava formálně zbavuje soustavu úrovnových konstant c_0, i_0, r_0 .

Snadno nahlédneme, že model je konstruován dosud nezvyklým způsobem závislosti výstupních veličin. Současná hodnota y_t je závislá na současné hodnotě y_t . Takové dynamické systémy jsou v ekonomii celkem běžné a označují se jako *interdependentní* nebo také *implicitní* systémy. V případě lineárních systémů je lze převést na explicitní. Rovnici (4.2) upravíme následujícím způsobem:

$$\begin{aligned}
 y_t - \alpha_0 y_t &= \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + \alpha_4 y_{t-4} + \beta_0 u_t \iff \\
 (I - \alpha_0) y_t &= \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + \alpha_4 y_{t-4} + \beta_0 u_t \iff \\
 y_t &= (I - \alpha_0)^{-1} \alpha_1 y_{t-1} + (I - \alpha_0)^{-1} \alpha_2 y_{t-2} + (I - \alpha_0)^{-1} \alpha_3 y_{t-3} + \\
 &\quad + (I - \alpha_0)^{-1} \alpha_4 y_{t-4} + (I - \alpha_0)^{-1} \beta_0 u_t \iff \\
 y_t &= a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + a_4 y_{t-4} + b_0 u_t
 \end{aligned}$$

Tím jsme dostali model ARX, který zapíšeme blokově takto

$$\begin{bmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ y_{t-3} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ y_{t-4} \end{bmatrix} + \begin{bmatrix} b_0 \\ 0 \\ 0 \\ 0 \end{bmatrix} u_t \quad (4.3)$$

kde

$$\begin{aligned}
 a_1 = (I - \alpha_0)^{-1}\alpha_1 &= \begin{bmatrix} 0.9904 & 0 & 0 & 0.0152 \\ 0.2006 & 0 & 0 & 0.2653 \\ 0.0167 & 0 & 0.4259 & -0.0096 \\ 1.1910 & 0 & 0 & 0.2805 \end{bmatrix} \\
 a_2 = (I - \alpha_0)^{-1}\alpha_2 &= \begin{bmatrix} 0 & 0 & 0 & -0.0152 \\ 0 & 0 & 0 & -0.2653 \\ 0 & 0 & 0.4259 & -0.0039 \\ 0 & 0 & 0 & -0.2805 \end{bmatrix} \\
 a_3 = (I - \alpha_0)^{-1}\alpha_3 &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\
 a_4 = (I - \alpha_0)^{-1}\alpha_4 &= \begin{bmatrix} 0 & 0 & -0.7838 & 0 \\ 0 & 0 & -13.7049 & 0 \\ 0 & 0 & -0.2030 & 0 \\ 0 & 0 & -14.4887 & 0 \end{bmatrix} \\
 b_0 = (I - \alpha_0)^{-1}\beta_0 &= \begin{bmatrix} 0 & 0.0696 & -14.7178 \\ 0 & 0.2166 & -82.5799 \\ -0.0853 & 0.0180 & -1.9192 \\ 0 & 1.2862 & -97.2977 \end{bmatrix}
 \end{aligned}$$

což můžeme zkráceně napsat jako

$$x_t = Ax_{t-1} + Bu_t \quad (4.4)$$

Rovnice (4.4) je stavovou rovnicí našeho systému. Přitom $x \in \mathbb{R}^{16}$, $u \in \mathbb{R}^3$.

Rozptyly s_1^2, s_2^2, s_3^2 jsou nestrannými odhady rozptylů stochastické složky původního modelu (4.1). Tyto složky předpokládáme nezávislé a normální. Označme e šum původního modelu

$$e \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 11.230^2 & 0 & 0 \\ 0 & 23.864^2 & 0 \\ 0 & 0 & 0.854^2 \end{bmatrix} \right)$$

Po restrukturalizaci se tento šum změnil na šum v , pro který platí

$$\begin{aligned}
 \mathbb{E}v &= (I - \alpha_0)^{-1}\mathbb{E}e = 0 \\
 \text{var } v &= (I - \alpha_0)^{-1}\text{var } e((I - \alpha_0)^{-1})^T
 \end{aligned}$$

Matice $\text{var } v$ nemusí být nutně diagonální a složky šumu v jsou tedy stochasticky závislé.

Pro zkoumání stability systému je podle (3.9) nutné znát spektrum matice A . To bylo

spočteno programem Matlab:

$$\Lambda = \begin{bmatrix} -0.4905 + 0.4354i \\ -0.4905 - 0.4354i \\ 0.9878 \\ 0.9026 \\ 0.3937 + 0.5633i \\ 0.3937 - 0.5633i \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Simulace systému (4.1) se rozpadá do čtyř relativně nezávislých částí

- Příprava řídicích strategií, nejlépe ve více variantách
- Funkce $z = \text{pindyck}(x, u)$
- Funkce $[Y, X] = \text{sim}(x_0, U)$
- Grafické zpracování výsledků, tj. prezentace matic U, X, Y .

Kapitola 5

OPTIMÁLNÍ ODHAD STAVŮ SYSTÉMU

5.1 Optimální estimátor

Dynamické systémy, které jsme až doposud studovali, měly svůj vnitřní popis odvozený z popisu vnějšího a vektor stavů měl charakter umělé proměnné, která zprostředkovala zmíněný převod. Řada systémů je však často zadána přímo svým vnitřním popisem, ve kterém má vektor stavů svou konkrétní obsahovou interpretaci. Jako příklad je možno uvést celou řadu fyzikálních dějů, které probíhají v izolovaném prostředí. Pozorovateli je skryta přímá informace o hodnotách stavových proměnných. Namísto toho jsou do systému instalovány měřicí přístroje, které podávají zprostředkovanou informaci o stavu systému. Vystává tak problém, jak z napozorovaných výstupů y určit co nejpřesněji stavy x . Protože se předpokládá, že jak samotný proces tak i měření jsou zatíženy náhodnými chybami, jedná se o poměrně komplikovaný problém na poli matematické statistiky. Tento problém byl zatím úspěšně vyřešen pro případ lineárních systémů a Gaussovských šumů. Příslušný algoritmus pro výpočet neznámých stavů je znám pod názvem *Kalmanův filtr*. Pro nelineární systémy jsou známy pouze přibližné metody.

Předpokládejme neznámý náhodný vektor $X \in \mathcal{L}n_x$ a vektor naměřených dat $z \in \mathbb{R}^{n_z}$, který je realizací náhodného vektoru $Z \in \mathcal{L}n_z$. Předpokládejme dále, že známe simultánní hustotu pravděpodobnosti $p(x, z)$.

Předtím, než proběhne měření z , je naše informace o X obsažena pouze v hustotě $p(x, y)$ a pravděpodobnostní rozložení náhodného vektoru X udává hustota

$$p_0(x) = \int p(x, z) dz$$

Tato hustota se nazývá *apriorní hustota* náhodného vektoru X .

Po provedení měření z je naše informace o X obohacena a pravděpodobnostní rozložení vektoru X dáno podmíněnou hustotou pravděpodobnosti

$$p_1(x) = p(x|y)$$

kterou nazýváme *aposteriorní hustotou* náhodného vektoru X (viz B.4).

Často se namísto hustot určují pouze první a druhé momenty, tj. podmíněné střední hodnoty a podmíněné rozptyly, které mají některé významné vlastnosti.

Tyto úvahy jsou jádrem takzvaného *Bayesovského přístupu* k optimálnímu odhadu, který hojně používá podmíněné charakteristiky náhodných veličin.

Z uvedeného se zdá, že jsme pro tentýž náhodný vektor zavedli dvě pravděpodobnostní míry, které si navzájem konkurují. Ve skutečnosti je jediným skutečným rozložením rozložení apriorní, avšak aposteriorní rozložení má oproti apriornímu některé výhodné vlastnosti právě s ohledem na co nejpřesnější odhad *realizace* vektoru X , nikoliv hustoty vektoru X .

Přesnost odhadu se nejčastěji měří podle kritéria *nejmenších čtverců*.

Definice 5.1. Buď $\mathcal{E} \subset \mathbb{R}^{n_x}$ množina přípustných estimátorů náhodného vektoru X . Pro každý přípustný estimátor $y \in \mathcal{E}$ zavedme *střední kvadratickou chybu* vzorcem

$$J(y) = \mathbb{E}((X - y)^T(X - y))$$

Estimátor $\hat{X} \in \mathcal{E}$ nazveme *optimálním* podle kritéria nejmenších čtverců (least mean square), jestliže platí

$$J(\hat{X}) = \min_{y \in \mathcal{E}} J(y)$$

\hat{X} se také nazývá optimální mean square estimátor (MS estimátor).

Veta 5.2 (optimální apriorní estimátor). Střední hodnota $\mathbb{E} X$ je optimálním estimátorem na množině všech apriorních estimátorů.

Důkaz. Napišme si střední kvadratickou chybu

$$\begin{aligned} J(y) &= \mathbb{E}(X - y)^T(X - y) = \\ &= \mathbb{E} X^T X - y^T \mathbb{E} X - \mathbb{E} X^T y + y^T y = \\ &= \mathbb{E} X^T X - 2y^T \mathbb{E} X + y^T y \end{aligned}$$

kde jsme použili $\mathbb{E} y = y$ protože y není náhodná veličina, ale reálný vektor. Vidíme, že $J(y)$ je kvadratická funkce v proměnné y a globálního minima dosahuje v bodě \hat{X} , kde diferenciál je roven nule. Zapsáno

$$\frac{\partial J}{\partial y} = -2\mathbb{E} X + 2y = 0$$

nebo také

$$\hat{X} = \mathbb{E} X \tag{5.1}$$

□

Tímto jsme odvodili možná samozřejmý a určitě intuitivní výsledek, že při znalosti $p(x)$ je nejlepší konstantou pro odhad X střední hodnota X .

Zavedme chybu estimace $\tilde{X} \in \mathcal{L}_{n_x}$ jako rozdíl

$$\tilde{X} = X - \hat{X}$$

Pak platí

$$\mathbb{E} \tilde{X} = \mathbb{E} X - \hat{X} = \mathbb{E} X - \mathbb{E} X = 0,$$

takže střední chyba estimace je rovna nule. Říkáme, že odhad \hat{X} je *nestranný* (nevychýlený, unbiased). Ekvivalentně také platí

$$\mathbb{E} \tilde{X} = 0 \Leftrightarrow \mathbb{E} \hat{X} = \mathbb{E} X$$

Jako míru přesnosti naší estimace spočtíme rozptyl chyby

$$\begin{aligned}\text{var } \tilde{X} &= \text{E}(\tilde{X} - \text{E} \tilde{X})^T (\tilde{X} - \text{E} \tilde{X}) = \\ &= \text{E} \tilde{X}^T \tilde{X} = \\ &= \text{E}(X - \hat{X})^T (X - \hat{X}) = \\ &= \text{E}(X - \text{E} X)^T (X - \text{E} X) = \\ &= \text{var } X\end{aligned}$$

Bude-li rozptyl malý, je náš odhad dobrý. Inverzní matici $(\text{var } \tilde{X})^{-1}$ nazýváme *informační maticí*.

Věta 5.3 (optimální aposteriorní estimátor). Podmíněná střední hodnota $\text{E}(X|z)$ je optimálním estimátorem na množině všech aposteriorních estimátorů.

Důkaz. Předpokládejme, že vektor z je neprázdný, tj. $n_z \geq 1$, jinak by se jednalo o apriorní estimátor. Přitom známe hustotu $p(x, z)$. Optimální estimátor bude deterministickou funkcí naměřeného vektoru z , tj. $\hat{X} : \mathbb{R}^{n_z} \rightarrow \mathbb{R}$. Napišme funkcionál J

$$J = \int \int (x - \hat{X}(z))^T (x - \hat{X}(z)) p(x, z) dx dz$$

kde integrujeme přes celý prostor \mathbb{R}^{n_x} , případně \mathbb{R}^{n_z} . Podle řetězového pravidla (B.9) je

$$p(x, z) = p(x|z)p(z)$$

a funkcionál přejde do tvaru

$$J = \int p(z) \int (x - \hat{X}(z))^T (x - \hat{X}(z)) p(x|z) dx dz$$

Protože $p(z)$ a vnitřní integrál jsou nezáporné funkce, můžeme J minimalizovat tak, že minimalizujeme pro každé pevné z vnitřní integrál

$$H = \int (x - \hat{X}(z))^T (x - \hat{X}(z)) p(x|z) dx$$

Integrál H můžeme napsat jako

$$\begin{aligned}H &= \text{E}((X - \hat{X}(z))^T (X - \hat{X}(z)) | z) = \\ &= \text{E}(X^T X | z) - \hat{X}(z)^T \text{E}(X | z) - \text{E}(X^T | z) \hat{X}(z) + \hat{X}^T(z) \hat{X}(z) = \\ &= \text{E}(X^T X | z) - 2\hat{X}(z)^T \text{E}(X | z) + \hat{X}(z)^T \hat{X}(z)\end{aligned}$$

Stejným způsobem jako minule dostáváme

$$\frac{\partial H}{\partial \hat{X}} = -2\text{E}(X^T X | Z) + 2\hat{X}(Z) = 0$$

což dává

$$\hat{X}(z) = \text{E}(X | z) \tag{5.2}$$

□

To už není tak samozřejmý výsledek jako vzorec (5.1), ale je také intuitivní. Jestliže máme k dispozici realizaci z náhodného vektoru Z , pak nejlepším MS-estimátorem náhodného vektoru X je střední hodnota podmíněná $Z = z$.

Rovnice (5.2) vypadá velmi jednoduše, avšak jednoduchost této rovnice skrývá komplikace při jejím konkrétním použití. Ve skutečnosti může být výpočet podmíněné střední hodnoty podstatným problémem. Navíc se předpokládá, že známe přesně simultánní rozložení $p(x, z)$, což nemusí být pravda. V některých případech známe pouze první a druhé momenty.

Optimální aposteriorní estimátor je *nestranný*, což dokazuje výpočet

$$\begin{aligned} E \tilde{X} &= E(E(\tilde{X}|Z)) = E(E(X - E(X|Z)|Z)) = \\ &= E(X - E(X|Z)|Z) = E(X|Z) - E(X|Z) = 0 \end{aligned}$$

Míra spolehlivosti estimátoru je dána rozptylem chyby estimace, který se spočte jako

$$\begin{aligned} \text{var } \tilde{X} &= \text{var}[(X - \hat{X})|Z] = \\ &= E[(X - E X|Z)(X - E X|Z)^T|Z] \end{aligned}$$

Výraz v hranatých závorkách je ale podle definice podmíněný rozptyl $\text{var}(X|Z)$, takže rozptyl chyby můžeme psát jako

$$\text{var } \tilde{X} = E(\text{var}(X|Z)) \quad (5.3)$$

Jestliže X a Z jsou nezávislé vektory, pak $p(x, z) = p(x)p(z)$ a podle (B.13) platí

$$\begin{aligned} E(X|z) &= \int x \frac{p(x)p(z)}{p(z)} dx = \\ &= \int xp(x) dx = \\ &= E X \end{aligned}$$

Při nezávislosti tedy aposteriorní a apriorní estimátor splývají, což můžeme interpretovat tak, že měření Z nám nepřináší žádnou novou informaci ohledně X .

Veta 5.4 (gaussovské pozorování). Nechť platí

$$\begin{bmatrix} X \\ Z \end{bmatrix} \sim N \left(\begin{bmatrix} \mu_x \\ \mu_z \end{bmatrix}, \begin{bmatrix} \Sigma_x & \Sigma_{xz} \\ \Sigma_{zx} & \Sigma_z \end{bmatrix} \right)$$

Pak aposteriorní náhodná veličina $X|z$ má normální rozložení se střední hodnotou

$$E(X|z) = \mu_x + \Sigma_{xz}\Sigma_z^{-1}(z - \mu_z) \quad (5.4)$$

a varianční maticí

$$\text{var}(X|z) = \Sigma_x - \Sigma_{xz}\Sigma_z^{-1}\Sigma_{zx} \quad (5.5)$$

Důkaz. Definujme náhodný vektor $Y \in \mathcal{L}n_x + n_z$ jako

$$Y = \begin{bmatrix} \mu_x \\ \mu_z \end{bmatrix}$$

Hustota pravděpodobnosti náhodného vektoru Y je dána vzorcem

$$p(y) = \frac{1}{\sqrt{(2\pi)^{n_x+n_z} |\Sigma_y|}} e^{-\frac{1}{2}(y-\mu_y)^T \Sigma_y^{-1} (y-\mu_y)}$$

kde

$$\mu_y = \begin{bmatrix} \mu_x \\ \mu_z \end{bmatrix} \quad \Sigma_y = \begin{bmatrix} \Sigma_x & \Sigma_{xz} \\ \Sigma_{zx} & \Sigma_z \end{bmatrix}$$

Podle (A.5) je

$$\Sigma_y^{-1} = \begin{bmatrix} D^{-1} & -D^{-1}\Sigma_{xz}\Sigma_z^{-1} \\ -\Sigma_z^{-1}\Sigma_{zx}D^{-1} & \Sigma_z^{-1} + \Sigma_z^{-1}\Sigma_{zx}D^{-1}\Sigma_{xz}\Sigma_z^{-1} \end{bmatrix}$$

kde Schurův komplement k Σ_x je $D = \Sigma_x - \Sigma_{xz}\Sigma_z^{-1}\Sigma_{zx}$. Podle (B.13) je

$$p(x|z) = \frac{p(y)}{p(z)}$$

neboli

$$\begin{aligned} p(x|z) &= \frac{\sqrt{(2\pi)^{n_z} |\Sigma_z|}}{\sqrt{(2\pi)^{n_x+n_z} |\Sigma_y|}} e^{-\frac{1}{2}[(y-\mu_y)^T \Sigma_y^{-1} (y-\mu_y) - (z-\mu_z)^T \Sigma_z^{-1} (z-\mu_z)]} = \\ &= \frac{1}{\sqrt{(2\pi)^{n_x} |\Sigma_y| / |\Sigma_z|}} e^{-\frac{1}{2}[(y-\mu_y)^T \Sigma_y^{-1} (y-\mu_y) - (z-\mu_z)^T \Sigma_z^{-1} (z-\mu_z)]} \end{aligned}$$

Napišme si exponent a upravujme ho

$$\begin{aligned} (y - \mu_y)^T \Sigma_y^{-1} (y - \mu_y) - (z - \mu_z)^T \Sigma_z^{-1} (z - \mu_z) &= \\ &= \begin{bmatrix} x - \mu_x \\ z - \mu_z \end{bmatrix}^T \Sigma_y^{-1} \begin{bmatrix} x - \mu_x \\ z - \mu_z \end{bmatrix} - (z - \mu_z)^T \Sigma_z^{-1} (z - \mu_z) = \\ &= (x - \mu_x)^T D^{-1} (x - \mu_x) - (x - \mu_x)^T D^{-1} \Sigma_{xz} \Sigma_z^{-1} (z - \mu_z) - \\ &- (z - \mu_z)^T \Sigma_z^{-1} \Sigma_{zx} D^{-1} (x - \mu_x) + (z - \mu_z)^T (\Sigma_z^{-1} + \Sigma_z^{-1} \Sigma_{zx} D^{-1} \Sigma_{xz} \Sigma_z^{-1}) (z - \mu_z) - \\ &- (z - \mu_z)^T \Sigma_z^{-1} (z - \mu_z) = \\ &= (x - \mu_x)^T D^{-1} (x - \mu_x) - (x - \mu_x)^T D^{-1} \Sigma_{xz} \Sigma_z^{-1} (z - \mu_z) - \\ &- (z - \mu_z)^T \Sigma_z^{-1} \Sigma_{zx} D^{-1} (x - \mu_x) + (z - \mu_z)^T \Sigma_z^{-1} \Sigma_{zx} D^{-1} \Sigma_{xz} \Sigma_z^{-1} (z - \mu_z) = \\ &= (x - \mu_x)^T D^{-1} (x - \mu_x) - (x - \mu_x)^T D^{-1} [\Sigma_{xz} \Sigma_z^{-1} (z - \mu_z)] - \\ &- [\Sigma_{xz} \Sigma_z^{-1} (z - \mu_z)]^T D^{-1} (x - \mu_x) + [\Sigma_{xz} \Sigma_z^{-1} (z - \mu_z)]^T D^{-1} [\Sigma_{xz} \Sigma_z^{-1} (z - \mu_z)] \end{aligned}$$

Z prvních dvou členů vzniklého výrazu vytkneme $(x - \mu_x)^T D^{-1}$ a z dalších dvou členů vytkneme $[\Sigma_{xz} \Sigma_z^{-1} (z - \mu_z)]^T D^{-1}$ a dostaneme

$$(x - \mu_x)^T D^{-1} [(x - \mu_x) - \Sigma_{xz} \Sigma_z^{-1} (z - \mu_z)] - [\Sigma_{xz} \Sigma_z^{-1} (z - \mu_z)]^T D^{-1} [(x - \mu_x) - \Sigma_{xz} \Sigma_z^{-1} (z - \mu_z)]$$

Nyní můžeme vytknout zprava výraz $[(x - \mu_x) - \Sigma_{xz} \Sigma_z^{-1} (z - \mu_z)]$, čímž dostaneme

$$\begin{aligned} &[(x - \mu_x) - \Sigma_{xz} \Sigma_z^{-1} (z - \mu_z)]^T D^{-1} [(x - \mu_x) - \Sigma_{xz} \Sigma_z^{-1} (z - \mu_z)] = \\ &= [x - (\mu_x + \Sigma_{xz} \Sigma_z^{-1} (z - \mu_z))]^T D^{-1} [x - (\mu_x + \Sigma_{xz} \Sigma_z^{-1} (z - \mu_z))] \end{aligned}$$

Podle (A.9) platí, že

$$|\Sigma_y| = |\Sigma_z| |D| \implies \frac{|\Sigma_y|}{|\Sigma_z|} = |D|$$

Položíme-li $\mu_{x|z} = \mu_x + \Sigma_{xz}\Sigma_z^{-1}(z - \mu_z)$, dostáváme nakonec, že

$$p(x|z) = \frac{1}{\sqrt{(2\pi)^{n_x} |D|}} e^{-\frac{1}{2}(x-\mu_{x|z})^T D^{-1}(x-\mu_{x|z})}$$

tedy podmíněný náhodný vektor $X|z$ má normální rozložení se střední hodnotou

$$\mu_{x|z} = \mu_x + \Sigma_{xz}\Sigma_z^{-1}(z - \mu_z)$$

a rozptylem

$$D = \Sigma_x - \Sigma_{xz}\Sigma_z^{-1}\Sigma_{zx}$$

což bylo potřeba dokázat \square

Právě uvedený výsledek je velmi důležitý a bude všestranně užitečný pro další práci. Ze vzorce (5.4) vyplývá, že $\text{var}(X|z) \leq \text{var} X$. Jestliže nastane speciální případ $\text{var}(X|z) < \text{var} X$, pak jsme po provedení měření z více jistí správnou hodnotou X , než jsme byli předtím. To znamená, že aposteriorní náhodná veličina $X|z$ je těsněji rozložena kolem své střední hodnoty $E(X|z)$, než apriorní náhodná veličina X kolem své střední hodnoty EX .

Rozptyl náhodné chyby odhadu je roven

$$\text{var}((X - E(X|z))|z) = \text{var}(X|z) - \text{var}(E(X|z)) = \text{var}(X|z) \quad (5.6)$$

Důležitá skutečnost je ta, že zatímco $E(X|z)$ závisí na změřené hodnotě z , varianční matice $\text{var}(X|z)$ na změřené hodnotě z *nezávisí*, proto ji můžeme spočítat ještě předtím, než provedeme příslušné měření.

Vraťme se nyní opět k obecnému vztahu mezi X a Z , danému simultánní hustotou pravděpodobnosti $p(x, z)$. Ukázali jsme, že nejlepším estimátorem X při daném z je $\hat{X}_{MS} = E(X|z)$. Výpočet tohoto estimátoru může být v některých případech velmi obtížný. Abychom tento výpočet výrazně zjednodušili, omezme množinu přípustných estimátorů jen na ty estimátory, které jsou lineární funkcí naměřeného vektoru z .

Protože třída přípustných estimátorů je nyní omezena, lineární MS estimátor nebude obecně tak dobrý jako nelineární MS estimátor.

Veta 5.5 (lineární MS estimátor). Optimální MS estimátor na přípustné množině lineárních estimátorů $\mathcal{E} = \{y = Az + b : A \in \mathbb{R}^{n_x \times n_z}, b \in \mathbb{R}^{n_x}\}$ je estimátor

$$\hat{X}_{LMS} = \mu_x + \Sigma_{xz}\Sigma_z^{-1}(z - \mu_z) \quad (5.7)$$

kde

$$\mu_x = EX, \quad \mu_z = EZ, \quad \Sigma_{xz} = \text{cov}(X, Z), \quad \Sigma_z = \text{var} Z$$

Důkaz. Počítejme opět střední kvadratickou chybu

$$J = E(X - \hat{X}(Z))^T (X - \hat{X}(Z)) = E(X - AZ - b)^T (X - AZ - b)$$

což můžeme vzhledem k (A.21) napsat jako

$$\begin{aligned} J &= E(\text{tr}(X - AZ - b)(X - AZ - b)^T) = \\ &= \text{tr} E(X - AZ - b)(X - AZ - b)^T = \\ &= \text{tr} E[(X - \mu_x) - (AZ + b - \mu_x)][(X - \mu_x) - (AZ + b - \mu_x)]^T = \\ &= \text{tr}[E(X - \mu_x)(X - \mu_x)^T - E(X - \mu_x)(AZ + b - \mu_x)^T - \\ &\quad - E(AZ + b - \mu_x)(X - \mu_x)^T + E(AZ + b - \mu_x)(AZ + b - \mu_x)^T] \end{aligned}$$

Označme $\Sigma_x = \text{var } X = E(X - \mu_x)(X - \mu_x)^T$. Předpokládejme navíc, že matice A je symetrická. Potom platí

$$\begin{aligned} J &= \text{tr} [\Sigma_x - \Sigma_{xz}A^T - A\Sigma_{zx} + E(AZ + b - \mu_x)(AZ + b - \mu_x)^T] = \\ &= \text{tr} [\Sigma_x - 2A\Sigma_{zx} + E(A(Z - \mu_z) + (A\mu_z + b - \mu_x))(A(Z - \mu_z) + (A\mu_z + b - \mu_x))^T] = \\ &= \text{tr} [\Sigma_x - 2A\Sigma_{zx} + A\Sigma_zA^T + (A\mu_z + b - \mu_x)(A\mu_z + b - \mu_x)^T] = \\ &= \text{tr} [\Sigma_x + A(\Sigma_z + \mu_z\mu_z^T)A^T + (b - \mu_x)(b - \mu_x)^T + 2A\mu_z(b - \mu_x)^T - 2A\Sigma_{zx}] \end{aligned}$$

Podle (A.16) a (A.17) platí

$$\begin{aligned} \frac{\partial}{\partial A} \text{tr} (A(\Sigma_z + \mu_z\mu_z^T)A^T) &= 2A(\Sigma_z + \mu_z\mu_z^T) \\ \frac{\partial}{\partial A} \text{tr} (2A\Sigma_{zx}) &= 2\Sigma_{xz} \\ \frac{\partial}{\partial A} \text{tr} (2A\mu_z(b - \mu_x)^T) &= 2(b - \mu_x)\mu_z^T \end{aligned}$$

Podle (A.21) a (A.10) pak máme

$$\begin{aligned} \frac{\partial}{\partial b} \text{tr} ((b - \mu_x)(b - \mu_x)^T) &= 2(b - \mu_x) \\ \frac{\partial}{\partial b} \text{tr} (2A\mu_z(b - \mu_x)^T) &= 2A\mu_z \end{aligned}$$

Můžeme tedy napsat parciální derivace

$$\frac{\partial J}{\partial A} = 2A(\Sigma_z + \mu_z\mu_z^T) - 2\Sigma_{xz} + 2(b - \mu_x)\mu_z^T = 0 \quad (5.8)$$

$$\frac{\partial J}{\partial b} = 2(b - \mu_x) + 2A\mu_z = 0 \quad (5.9)$$

Z rovnice (5.9) vypočteme

$$b = \mu_x - A\mu_z$$

a dosazením do (5.8) dostaneme

$$A\Sigma_z - \Sigma_{xz} = 0$$

což dává

$$A = \Sigma_{xz}\Sigma_z^{-1}$$

Optimální lineární MS estimátor je potom dán vzorcem

$$\hat{X}_{LMS} = \mu_x + \Sigma_{xz}\Sigma_z^{-1}(z - \mu_z)$$

což je (5.7) □

Prozkoumejme strukturu vzorce (5.7). Definujeme-li výstupní reziduum jako

$$\tilde{Z} = Z - EZ \quad (5.10)$$

pak vzorec (5.7) přejde v $\hat{X}_{LMS} = \mu_x + \Sigma_{xz}\Sigma_z^{-1}\tilde{z}$. Člen μ_x představuje apriorní estimátor pro X a druhý člen je korekce, založená na výstupním reziduu. Bude-li závislost mezi X a Z

velká, bude apriorní odhad silněji korigován. Naopak bude-li veliká nejistota ohledně Z (velký rozptyl Σ_z), bude korekční člen potlačen. Konečně změříme-li $z = E Z$, pak se v podstatě náš apriorní odhad potvrdil a k žádné korekci nedojde.

Srovnáme-li nelineární odhad s lineárním, pak vidíme, že zatímco ke konstrukci nelineárního odhadu jsme potřebovali znát dokonale simultánní hustotu $p(x, z)$, v případě lineárního odhadu úplně postačí první a druhé momenty. To je výrazná úspora dat při reálném výpočtu.

Věta 5.6. Estimátor \hat{X}_{LMS} je nestranný.

Důkaz.

$$\begin{aligned} E(\hat{X}_{LMS}) &= E\mu_x + \Sigma_{xz}\Sigma_z^{-1}E(Z - \mu_z) = \\ &= \mu_x + \Sigma_{xz}\Sigma_z^{-1}(\mu_z - \mu_z) = \\ &= E X \end{aligned}$$

□

Odtud plyne, že $E\tilde{X} = 0$ a můžeme psát

$$\begin{aligned} \text{var } \tilde{X} &= E(\tilde{X} - E\tilde{X})(\tilde{X} - E\tilde{X})^T = E\tilde{X}\tilde{X}^T = E(X - \hat{X})(X - \hat{X})^T = \\ &= E((X - \mu_x) - \Sigma_{xz}\Sigma_z^{-1}(Z - \mu_z))((X - \mu_x) - \Sigma_{xz}\Sigma_z^{-1}(Z - \mu_z))^T = \\ &= \Sigma_x - \Sigma_{xz}\Sigma_z^{-1}\Sigma_{zx} - \Sigma_{xz}\Sigma_z^{-1}\Sigma_{zx} + \Sigma_{xz}\Sigma_z^{-1}\Sigma_z\Sigma_z^{-1}\Sigma_{zx} = \\ &= \Sigma_x - \Sigma_{xz}\Sigma_z^{-1}\Sigma_{zx} \end{aligned}$$

Celkem tedy

$$\text{var } \tilde{X} = \Sigma_x - \Sigma_{xz}\Sigma_z^{-1}\Sigma_{zx} \quad (5.11)$$

Porovnáním vzorců (5.7), (5.11) se vzorci (5.4), (5.5) dospějeme k důležitému závěru:

Je-li vektor $(X^T, Z^T)^T$ rozložen normálně, pak estimátory \hat{X}_{MS} a \hat{X}_{LMS} splývají.

Příklad 5.7. Necht' náhodná veličina X má rozložení $X \sim Rs(0, 1)$. Tuto náhodnou veličinu měříme pomocí náhodné veličiny Z . Vztah mezi X a Z je určen rovnicí měření

$$Z = \ln \frac{1}{X} + V$$

kde $V \sim Ex(1)$ je šum měření. Spočítáme optimální estimátory \hat{X}_{MS} a \hat{X}_{LMS} . Nejprve si napíšeme hustoty pravděpodobnosti $p(x), p(v)$:

$$\begin{aligned} p(x) = p_X(x) &= \begin{cases} 1 & \text{pro } 0 \leq x \leq 1 \\ 0 & \text{jinak} \end{cases} \\ p(v) = p_V(v) &= \begin{cases} e^{-x} & \text{pro } x \geq 0 \\ 0 & \text{pro } x < 0 \end{cases} \end{aligned}$$

Hustota p_X je apriorní hustota a její střední hodnota $\frac{1}{2}$ je apriorním estimátorem náhodné veličiny X . Rozptyl chyby činí $\frac{1}{12}$.

Ze znalosti výstupní rovnice a rozložení šumu můžeme ihned určit podmíněnou hustotu $p(z|x)$:

$$p(z|x) = p_{Z|X}(z, x) = p_V(z - \ln \frac{1}{x}) = \begin{cases} e^{-(z - \ln \frac{1}{x})} = \frac{1}{x} e^{-z} & \text{pro } x \geq e^{-z} \\ 0 & \text{pro } x < e^{-z} \end{cases}$$

sdružená hustota $p(x, z)$ je dána vzorcem $p(x, z) = p(z|x)p(x)$, neboli

$$p(x, z) = \begin{cases} \frac{1}{x} e^{-z} & \text{pro } x \geq e^{-z}, 0 \leq x \leq 1 \\ 0 & \text{jinak} \end{cases}$$

Marginalizací dostaneme hustotu $p(z) = p_Z(z)$:

$$\begin{aligned} p_Z(z) &= \int p(x, z) dx = \\ &= \int_{e^{-z}}^1 \frac{1}{x} e^{-z} dx = \\ &= e^{-z} [\ln x]_{e^{-z}}^1 = \\ &= e^{-z} (0 - \ln e^{-z}) = \\ &= z e^{-z} \text{ pro } z \geq 0 \end{aligned}$$

Nyní už snadno spočteme aposteriorní hustotu $p(x|z)$ jako

$$\begin{aligned} p_{X|Z}(x, z) &= \frac{p(x, z)}{p(z)} = \\ &= \frac{\frac{1}{x} e^{-z}}{z e^{-z}} = \\ &= \begin{cases} \frac{1}{xz} & \text{pro } x \geq e^{-z}, 0 \leq x \leq 1 \\ 0 & \text{jinak} \end{cases} \end{aligned}$$

Aposteriorní estimátor je střední hodnotou aposteriorní náhodné veličiny $X|Z$ s hustotou $p_{X|Z}$, tedy

$$\begin{aligned} \hat{X}_{MS}(z) &= E(X|z) = \int xp(x|z) dx = \\ &= \int_{e^{-z}}^1 x \frac{1}{xz} dx = \\ &= \frac{1}{z} (1 - e^{-z}) = \\ &= \frac{1 - e^{-z}}{z} \end{aligned}$$

Spočteme rozptyl aposteriorní chyby. Podle (5.3) máme

$$\begin{aligned}
 \text{var } \tilde{X} &= \text{E}(\text{var}(X|Z)) = \\
 &= \int \int (x - \text{E}(X|z))^2 p(x|z) dx dz = \\
 &= \int \int x^2 p(x|z) dx - \text{E}(X|z)^2 dz = \\
 &= \int_0^\infty \int_{e^{-z}}^1 x^2 \frac{1}{xz} dx - \frac{(1 - e^{-z})^2}{z^2} dz = \\
 &= \int_0^\infty \frac{1}{z} \int_{e^{-z}}^1 x dx - \frac{(1 - e^{-z})^2}{z^2} dz = \\
 &= \int_0^\infty \frac{1}{2z} (1 - e^{-2z}) - \frac{(1 - e^{-z})^2}{z^2} dz = \\
 &= \int_0^\infty \frac{1 - e^{-2z}}{2z} - \frac{1 - 2e^{-z} + e^{-2z}}{z^2} dz = \\
 &= \int_0^\infty \frac{(z - 2) + 4e^{-z} - (z - 2)e^{-2z}}{2z^2} dz =
 \end{aligned}$$

Nyní se budeme věnovat lineárnímu estimátoru \hat{X}_{LMS} , který je určen vzorcem (5.7). Podle předchozího $\mu_x = 0.5$. Zbývá spočítat $\mu_z, \Sigma_z, \Sigma_{xz}$:

$$\begin{aligned}
 \mu_z &= \int zp_Z(z) dz = \\
 &= \int_0^\infty z^2 e^{-z} dz = \\
 &= 2
 \end{aligned}$$

$$\begin{aligned}
 \text{E } XZ &= \int \int xzp(x, z) dx dz = \\
 &= \int_0^\infty \int_{e^{-z}}^1 xz e^{-z} dx dz = \\
 &= \frac{1}{2} \int_0^\infty z^2 e^{-z} (1 - e^{-2z}) dz = \\
 &= \frac{3}{4} = 0.75
 \end{aligned}$$

$$\begin{aligned}
 \Sigma_{xz} &= \text{E } XZ - \mu_x \mu_z = \\
 &= \frac{3}{4} - \frac{2}{2} = \\
 &= -\frac{1}{4} = -0.25
 \end{aligned}$$

$$\begin{aligned}
E Z^2 &= \int z^2 p_Z(z) dz = \\
&= \int_0^\infty \int_{e^{-z}}^1 z^2 z e^{-z} dx dz = \\
&= \int_0^\infty z^2 z e^{-z} (1 - e^{-z}) dz = \\
&= 6
\end{aligned}$$

$$\begin{aligned}
\Sigma_z &= E Z^2 - \mu_z^2 = \\
&= 6 - 2^2 = \\
&= 2
\end{aligned}$$

Podle (5.7) se lineární estimátor vypočte jako

$$\begin{aligned}
\hat{X}_{LMS} &= \mu_x + \Sigma_{xz} \Sigma_z^{-1} (z - \mu_z) = \\
&= \frac{1}{2} - \frac{1}{4} \frac{1}{2} (z - 2) = \\
&= \frac{3}{4} - \frac{1}{8} z = 0.75 - 0.125z
\end{aligned}$$

5.2 Odhad stavu dynamického systému

Zkoumejme nyní diskrétní dynamický systém (3.4) z hlediska optimálního odhadu stavů. Dynamický systém nechť je definován rovnicemi

$$x_{t+1} = f(x_t, u_t, v_t) \quad (5.12a)$$

$$y_t = g(x_t, u_t, w_t) \quad (5.12b)$$

a počátečním rozložením

$$p(x_0) \quad (5.12c)$$

Budeme postupně konstruovat optimální odhady stavů x_0, x_1, \dots, x_t rekurentním způsobem. Pro odhad stavů máme k dispozici pouze časovou řadu vstupů a časovou řadu výstupů. Rekurentní způsob výpočtu je optimální pro potřeby praxe. V realitě je předmětem zájmu aktuální stav x_t , který se jednoduše vypočte na základě minulého odhadu x_{t-1} a nových dat u_t, y_t .

Protože rozložení náhodného šumu v_t pokládáme za známé, je rovnicí (5.12a) jednoznačně určena podmíněná hustota pravděpodobnosti $p(x_{t+1}|x_t, u_t)$. Analogicky je rovnicí (5.12b) jednoznačně určena hustota $p(y_t|x_t, u_t)$.

Označme symbolem D_t data, vzniklá pozorováním systému na horizontu $0, \dots, t$

$$D_t = \{u_0, y_0, \dots, u_t, y_t\} \quad (5.13)$$

Protože jsou obě rovnice (5.12a) a (5.12b) nezávislé na datech D_{t-1} , můžeme podle (B.11) psát

$$p(x_{t+1}|x_t, u_t, D_{t-1}) = p(x_{t+1}|x_t, u_t) \quad (5.14)$$

$$p(y_t|x_t, u_t, D_{t-1}) = p(y_t|x_t, u_t) \quad (5.15)$$

Vztah (5.15) říká, že stav systému x_t obsahuje veškerou informaci obsaženou v datech D_{t-1} , která je potřebná k predikci výstupu y_t . Tuto vlastnost můžeme pokládat za ekvivalentní definici stavu.

Z rovnice (5.12a) dále vyplývá, že řízení u_t neovlivňuje přímo stav x_t , což můžeme zapsat jako

$$p(x_t|D_{t-1}, u_t) = p(x_t|D_{t-1}) \quad (5.16)$$

Tato vlastnost se někdy označuje jako *přirozené podmínky řízení*¹.

Předpokládejme, že se v současnosti nacházíme v časovém okamžiku $t \in \mathbb{N}$ a máme k dispozici data D_{t-1} . Chtěli bychom nyní na základě změřených dat D_{t-1} co nejlépe odhadnout neznámou hodnotu stavu x_t . Tuto hodnotu nejlépe aproximuje podmíněná hustota pravděpodobnosti $p(x_t|D_{t-1})$. Nazveme ji *apriorní hustota*. Předpokládejme dále, že apriorní hustota je známá funkce. Po doplnění datového souboru aktuálními hodnotami u_t, y_t máme k dispozici zaktualizovaná data D_t , na jejichž základě máme možnost zaktualizovat také náš odhad stavu x_t . Hledáme tedy novou podmíněnou hustotu $p(x_t|D_t)$ —tzv. *aposteriorní hustotu*.

Veta 5.8 (aposteriorní hustota). Aposteriorní hustota $p(x_t|D_t)$ je dána vzorcem

$$p(x_t|D_t) = \frac{p(y_t|x_t, u_t)}{p(y_t|u_t, D_{t-1})}p(x_t|D_{t-1}) \quad (5.17)$$

Důkaz. Použitím Bayesova vzorce (B.10) a vztahů (5.14),(5.16) dostáváme

$$\begin{aligned} p(x_t|D_t) &= p(x_t|y_t, u_t, D_{t-1}) = \\ &= \frac{p(y_t|x_t, u_t, D_{t-1})p(x_t|u_t, D_{t-1})}{p(y_t|u_t, D_{t-1})} = \\ &= \frac{p(y_t|x_t, u_t)}{p(y_t|u_t, D_{t-1})}p(x_t|D_{t-1}), \end{aligned}$$

což je vzorec (5.17) □

Jmenovatel $p(y_t|u_t, D_{t-1})$ je konstanta, protože y_t, u_t, D_{t-1} jsou známé hodnoty. Toto číslo funguje jako normalizační konstanta.

Přechod od apriorní hustoty k aposteriorní hustotě pomocí (5.17) se nazývá *datový krok*. Je dobré si všimnout, že jsme při něm použili pouze znalost výstupní rovnice (5.12b).

Aby bylo možné počítat rekurentně aposteriorní hustoty, je potřeba odvodit přechod od aposteriorní hustoty $p(x_t|D_t)$ k apriorní hustotě $p(x_{t+1}|D_t)$. Tento přechod se někdy nazývá *časový* nebo též *predikční* krok odhadu stavu.

Veta 5.9. Predikční krok odhadu stavu je dán vztahem

$$p(x_{t+1}|D_t) = \int p(x_{t+1}|x_t, u_t)p(x_t|D_t) dx_t \quad (5.18)$$

Důkaz. Podle řetězového pravidla (B.9) platí

$$p(x_{t+1}, x_t|D_t) = p(x_{t+1}|x_t, D_t)p(x_t|D_t)$$

¹Poznamenejme ovšem, že zejména u ekonomických systémů nebývá často tato podmínka splněna - viz příklad 4.2.1.

Vztah (5.14) dává

$$p(x_{t+1}|x_t, D_t) = p(x_{t+1}|x_t, u_t, y_t)$$

stav x_{t+1} je však nezávislý na výstupu y_t , takže můžeme psát

$$p(x_{t+1}|x_t, D_t) = p(x_{t+1}|x_t, u_t)$$

marginalizací podle (B.1) pak dostaneme

$$p(x_{t+1}|D_t) = \int p(x_{t+1}|x_t, u_t)p(x_t|D_t) dx_t$$

což jsme měli dokázat □

Posledním problémem je zahájení rekurze. Je nutné znát apriorní hustotu

$$p(x_0|D_{-1}) = p(x_0) \tag{5.19}$$

Apriorní hustota $p(x_0)$ je buďto předem známa, nebo se určuje zkusmo a podle zkušenosti, což zanáší do výpočtu subjektivní prvek.

Datový a časový krok algoritmu lze spojit do jediného kroku

$$\begin{aligned} p(x_{t+1}|D_t) &= \int \frac{p(x_{t+1}|x_t, u_t, y_t)p(y_t|x_t, u_t)}{p(y_t|u_t, D_{t-1})} p(x_t|D_{t-1}) dx_t = \\ &= \frac{p(x_{t+1}, y_t|u_t, D_{t-1})}{p(y_t|u_t, D_{t-1})} \end{aligned}$$

Tímto způsobem jsme ale ztratili aposteriorní hustoty, které jsou obecně lepší než hustoty apriorní.

5.3 Kalmanův filtr

Kalmanův filtr je analogií výše uvedené nelineární rekurze pro speciální případ, kdy jak stavy x tak výstupy y jsou rozloženy normálně. Pak je podle 5.4 apriorní i aposteriorní hustota opět normální.

Normálně rozdělenou náhodnou veličinu plně určuje její střední hodnota a rozptyl. Stačí tedy po celou dobu výpočtu sledovat pouze tyto dvě číselné charakteristiky, což v konečném důsledku umožňuje nahradit funkcionální rekurze rekurzemi algebraickými. Celý výpočet se tak snadno implementuje na počítači.

Mají-li být stavy rozloženy normálně, musí být příslušný dynamický systém nutně *lineární* s gaussovskými šumy. To je největší slabinou Kalmanova filtru, kterou se snaží odstranit některé novější algoritmy, které jsou ovšem buď nepřesné nebo podstatně složitější.

Veta 5.10 (Kalmanův filtr). Buď dán lineární diskrétní stochastický systém, popsany rovnicemi

$$x_{t+1} = Ax_t + Bu_t + v_t \tag{5.20a}$$

$$y_t = Cx_t + Du_t + w_t \tag{5.20b}$$

splňující tyto podmínky:

- (i) $x_0 \sim N(\mu_0, \Sigma_0)$
- (ii) $v_t \sim N(0, \Sigma_v)$ pro všechna t
- (iii) $w_t \sim N(0, \Sigma_w)$ pro všechna t
- (iv) $Ev_t w_t^T = 0$
- (v) $Ex_0 v_t^T = 0$

kde vektor μ_0 a matice $\Sigma_0, \Sigma_v, \Sigma_w$ jsou známé. Dále necht' jsou známa data D_t . Označme $x_{t|k} = x_t | D_k$. Pak apriorní estimátor $x_{t|t-1}$ a aposteriorní estimátor $x_{t|t}$ jsou normální pro všechna t

$$x_{t|t-1} \sim N(\mu_{t|t-1}, \Sigma_{t|t-1}) \quad (5.21)$$

$$x_{t|t} \sim N(\mu_{t|t}, \Sigma_{t|t}) \quad (5.22)$$

kde střední hodnoty $\mu_{t|t-1}, \mu_{t|t}$ a varianční matice $\Sigma_{t|t-1}, \Sigma_{t|t}$ se vypočtou podle vzorců

$$\mu_{t|t} = \mu_{t|t-1} + K_t(y_t - C\mu_{t|t-1} - Du_t) \quad (5.23)$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - K_t C \Sigma_{t|t-1} \quad (5.24)$$

$$\mu_{t+1|t} = A\mu_{t|t} + Bu_t \quad (5.25)$$

$$\Sigma_{t+1|t} = A\Sigma_{t|t}A^T + \Sigma_v \quad (5.26)$$

$$K_t = \Sigma_{t|t-1}C^T(C\Sigma_{t|t-1}C^T + \Sigma_w)^{-1} \quad (5.27)$$

Všimněme si rekurentní charakter vztahů (5.23)-(5.27). Na základě počáteční podmínky i a dat D_t jsme schopni postupně spočítat odhady $x_{1|0}, x_{1|1}, x_{2|1}, x_{2|2}, \dots, x_{t|t-1}, x_{t|t}$. Přitom ve výpočtu vystupují pouze první a druhé momenty těchto veličin.

Rovnice (5.23),(5.24) tvoří *datový krok* algoritmu, což je přechod od $x_{t|t-1}$ k $x_{t|t}$. Naproti tomu rovnice (5.25),(5.26) tvoří *predikční krok*, tedy přechod od $x_{t|t}$ k $x_{t+1|t}$.

Dosadíme-li (5.27) do rovnice (5.24)(viz (5.29)), dostaneme diferenční rovnici pro $\Sigma_{t|t}$, která je maticovou diskretní variantou *Riccatiho diferenční rovnice* tvaru $\dot{x} = a_0 + a_1x + a_2x^2$.

Důkaz. Předpokládejme, že platí (5.21). Označme $y_{t|k} = y_t | D_k$. Pak

$$\begin{aligned} y_{t|t-1} &= y_t | D_{t-1} = \\ &= (Cx_t + Du_t + w_t) | D_{t-1} = \\ &= Cx_{t|t-1} + w_t | D_t = \\ &= Cx_{t|t-1} + w_t \end{aligned}$$

Protože $y_{t|t-1}$ je lineární kombinací náhodných vektorů $x_{t|t-1}, w_t$, které jsou podle předpokladu normální, je $y_{t|t-1}$ podle (B.6) normální se střední hodnotou $C\mu_{t|t-1} + Du_t$ a rozptylem

$$\begin{aligned} \text{var}(y_{t|t-1}) &= \text{var}(Cx_{t|t-1} + Du_t + w_t) = \\ &= C\text{var}(x_{t|t-1})C^T + \text{var}w_t = \\ &= C\Sigma_{t|t-1}C^T + \Sigma_w \end{aligned}$$

Spočítejme kovarianční matici

$$\begin{aligned} \text{cov}(x_{t|t-1}, y_{t|t-1}) &= \text{cov}(x_{t|t-1}, Cx_{t|t-1} + Du_t + w_t) = \\ &= \text{var}(x_{t|t-1})C^T + \text{cov}(x_{t|t-1}, w_t) = \\ &= \Sigma_{t|t-1}C^T \end{aligned}$$

Spojený náhodný vektor $(x_{t|t-1}, y_{t|t-1})$ je normálně rozdělen

$$\begin{bmatrix} x_{t|t-1} \\ y_{t|t-1} \end{bmatrix} \sim N \left(\begin{bmatrix} \mu_{t|t-1} \\ C\mu_{t|t-1} + Du_t \end{bmatrix}, \begin{bmatrix} \Sigma_{t|t-1} & \Sigma_{t|t-1}C^T \\ C\Sigma_{t|t-1} & C\Sigma_{t|t-1}C^T + \Sigma_w \end{bmatrix} \right)$$

Jsou tedy splněny předpoklady věty 5.4 a pro podmíněnou náhodnou veličinu

$$x_{t|t-1}|y_t = x_t|(y_t, D_{t-1}) = x_t|(y_t, u_t, D_{t-1}) = x_t|D_t = x_{t|t}$$

platí, že je normální

$$x_{t|t} \sim N(\mu_{t|t}, \Sigma_{t|t})$$

s parametry

$$\mu_{t|t} = \mu_{t|t-1} + \Sigma_{t|t-1}C^T(C\Sigma_{t|t-1}C^T + \Sigma_w)^{-1}(y_t - C\mu_{t|t-1} - Du_t) \quad (5.28)$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - \Sigma_{t|t-1}C^T(C\Sigma_{t|t-1}C^T + \Sigma_w)^{-1}C\Sigma_{t|t-1} \quad (5.29)$$

Položíme-li

$$K_t = \Sigma_{t|t-1}C^T(C\Sigma_{t|t-1}C^T + \Sigma_w)^{-1}$$

pak můžeme psát

$$\begin{aligned} \mu_{t|t} &= \mu_{t|t-1} + K_t(y_t - C\mu_{t|t-1} - Du_t) \\ \Sigma_{t|t} &= \Sigma_{t|t-1} - K_tC\Sigma_{t|t-1} \end{aligned}$$

což jsou rovnice (5.23) a (5.24), které tvoří datový krok algoritmu. Nyní odvodíme predikční krok.

$$\begin{aligned} x_{t+1|t} &= (Ax_t + Bu_t + v_t)|D_t = \\ &= A(x_t|D_t) + Bu_t + (v_t|D_t) = \\ &= Ax_{t|t} + Bu_t + v_t \end{aligned}$$

Vidíme, že $x_{t+1|t}$ je lineární kombinací $x_{t|t}$ a v_t , které jsou normální. $x_{t+1|t}$ je tedy podle (B.6) také normální s parametry

$$\begin{aligned} \mu_{t+1|t} &= A\mu_{t|t} + Bu_t \\ \Sigma_{t+1|t} &= A\Sigma_{t|t}A^T + \Sigma_v \end{aligned}$$

což jsou rovnice (5.25),(5.26). Tím je věta dokázána □

Příklad 5.11. Loď se pohybuje po rovníku východním směrem rychlostí 10 námořních mil za hodinu. Okamžitou rychlost lodi ovšem ovlivňují náhodné poryvy větru a nárazy vln. Navigátor lodi má za úkol každou hodinu odhadnout zeměpisnou délku l v minutách a rychlost lodi $s = dl/dt$ v mph.

V čase $t = 0$ navigátor odhadnul polohu $l_0 = 0$ a rychlost $s_0 = 10$. Každou hodinu pak změřil sextantem zeměpisnou šířku a údaje zaznamenal do tabulky:

| Čas | Poloha |
|-----|--------|
| 1 | 9.0" |
| 2 | 19.5" |
| 3 | 29.0" |

Označíme-li l_t, s_t polohu a rychlost lodi v čase t , pak úloha navigátora je úlohou optimálního odhadu veličin l_t, s_t .

Počáteční odhady navigátora můžeme matematicky modelovat jako nezávislé náhodné veličiny s normálním rozložením. Rozptyly navigátorových odhadů jsou dlouhodobě sledovány a jejich hodnoty jsou $\text{var } l_0 = 2$, $\text{var } s_0 = 3$.

První krok při řešení problému je namodelovat dynamiku systému. Během hodiny t se loď pohybuje rychlostí s_t mph, takže její poloha se změní na²

$$l_{t+1} = l_t + s_t \quad (5.30)$$

Veličina s_{t+1} by měla být rovna s_t , protože lodivod udržuje konstantní rychlost 10 mph. Protože ale je rychlost náhodně ovlivněna účinkem větru a vln, přidáme do rovnice bílý gaussovský šum e , jehož rozptyl byl dlouhodobým měřením odhadnut na číslo 1.

$$s_{t+1} = s_t + e_t \quad (5.31)$$

Definujeme-li stavový vektor x_t jako

$$x_t = \begin{bmatrix} l_t \\ s_t \end{bmatrix},$$

můžeme (5.30),(5.31) napsat jako jedinou stavovou rovnici systému

$$x_{t+1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} x_t + v_t, \quad v_t \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) \quad (5.32)$$

Šum procesu v_t je singulární gaussovský šum. Počáteční podmínka je podle předchozího

$$x_0 \sim N(\mu_0, \Sigma_0) = N\left(\begin{bmatrix} 0 \\ 10 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}\right)$$

Tímto je kompletně popsána dynamika systému. Rovnici měření zapíšeme jako

$$y_t = \begin{bmatrix} 1 & 0 \end{bmatrix} x_t + w_t, \quad w_t \sim N(0, 2) \quad (5.33)$$

kde $\Sigma_w = 2$ je rozptyl měření sextantu, udávaný výrobcem. Z rovnic (5.32),(5.33) je zřejmé, že se jedná o lineární gaussovský systém bez deterministického vstupu u_t . Teď je vše připraveno k

²Jedna ujetá námořní míle odpovídá změně zeměpisné šířky o jednu úhlovou minutu.

optimálnímu odhadu stavů Kalmanovým filtrem. Použijeme obecné rekurentní vztahy (5.23)–(5.27) s počáteční podmínkou

$$\mu_{0|-1} = \begin{bmatrix} 0 \\ 10 \end{bmatrix}, \quad \Sigma_{0|-1} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \quad (5.34)$$

Výsledky uvedeme přehledně v tabulce

| t | y_t | $\mu_{t t-1}$ | $\Sigma_{t t-1}$ | | $\mu_{t t}$ | $\Sigma_{t t}$ | |
|-----|-------|---------------|------------------|-------|-------------|----------------|-------|
| 0 | — | 0.000 | 2.000 | 0.000 | 0.000 | 2.000 | 0.000 |
| | | 10.000 | 0.000 | 3.000 | 10.000 | 0.000 | 3.000 |
| 1 | 9.0 | 10.000 | 5.000 | 3.000 | 9.286 | 1.429 | 0.857 |
| | | 10.000 | 3.000 | 4.000 | 9.571 | 0.857 | 2.714 |
| 2 | 19.5 | 18.857 | 5.857 | 3.571 | 19.336 | 1.491 | 0.909 |
| | | 9.571 | 3.571 | 3.714 | 9.864 | 0.909 | 2.091 |
| 3 | 29.0 | 29.200 | 5.400 | 3.000 | 29.054 | 1.460 | 0.811 |
| | | 9.864 | 3.000 | 3.091 | 9.783 | 0.811 | 1.875 |

Všimněme si, že variance v modelovém kroku vzroste v důsledku nejistoty stavové rovnice. Naproti tomu v datovém kroku variance klesá v důsledku zapracování dodatečné informace do naší předpovědi.

5.4 Smoothing

Smoothing je zkráceně řečeno optimální odhad stavu x_t dynamického systému (5.12), založený na datech D_N , kde $N > t$. Výsledkem je tedy podmíněný náhodný vektor $x_{t|N}$. V porovnání s apriorním odhadem $x_{t|t-1}$ a aposteriorním odhadem $x_{t|t}$ máme pro odhad stavu k dispozici více informace a výsledný estimátor tedy bude ještě *lepší* než oba uvedené typy. Naproti tomu smoothing nelze provádět v aktuálním čase t , zatímco rekurzivní odhady ano. Smoothing se tedy nejspíše uplatní tam, kde nás kromě aktuálního stavu x_t zajímají také minulé stavy x_0, \dots, x_{t-1} . Jak si ukážeme, lze odhady $x_{0|t}, \dots, x_{t-1|t}$ získat úpravou již existujících aposteriorních odhadů $x_{0|0}, \dots, x_{t-1|t-1}$. Algoritmus přitom postupuje zpětně v čase.

Věta 5.12. Nechť $t < N$. Pro podmíněnou hustotu $p(x_t|D_N)$ platí:

$$p(x_t|D_N) = p(x_t|D_t) \int p(x_{t+1}|x_t, u_t) \frac{p(x_{t+1}|D_N)}{p(x_{t+1}|D_t)} dx_{t+1} \quad (5.35)$$

Důkaz.

$$p(x_t|D_N) = \int p(x_t, x_{t+1}|D_N) dx_{t+1} \quad (5.36)$$

Počítejme

$$p(x_t, x_{t+1}|D_N) = p(x_t|x_{t+1}, D_N)p(x_{t+1}|D_N) \quad (5.37)$$

Rozberme nyní podrobněji podmíněnou hustotu $p(x_t|x_{t+1}, D_N)$. Veškerá informace uložená v datech $\{y_{t+1}, u_{t+1}, \dots, y_N, u_N\}$ potřebná k predikci x_t je koncentrována v hodnotě x_{t+1} , což vyplývá z konstrukce modelových rovnic (5.12). Můžeme tedy prohlásit, že

$$p(x_t|x_{t+1}, D_N) = p(x_t|x_{t+1}, D_t) \quad (5.38)$$

přítom

$$\begin{aligned} p(x_t|x_{t+1}, D_t) &= p(x_{t+1}|x_t, D_t) \frac{p(x_t|D_t)}{p(x_{t+1}|D_t)} = \\ &= p(x_{t+1}|x_t, u_t) \frac{p(x_t|D_t)}{p(x_{t+1}|D_t)} \end{aligned}$$

dosazením do (5.37) a (5.36) nakonec dostáváme

$$\begin{aligned} p(x_t|D_N) &= \int p(x_{t+1}|x_t, u_t) \frac{p(x_t|D_t)}{p(x_{t+1}|D_t)} p(x_{t+1}|D_N) dx_{t+1} = \\ &= p(x_t|D_t) \int p(x_{t+1}|x_t, u_t) \frac{p(x_{t+1}|D_N)}{p(x_{t+1}|D_t)} dx_{t+1} \end{aligned}$$

což jsme měli dokázat □

Povšimněme si, že ve výpočtu nikde nevyužíváme naměřená data. Využíváme je pouze zprostředkovaně, a to prostřednictvím estimátorů $x_{t+1|N}$ a $x_{t|t}$. Tyto dva estimátory jsou také postačující pro výpočet estimátoru $x_{t|N}$.

Rekurzivní vztah (5.35) se v případě gaussovského systému opět rozpadne na maticové vzorce pro střední hodnotu a rozptyl estimátoru. Tyto vzorce, známé pod názvem Rauch-Tung-Striebel Smoother, uvedeme bez důkazu.

Veta 5.13 (Rauch-Tung-Striebel Smoother). Pro lineární systém 5.20a se odhad $x_{t|N} \sim N(\mu_{t|N}, \Sigma_{t|N})$ vypočte podle vzorců

$$\mu_{t|N} = \mu_{t|t} + F_t(\mu_{t+1|N} - \mu_{t+1|t}) \quad (5.39)$$

$$\Sigma_{t|N} = \Sigma_{t|t} - F_t(\Sigma_{t+1|t} - \Sigma_{t+1|N})F_t^T \quad (5.40)$$

$$F_t = \Sigma_{t|t}A^T\Sigma_{t+1|t}^{-1} \quad (5.41)$$

5.5 Rozšířený Kalmanův filtr

Myšlenka rozšířeného Kalmanova filtru (Extended Kalman Filter) spočívá v tom, že nelineární dynamický systém v každém kroku přibližně nahradíme *lineárním* dynamickým systémem a na tento pak aplikujeme obyčejný Kalmanův filtr. Výsledky věty 5.10 však nelze použít přímo, protože linearizace vede na poněkud jiný typ DS.

Lemma 5.14. Bud' dán lineární diskretní stochastický systém, popsany rovnicemi

$$x_{t+1} = A_t x_t + b_t + v_t \quad (5.42)$$

$$y_t = C_t x_t + d_t + w_t \quad (5.43)$$

splňující podmínky (i)–(v) z věty 5.10. Pak pro apriorní estimátor $x_{t|t-1} \sim N(\mu_{t|t-1}, \Sigma_{t|t-1})$ a posteriorní estimátor $x_{t|t} \sim N(\mu_{t|t}, \Sigma_{t|t})$ platí

$$\mu_{t|t} = \mu_{t|t-1} + K_t(y_t - C_t \mu_{t|t-1} - d_t) \quad (5.44)$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - K_t C_t \Sigma_{t|t-1} \quad (5.45)$$

$$\mu_{t+1|t} = A_t \mu_{t|t} + b_t \quad (5.46)$$

$$\Sigma_{t+1|t} = A_t \Sigma_{t|t} A_t^T + \Sigma_v \quad (5.47)$$

$$K_t = \Sigma_{t|t-1} C_t^T (C_t \Sigma_{t|t-1} C_t^T + \Sigma_w)^{-1} \quad (5.48)$$

Důkaz. Důkaz je pouze drobnou modifikací důkazu věty 5.10. \square

Vezměme obecný nelineární dynamický systém ve stavovém tvaru

$$\begin{aligned}x_{t+1} &= f(x_t, u_t) + v_t \\ y_t &= g(x_t, u_t) + w_t\end{aligned}$$

Linearizujeme-li funkce $f(x, u)$ a $g(x, u)$ proměnné x podle Taylorova rozvoje v bodě $x = \xi$, dostaneme

$$f(x, u) = f(\xi, u) + \frac{\partial f}{\partial x}(\xi)(x - \xi) + O(x^2) \quad (5.49)$$

$$g(x, u) = g(\xi, u) + \frac{\partial g}{\partial x}(\xi)(x - \xi) + O(x^2) \quad (5.50)$$

Aproximace bude tím lepší, čím menší bude rozdíl $x - \xi$. Zanedbáme-li členy $O(x^2)$, můžeme původní systém přibližně nahradit lineárním systémem

$$x_{t+1} = f(\xi, u_t) + A_t(x_t - \xi) + v_t \quad A_t = \frac{\partial f}{\partial x}(\xi) \quad (5.51)$$

$$y_t = g(\xi, u_t) + C_t(x_t - \xi) + w_t \quad C_t = \frac{\partial g}{\partial x}(\xi) \quad (5.52)$$

Předpokládejme, že máme k dispozici estimátor $x_{t|t-1}$ prostřednictvím jeho momentů $\mu_{t|t-1}, \Sigma_{t|t-1}$. Linearizujeme výstupní rovnici v okolí bodu $\mu_{t|t-1}$:

$$y_t = C_t x_t + (g(\mu_{t|t-1}, u_t) - C_t \mu_{t|t-1}) + w_t \quad \text{kde } C_t = \frac{\partial g}{\partial x}(\mu_{t|t-1}) \quad (5.53)$$

Pak podle lemmatu 5.14 je datový krok dán rovnicí

$$\mu_{t|t} = \mu_{t|t-1} + K_t(y_t - g(\mu_{t|t-1}, u_t)) \quad (5.54)$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - K_t C_t \Sigma_{t|t-1} \quad (5.55)$$

kde jsme použili *Kalmanovo zesílení*

$$K_t = \Sigma_{t|t-1} C_t^T (C_t \Sigma_{t|t-1} C_t^T + \Sigma_w)^{-1}$$

Nyní máme k dispozici aposteriorní odhad $x_{t|t}$. Linearizujeme-li stavovou rovnici v okolí bodu $\mu_{t|t}$, obdržíme lineární stavovou rovnici

$$x_{t+1} = A_t x_t + (f(\mu_{t|t}, u_t) - A_t \mu_{t|t}) + v_t \quad \text{kde } A_t = \frac{\partial f}{\partial x}(\mu_{t|t}) \quad (5.56)$$

Podle lemmatu 5.14 napíšeme predikční krok

$$\mu_{t+1|t} = f(\mu_{t|t}, u_t) \quad (5.57)$$

$$\Sigma_{t+1|t} = A_t \Sigma_{t|t} A_t^T + \Sigma_v \quad (5.58)$$

Někdy se odhady vylepšují ještě smoothingem podle věty 5.13, takže algoritmus rozšířeného Kalmanova filtru se pak rozpadne na dopředný běh (filtr) a zpětný běh (smoother). Dosažené výsledky přehledně zformulujeme v definici

Definice 5.15. Pro nelineární dynamický systém

$$\begin{aligned}x_{t+1} &= f(x_t, u_t) + v_t \\ y_t &= g(x_t, u_t) + w_t\end{aligned}$$

s počáteční podmínkou $x_0 \sim N(\mu_0, \Sigma_0)$ definujeme *Rozšířený Kalmanův filtr* jako algoritmus výpočtu estimátorů

$$\begin{aligned}x_{t|t-1} &\sim N(\mu_{t|t-1}, \Sigma_{t|t-1}) \\ x_{t|t} &\sim N(\mu_{t|t}, \Sigma_{t|t}) \\ x_{t|N} &\sim N(\mu_{t|N}, \Sigma_{t|N})\end{aligned}$$

podle vzorců

$$\mu_{t|t} = \mu_{t|t-1} + K_t(y_t - g(\mu_{t|t-1}, u_t)) \quad (5.59)$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - K_t C_t \Sigma_{t|t-1} \quad (5.60)$$

$$\mu_{t+1|t} = f(\mu_{t|t}, u_t) \quad (5.61)$$

$$\Sigma_{t+1|t} = A_t \Sigma_{t|t} A_t^T + \Sigma_v \quad (5.62)$$

$$K_t = \Sigma_{t|t-1} C_t^T (C_t \Sigma_{t|t-1} C_t^T + \Sigma_w)^{-1} \quad (5.63)$$

$$\mu_{t|N} = \mu_{t|t} + F_t(\mu_{t+1|N} - \mu_{t+1|t}) \quad (5.64)$$

$$\Sigma_{t|N} = \Sigma_{t|t} - F_t(\Sigma_{t+1|t} - \Sigma_{t+1|N})F_t^T \quad (5.65)$$

$$F_t = \Sigma_{t|t} A^T \Sigma_{t+1|t}^{-1} \quad (5.66)$$

kde matice A_t, C_t jsou Jakobiány

$$A_t = \frac{\partial f}{\partial x}(\mu_{t|t}) \quad C_t = \frac{\partial g}{\partial x}(\mu_{t|t-1}) \quad (5.67)$$

5.6 Odhad nelineárních systémů metodou Monte Carlo

Rozšířený Kalmanův filtr je průmyslovým standartem pro nelineární rekurzivní estimaci. V roce 1992 se v článku [?] objevil zcela nový algoritmus, založený na aplikaci známé metody Monte Carlo při bayesovské estimaci. Metoda *Monte Carlo* je založena na tom, že informaci o rozložení náhodné veličiny nese *náhodný výběr* z tohoto rozložení. Čím je náhodný výběr delší, tím je informace přesnější. V předchozích metodách informaci o pravděpodobnostním rozložení estimátoru obsahovala buďto hustota pravděpodobnosti, nebo první dva momenty v gaussovském případě. Pro náhodný vektor $x \in \mathcal{L}n_x$ máme náhodný výběr $S_n = (x_1, \dots, x_n)$ délky n , obsahující vzorky $x_i \in \mathbb{R}^{n_x}$. Empirická hustota pravděpodobnosti, získaná z náhodného výběru, je aproximací skutečné hustoty pravděpodobnosti náhodného vektoru x . Empirickou hustotu pravděpodobnosti můžeme napsat jako

$$p_n(x) = \frac{1}{n} \sum_{i=1}^n \delta(x - x_i) \quad (5.68)$$

kde $\delta(x) : \mathbb{R}^{n_x} \rightarrow \mathbb{R}$ je tzv. *Diracova funkce*, definovaná limitou

$$\delta(x) = \lim_{h \rightarrow 0} \delta_h(x), \quad \delta_h(x) = \begin{cases} \frac{1}{h} & \text{pro } 0 < x < h \\ 0 & \text{jinak} \end{cases} \quad (5.69)$$

V článku [?] je metoda Monte Carlo ještě vylepšena tím, že se jednotlivým vzorkům x_i přiřadí váha $w_i \geq 0$ tak, aby součet jednotlivých vah dával jedničku. Empirická hustota je potom

$$p_n(x) = \sum_{i=1}^n w_i \delta(x - x_i) \quad (5.70)$$

Rozložení náhodného vektoru x je tedy popsáno náhodným výběrem S_n a vektorem vah $w_n = (w_1, \dots, w_n)$. Připomeňme obecné rekurzivní vztahy pro nelineární estimaci:

$$p(x_t|D_t) = \frac{p(y_t|x_t, u_t)}{p(y_t|u_t, D_{t-1})} p(x_t|D_{t-1}) \quad (5.71)$$

$$p(x_{t+1}|D_t) = \int p(x_{t+1}|x_t, u_t) p(x_t|D_t) dx_t \quad (5.72)$$

$$p(x_t|D_N) = p(x_t|D_t) \int p(x_{t+1}|x_t, u_t) \frac{p(x_{t+1}|D_N)}{p(x_{t+1}|D_t)} dx_{t+1} \quad (5.73)$$

Pracujeme-li s empirickými hustotami, pak aplikace těchto rekurzivních vztahů vyústí v algoritmus přepočtu vah w_i a vzorků x_i , který má název *Weighted Bootstrap Algorithm*. Vysvětlíme si ve stručnosti princip tohoto algoritmu:

Buď dána apriorní hustota

$$p_n(x_{t|t-1}) = \sum_{i=1}^n w_i(t|t-1) \delta(x - x_i(t|t-1))$$

V datovém kroku hledáme aposteriorní hustotu $p_n(x_{t|t})$. Položíme $x_i(t|t) = x_i(t|t-1)$ a váhy $w_i(t|t)$ získáme pomocí (5.71). Člen $p(y_t|u_t, D_{t-1})$ funguje ve vzorci (5.71) jako normalizační konstanta. Když tento člen vypustíme, obdržíme váhy

$$\bar{w}_i(t|t) = p(y(t)|x_i(t|t-1), u(t)) w_i(t|t-1) \quad (5.74)$$

z nichž normalizací dostaneme výsledné váhy

$$w_i(t|t) = \frac{\bar{w}_i(t|t)}{\sum_{j=1}^n \bar{w}_j(t|t)} \quad (5.75)$$

Tím je zakončen datový krok algoritmu. Při predikčním kroku nepoužijeme vztah (5.72), ale pro modelový krok využijeme vztahu

$$x_{t+1|t} = f(x_{t|t}, u(t)) + v(t)$$

který aplikujeme na vzorky $x_i(t|t)$. Dostaneme pak

$$x_i(t+1|t) = f(x_i(t|t), u(t)) + v_i(t) \quad (5.76)$$

kde $v_i(t)$ jsou náhodně generované vzorky šumu procesu v_t . Váhy se při predikčním kroku nemění, tj. $w_i(t+1|t) = w_i(t|t)$. Takto jsme dostali predikci $x_{t+1|t}$. Konstrukce smoothingu má stejnou filozofii jako datový krok. Vzorky zůstávají stejné $x_i(t|N) = x_i(t|t)$ a upravují se jen váhy. Vypuštěním členu $p(x_{t+1}|D_t)$ ze vzorce (5.73) dostaneme zmenšené váhy $\bar{w}_i(t|N)$ podle vzorce

$$\bar{w}_i(t|N) = w_i(t|t) \sum_{j=1}^n w_j(t+1|N) p(x_j(t+1|N)|x_i(t|N), u(t)) \quad (5.77)$$

a váhy $w_i(t|N)$ dostaneme normalizací

$$w_i(t|N) = \frac{\bar{w}_i(t|t)}{\sum_{j=1}^n \bar{w}_j(t|t)} \quad (5.78)$$

V datovém kroku jsme potřebovali znalost výstupní rovnice skrze hustotu $p(y|x, u)$, v predikčním kroku jsme zase použili funkci f . Při smoothingu se neobejdeme bez podmíněné hustoty $p(x_{t+1}|x_t, u_t)$, která je určena stavovou rovnicí.

Poslední nevyřešenou otázkou je zahájení algoritmu. Předpokládáme, že rozložení stavu x_0 je známé. Pak v souladu s (5.19) položíme $x_{0|-1} = x_0$. Potřebujeme vzorky $x_i(0|-1)$ a váhy $w_i(0|-1)$. Nejčastěji se vzorky získávají generátorem náhodných čísel s požadovaným rozložením. Váhy jsou potom rovnoměrné, tedy $w_i(0|-1) = 1/n$. Jiná možnost je zvolit rovnoměrnou síť bodů $x_i(0|-1)$ ve vhodně zvolené oblasti a váhy volit $w_i(0|-1) = p_0(x_i(0|-1))$ a potom je normalizovat.

Dosažené výsledky shrneme v definici.

Definice 5.16 (Weighted Bootstrap Algorithm). Buď dán dynamický systém

$$\begin{aligned} x_{t+1} &= f(x_t, u_t, v_t) \\ y_t &= g(x_t, u_t, w_t) \end{aligned}$$

s počátečním stavem x_0 , který má empirickou hustotu

$$p_n(x) = \sum_{i=1}^n w_i(0|-1) \delta(x - x_i(0|-1))$$

Buď dána data D_t (viz (5.13)). Pak odhady $x_{t|t-1}, x_{t|t}, x_{t|N}$ s empirickými hustotami

$$\begin{aligned} p_n(x_{t|t-1}) &= \sum_{i=1}^n w_i(t|t-1) \delta(x - x_i(t|t-1)) \\ p_n(x_{t|t}) &= \sum_{i=1}^n w_i(t|t) \delta(x - x_i(t|t)) \\ p_n(x_{t|N}) &= \sum_{i=1}^n w_i(t|N) \delta(x - x_i(t|N)) \end{aligned}$$

jsou spočteny algoritmem *Weighted Bootstrap*, právě když platí

$$x_i(t|t) = x_i(t|t-1) \quad (5.79)$$

$$\bar{w}_i(t|t) = p(y(t)|x_i(t|t-1), u(t)) w_i(t|t-1) \quad (5.80)$$

$$w_i(t|t) = \frac{\bar{w}_i(t|t)}{\sum_{j=1}^n \bar{w}_j(t|t)} \quad (5.81)$$

$$x_i(t+1|t) = f(x_i(t|t), u(t)) + v_i(t) \quad (5.82)$$

$$w_i(t+1|t) = w_i(t|t) \quad (5.83)$$

$$x_i(t|N) = x_i(t|t) \quad (5.84)$$

$$\bar{w}_i(t|N) = w_i(t|t) \sum_{j=1}^n w_j(t+1|N) p(x_j(t+1|N)|x_i(t|N), u(t)) \quad (5.85)$$

$$w_i(t|N) = \frac{\bar{w}_i(t|t)}{\sum_{j=1}^n \bar{w}_j(t|t)} \quad (5.86)$$

5.7 Shrnutí

Optimální estimátor náhodného vektoru X na základě dat Z je podmíněná náhodná veličina $X|Z$ případně podmíněná střední hodnota $E(X|Z)$. Je-li sdružený vektor (X, Z) rozložen normálně, pak optimální estimátor je lineární. Lineární estimátor se dá použít i pro jiná rozdělení, je však obecně méně přesný než estimátor $X|Z$.

Při odhadu stavů dynamického systému (5.12) používáme apriorní odhad $x_{t|t-1}$, aposteriorní odhad $x_{t|t}$ a smoothing $x_{t|N}$. První dva estimátory se vypočtou algoritmem dopředné rekurze pomocí vzorců (5.18) a (5.17). Smoothing probíhá zpětnou rekurzí s využitím vzorce (5.35).

Pracujeme-li s lineárním gaussovským systémem (5.20a), pak odhady jsou normální a stačí spočítat jejich střední hodnoty a rozptyl. Dopředný algoritmus pro $x_{t|t-1}$ a $x_{t|t}$ se jmenuje *Kalmanův filtr* a zpětný algoritmus pro $x_{t|N}$ se jmenuje *Rauch-Tung-Striebel smoother*.

Pro nelineární systémy s diferencovatelnými funkcemi f, g se může použít rozšířený Kalmanův filtr. Ten v každém kroku lokálně linearizuje funkce f, g a na takto linearizovaný systém použije Kalmanových vzorců. Výsledné estimátory pak sice nejsou optimální, zato jejich výpočet je poměrně efektivní.

Přesný nelineární estimátor je dán vzorcí (5.18), (5.17), (5.35). Jejich numerický výpočet je však se současnými prostředky nedosažitelný. Proto saháme k přibližným metodám. V *Bootstrap* algoritmu je rozvedena myšlenka reprezentace hustoty pomocí váženého náhodného výběru, tj. metoda Monte Carlo.

5.8 Příklady k procvičení

Příklad 5.17. Napište v systému Matlab funkci `kf`, která provede algoritmus Kalmanova filtru na časovém horizontu $0 \dots N$ pro systém (5.20a). Funkce bude mít prototyp

`[Mp, Sp, Mf, Sf] = kf(A, B, C, D, Sv, Sw, m0, S0, Y, U)`, kde uvedené matice mají následující význam:

Mp Matice typu $n_x \times N + 1$, obsahující ve svých sloupcích predikce $\mu_{0|-1}, \dots, \mu_{N|N-1}$.

Sp Matice typu $n_x^2 \times N + 1$, obsahující ve sloupcích varianční matice $\Sigma_{t|t-1}$.

Mf Matice typu $n_x \times N + 1$, obsahující ve svých sloupcích filtrované střední hodnoty $\mu_{0|0}, \dots, \mu_{N|N}$.

Sf Matice typu $n_x^2 \times N + 1$, obsahující ve sloupcích varianční matice $\Sigma_{t|t}$.

A Matice A typu $n_x \times n_x$, definující stavovou rovnici (5.20a).

B Matice B typu $n_x \times n_u$, definující stavovou rovnici (5.20a).

C Matice C typu $n_y \times n_x$, definující výstupní rovnici (5.20b).

D Matice D typu $n_y \times n_u$, definující výstupní rovnici (5.20b).

Sv Varianční matice stavového šumu Σ_v typu $n_x \times n_x$, symetrická a pozitivně semidefinitní.

Sw Varianční matice výstupního šumu Σ_w typu $n_y \times n_y$, symetrická a pozitivně semidefinitní.

m0 Střední hodnota počátečního stavu x_0 typu $n_x \times 1$.

S0 Variance počátečního stavu x_0 typu $n_x \times n_x$.

Y Datová matice výstupů typu $n_y \times N + 1$, obsahující ve sloupcích t výstupy y_t .

U Datová matice vstupů typu $n_u \times N + 1$, obsahující ve sloupcích t vstupy u_t .

Kapitola 6

IDENTIFIKACE DYNAMICKÝCH SYSTÉMŮ

6.1 Úvod

Uvažujme diskrétní dynamický systém závislý na parametru θ :

$$x_{t+1} = f(x_t, u_t, v_t, \theta) = f_\theta(x_t, u_t, v_t) \quad (6.1a)$$

$$y_t = g(x_t, u_t, w_t, \theta) = g_\theta(x_t, u_t, w_t) \quad (6.1b)$$

Parametr $\theta \in \Theta$ je veličina, na které jsou závislé vlastnosti systému (6.1). Množina Θ se pak nazývá *parametrický prostor*.

Příklad 6.1. Mějme dynamický systém, o němž víme, že je lineární:

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + v_t \\ y_t &= Cx_t + Du_t + w_t, \end{aligned}$$

ale neznáme matice A, B, C, D . Tento systém je závislý na parametru $\theta = [A \ B \ C \ D]$ a parametrický prostor Θ je potom $\Theta = \mathbb{R}^{n_x \times n_x} \times \mathbb{R}^{n_x \times n_u} \times \mathbb{R}^{n_y \times n_x} \times \mathbb{R}^{n_y \times n_u}$ s dimenzí $p = n_x n_x + n_x n_u + n_y n_x + n_y n_u$.

Většinou se stává, že u dynamického systému neznáme konkrétně funkce f_θ, g_θ , ale známe strukturu systému f, g . Struktura systému spolu s vektorem θ jednoznačně určuje vlastnosti systému. Chybějí-li tedy informace o hodnotě θ , je žádoucí je doplnit. Jediným dostupným zdrojem informací přitom bývá chování systému v minulosti, tedy historie vstupů a výstupů. Techniky, které se zabývají odhadem parametrů systému na základě známé historie se nazývají techniky *identifikace*. Označíme-li $\hat{\theta}$ odhad parametru θ a $D_n = \{u_0, y_0, \dots, u_n, y_n\}$ historická data, pak identifikace systému (6.1) je úloha nalézt při známé struktuře f, g odhad $\hat{\theta} = F(D_n)$ parametru θ tak, aby se co nejméně lišil od skutečného parametru θ .

6.2 Klasický přístup

Při klasickém přístupu se neznámý parametr θ uvažuje jako reálný vektor dimenze p . Parametrický prostor Θ je potom roven $\Theta \subseteq \mathbb{R}^p$.

Při identifikaci je cílem dosáhnout maximální shody mezi odhadem $\hat{\theta}$ a skutečnou hodnotou parametru θ . Protože však rozdíl $\theta - \hat{\theta}$ není přímo měřitelný, používají se různá kritéria, která jsou vodítkem při určení $\hat{\theta}$.

6.2.1 Metoda nejmenších čtverců (LS)

Metoda nejmenších čtverců (least squares method, LS) říká, že odhad parametru je tím lepší, čím lepší je shoda modelu se skutečností. Označíme-li výstup modelu jako $\hat{y}_t = \hat{y}_t(\theta)$, pak shodu modelu se skutečností měříme podle kritéria nejmenších čtverců jako

$$J(\theta) = \sum_{t=0}^n |y_t - \hat{y}_t(\theta)|^2 \quad (6.2)$$

Namísto neměřitelného rozdílu $\theta - \hat{\theta}$ tak pracujeme s měřitelným rozdílem $y_t - \hat{y}_t$. Optimální odhad $\hat{\theta}$ bude takový vektor, pro který bude kritérium J minimální, tj.

$$J(\hat{\theta}) = \min_{\theta \in \Theta} J(\theta) \quad (6.3)$$

Funkce J je zobrazení $J : \mathbb{R}^p \rightarrow \mathbb{R}$. Tuto funkci minimalizujeme metodami matematického programování.

Příklad 6.2 (Lineární systém). Pro jednoduchost předpokládejme nejprve lineární systém s pozorovatelnými stavy, neboli systém

$$x_{t+1} = Ax_t + Bu_t + v_t \quad (6.4a)$$

$$y_t = x_t \quad (6.4b)$$

Máme identifikovat matice A, B na základě historie systému, zachycené v datech

$$D_n = \{u_0, y_0, \dots, u_n, y_n\} = \{u_0, x_0, \dots, u_n, x_n\}$$

Kritériální funkce J bude nyní vážit rozdíl mezi stavem x_{t+1} a jeho θ -předpovědí $Ax_t + Bu_t$:

$$J(\theta) = J(A, B) = \sum_{t=0}^{n-1} |x_{t+1} - Ax_t - Bu_t|^2$$

Označíme-li $e_t = x_{t+1} - Ax_t - Bu_t$ a matice e_t, x_{t+1}, A, B rozepíšeme po řádcích

$$e_t = \begin{bmatrix} e_{1,t} \\ \vdots \\ e_{n_x,t} \end{bmatrix} \quad x_{t+1} = \begin{bmatrix} x_{1,t+1} \\ \vdots \\ x_{n_x,t+1} \end{bmatrix} \quad A = \begin{bmatrix} a_1 \\ \vdots \\ a_{n_x} \end{bmatrix} \quad B = \begin{bmatrix} b_1 \\ \vdots \\ b_{n_x} \end{bmatrix}$$

pak platí

$$J(A, B) = \sum_{t=0}^{n-1} |e_t|^2 = \sum_{t=0}^{n-1} \sum_{j=1}^{n_x} e_{jt}^2 = \sum_{j=1}^{n_x} \sum_{t=0}^{n-1} (x_{j,t+1} - a_j x_t - b_j u_t)^2 = \sum_{j=1}^{n_x} J_j(a_j, b_j)$$

kde jsme položili

$$J_j(a_j, b_j) = (x_{j,t+1} - a_j x_t - b_j u_t)^2$$

Úloha minimalizovat $J(A, B)$ se nám nyní rozpadla na n_x úloh minimalizovat $J_j(a_j, b_j)$ pro $j \in \{1, \dots, n_x\}$. Tyto dílčí úlohy jsou úlohami lineární regrese (viz B.6) ve tvaru

$$x_{j,t+1} = a_j x_t + b_j u_t \quad (6.5)$$

při označení

$$z_j = \begin{bmatrix} x_{j,1} \\ \vdots \\ x_{j,n} \end{bmatrix} \quad \beta_j = [a_j \quad b_j]^T \quad U = \begin{bmatrix} x_0^T & u_0^T \\ \vdots & \vdots \\ x_{n-1}^T & u_{n-1}^T \end{bmatrix}$$

můžeme tyto úlohy zapsat maticově

$$z_j = U \beta_j \quad (6.6)$$

Podle (B.18) je optimálním odhadem vektoru β_j vektor $\hat{\beta}_j$, daný vztahem

$$\hat{\beta}_j = (U^T U)^{-1} U^T z_j \quad (6.7)$$

Odhad $\hat{\theta} = [\hat{A} \quad \hat{B}]$ složíme s odhadů $\hat{\beta}_j$:

$$\begin{aligned} [\hat{A} \quad \hat{B}] &= \begin{bmatrix} \hat{\beta}_1^T \\ \vdots \\ \hat{\beta}_{n_x}^T \end{bmatrix} = [\hat{\beta}_1 \quad \dots \quad \hat{\beta}_{n_x}]^T = [(U^T U)^{-1} U^T z_1 \quad \dots \quad (U^T U)^{-1} U^T z_{n_x}]^T = \\ &= \{(U^T U)^{-1} U^T [z_1 \quad \dots \quad z_{n_x}]\}^T = [z_1 \quad \dots \quad z_{n_x}]^T U (U^T U)^{-1} = \\ &= \begin{bmatrix} x_{1,1} & \dots & x_{n_x,1} \\ \vdots & & \vdots \\ x_{1,n} & \dots & x_{n_x,n} \end{bmatrix}^T U (U^T U)^{-1} = XU (U^T U)^{-1} \end{aligned}$$

Celkem tedy máme

$$[\hat{A} \quad \hat{B}] = XU (U^T U)^{-1}, \quad (6.8)$$

kde jsme položili

$$X = \begin{bmatrix} x_1^T \\ \vdots \\ x_n^T \end{bmatrix} \quad U = \begin{bmatrix} x_0^T & u_0^T \\ \vdots & \vdots \\ x_{n-1}^T & u_{n-1}^T \end{bmatrix}$$

Příklad 6.3 (Nelineární systém). Pro jednoduchost předpokládejme opět systém s pozorovatelnými stavy, s konstantními parametry a s aditivním šumem

$$x_{t+1} = f(x_t, u_t, \theta) + v_t \quad (6.9a)$$

$$y_t = x_t \quad (6.9b)$$

Na základě známých dat D_n máme odhadnout parametr θ tak, aby součet čtverců

$$J(\theta) = \sum_{t=0}^{n-1} |x_{t+1} - f(x_t, u_t, \theta)|^2 \quad (6.10)$$

byl minimální na \mathbb{R}^p . Jedná se o hledání volného extrému funkce p proměnných. Je-li funkce f třídy C^1 , pak z analýzy víme, že je-li bod $\hat{\theta}$ bodem minima funkce J , pak platí, že diferenciál v tomto bodě je nulový a tedy všechny parciální derivace jsou nulové. Spočteme parciální derivace:

$$\frac{\partial J}{\partial \theta_i} = -2 \sum_{t=0}^{n-1} \sum_{j=1}^{n_x} \frac{\partial f_j}{\partial \theta_i}(x_t, u_t, \theta) (x_{j,t+1} - f_j(x_t, u_t, \theta)) \quad (6.11)$$

Označíme-li

$$\begin{aligned} A_t(\theta) &= \left[\frac{\partial f_j}{\partial \theta_i}(x_t, u_t, \theta) \right] \\ b_t(\theta) &= x_{j,t+1} - f_j(x_t, u_t, \theta) \\ F(\theta) &= \sum_{t=0}^{n-1} A_t(\theta) b_t(\theta) \end{aligned}$$

pak můžeme podmínku $\frac{\partial J}{\partial \theta} = 0$ zapsat jako soustavu nelineárních rovnic řádu p

$$F(\theta) = 0 \quad (6.12)$$

Tuto soustavu pak můžeme řešit některou ze známých numerických metod, jako je Newtonova metoda, Gauss-Seidelova metoda, apod. (viz [?]). V Matlabu je k dispozici funkce `fsolve`, do které za funkci `fun` dosadíme funkci F .

Jiný způsob je použití metod numerické optimalizace přímo pro minimalizaci kritéria (6.10). Známé jsou metody nejrychlejšího spádu, metoda konjugovaných gradientů, metoda proměnné metriky, Newtonova metoda (viz [?]). V Matlabu minimalizaci provádí funkce `fmins`, do které jako argument `F` dosadíme funkci J podle (6.10).

6.2.2 Metoda maximální věrohodnosti

Není-li kritérium nejmenších čtverců $J(\theta)$ funkce kvadratická v θ , je mnohdy výhodné místo kritéria nejmenších čtverců uvažovat kritérium maximální věrohodnosti.

6.3 Bayesův přístup

Bayesova metoda předpokládá, že parametr θ je náhodný vektor, který se vyvíjí v čase. Jinými slovy posloupnost $\{\theta_t\}_{t=0}^{n-1}$ je náhodný proces. Takový proces je popsán například sdruženou hustotou $p(\theta_0, \dots, \theta_{n-1})$. Tento náhodný proces je před identifikací již *známý* (je známa hustota pravděpodobnosti p).

Podoba náhodného procesu $\{\theta_t\}_{t=0}^{n-1}$ patří mezi apriorní informace, které ovlivní konečný odhad. Při konstrukci modelu vývoje parametru θ máme možnost ovlivnit závislost mezi parametry v různých obdobích i závislost mezi jednotlivými složkami parametru θ . Můžeme například předepsat, že parametry θ_t a θ_{t+1} budou silně korelovány, zatímco složky $\theta_{1,t}, \dots, \theta_{p,t}$

budou nezávislé a podobně. U Bayesovy metody se předpokládá, že náhodný proces $\{\theta_t\}_{t=0}^{n-1}$ je zadán jako dynamický systém (ve stavovém tvaru)

$$\theta_{t+1} = h(\theta_t, \varepsilon_t) \quad (6.13)$$

s počáteční podmínkou $\theta_0 = \zeta$, kde $\zeta \in \mathcal{L}p$ je náhodný vektor se známým rozložením a ε_t je stochastická složka náhodného procesu (6.13) (bílý šum).

Mezi nejčastější (a také nejjednodušší) modely vývoje parametrů patří model náhodné procházky (random walk) ve tvaru

$$\theta_{t+1} = \theta_t + \varepsilon_t \quad (6.14)$$

Model náhodné procházky vykazuje některé charakteristické vlastnosti:

- (i) $E(\theta_t) = E(\theta_0)$ pro $t \in \{0, \dots, n-1\}$.
- (ii) $\text{var}(\theta_t) = \text{var}(\theta_0) + \sum_{j=0}^{t-1} \text{var}(\varepsilon_j)$ pro $t \in \{0, \dots, n-1\}$.
- (iii) $\text{cov}(\theta_t, \theta_{t+s}) = \text{var}(\theta_t)$ pro $t \in \{0, \dots, n-1\}$, $s \in \{0, \dots, n-t-1\}$.

Předpokládejme, že je znám model vývoje parametrů (6.13) a je dán parametrický dynamický systém (6.1). Máme tedy systém ve tvaru

$$x_{t+1} = f(x_t, u_t, v_t, \theta_t) \quad (6.15a)$$

$$\theta_{t+1} = h(\theta_t, \varepsilon_t) \quad (6.15b)$$

$$y_t = g(x_t, u_t, w_t, \theta_t) \quad (6.15c)$$

Definujme náhodné vektory z_t, e_t předpisem

$$z_t = \begin{bmatrix} x_t \\ \theta_t \end{bmatrix} \quad e_t = \begin{bmatrix} v_t \\ \varepsilon_t \end{bmatrix} \quad (6.16)$$

Definujme dále funkci $F : \mathbb{R}^{n_x+p} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_v} \rightarrow \mathbb{R}^{n_x+p}$ následujícím způsobem:

$$F(z_t, u_t, e_t) = \begin{bmatrix} f(x_t, u_t, v_t, \theta_t) \\ h(\theta_t, \varepsilon_t) \end{bmatrix} \quad (6.17)$$

a nakonec funkci $G : \mathbb{R}^{n_x+p} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_w} \rightarrow \mathbb{R}^{n_y}$ předpisem

$$G(z_t, u_t, w_t) = g(x_t, u_t, w_t, \theta_t) \quad (6.18)$$

Potom dynamický systém (6.15) napíšeme pomocí (6.16), (6.17), (6.18) jako

$$z_{t+1} = F(z_t, u_t, e_t) \quad (6.19a)$$

$$y_t = G(z_t, u_t, w_t) \quad (6.19b)$$

Systém (6.19) je neparametrický dynamický systém ve stavovém tvaru.

Úloha optimálního odhadu stavu z_t systému (6.19) je ekvivalentní s úlohou optimálního odhadu stavu x_t systému (6.15) a současně optimálního odhadu parametrů θ_t systému (6.15).

Tímto jsme převedli odhad parametrů na odhad stavů, což je Bayesova metoda odhadu parametrů dynamického systému. Algoritmicky se tato úloha dá realizovat například použitím Rozšířeného Kalmanova filtru.

Příklad 6.4. Vezměme nelineární parametrický dynamický systém s pozorovatelnými stavy a se známým modelem vývoje parametrů

$$\begin{aligned}x_{t+1} &= f(x_t, u_t, v_t, \theta_t) \\ \theta_{t+1} &= h(\theta_t, \varepsilon_t) \\ y_t &= x_t\end{aligned}$$

Pro tento systém nechť známe data $D_n = \{u_0, y_0, \dots, u_n, y_n\}$. Odhadněme parametry θ_t Bayesovou metodou.

Zavedme výstupní proměnnou $\eta_t = x_{t+1}$ a stav $z_t = \theta_t$. Dále definujme funkci g jako

$$g(z_t, u_t, v_t, t) = f(x_t, u_t, v_t, z_t)$$

Funkce g je jednoznačně určena funkcí f a daty D_n . Nyní má systém podobu

$$\begin{aligned}z_{t+1} &= h(z_t, \varepsilon_t) \\ \eta_t &= g(z_t, u_t, v_t, t)\end{aligned}$$

Jedná se tedy o neautonomní systém s nepozorovatelnými stavy z_t s šumem procesu ε_t a výstupním šumem v_t , který byl původně šumem procesu. Pro odhad stavu z_t (parametru θ_t) máme k dispozici vstupy u_0, \dots, u_n a výstupy $\eta_0, \dots, \eta_{n-1}$. Časový horizont je tedy zkrácený o jednu periodu.

Příklad 6.5. Buď dán skalární lineární systém

$$\begin{aligned}x(t+1) &= a(t)x(t) + b(t)u(t) + v(t) \\ y(t) &= c(t)x(t) + d(t)u(t) + w(t)\end{aligned}$$

kde $v(t) \sim N(0, \sigma_v^2)$, $w(t) \sim N(0, \sigma_w^2)$ s časově proměnnými parametry

$$\theta(t) = \begin{bmatrix} a(t) \\ b(t) \\ c(t) \\ d(t) \end{bmatrix}$$

Předpokládáme model vývoje parametrů ve tvaru náhodné procházky

$$\theta_{t+1} = \theta(t) + \varepsilon(t)$$

kde $\varepsilon(t) \sim N(0, \sigma_\varepsilon^2)$ je bílý šum. Známe počáteční podmínky

$$\begin{aligned}x_0 &\sim N(\mu_0, \sigma_0^2) \\ \theta_0 &\sim N(\nu_0, \tau_0^2)\end{aligned}$$

Sestrojme algoritmus pro optimální odhad parametru $\theta(t)$ s použitím Bayesovy metody a Rozšířeného Kalmanova filtru. Definujme si stavový vektor

$$z(t) = \begin{bmatrix} x(t) \\ a(t) \\ b(t) \\ c(t) \\ d(t) \end{bmatrix}$$

a vektor šumu

$$e(t) = \begin{bmatrix} v(t) \\ \varepsilon_1(t) \\ \varepsilon_2(t) \\ \varepsilon_3(t) \\ \varepsilon_4(t) \end{bmatrix}$$

Dále definujeme funkci $F : \mathbb{R}^5 \times \mathbb{R} \times \mathbb{R}^5$ vzorcem

$$F(z(t), u(t), e(t)) = \begin{bmatrix} z_1(t)z_2(t) + z_3(t)u(t) + e_1(t) \\ z_2(t) + e_2(t) \\ z_3(t) + e_3(t) \\ z_4(t) + e_4(t) \\ z_5(t) + e_5(t) \end{bmatrix}$$

a funkci $G : \mathbb{R}^5 \times \mathbb{R} \times \mathbb{R}$ vzorcem

$$G(z(t), u(t), w(t)) = z_1(t)z_4(t) + z_5(t)u(t) + w(t)$$

Dynamický systém

$$\begin{aligned} z(t+1) &= F(z(t), u(t), e(t)) \\ y(t) &= G(z(t), u(t), w(t)) \end{aligned}$$

je nyní přepisem původního lineárního dynamického systému. V rozepsané podobě je to systém

$$\begin{aligned} z_1(t+1) &= z_1(t)z_2(t) + z_3(t)u(t) + e_1(t) \\ z_2(t+1) &= z_2(t) + e_2(t) \\ z_3(t+1) &= z_3(t) + e_3(t) \\ z_4(t+1) &= z_4(t) + e_4(t) \\ z_5(t+1) &= z_5(t) + e_5(t) \\ y(t) &= z_1(t)z_4(t) + z_5(t)u(t) + w(t), \end{aligned}$$

který zřejmě není lineární. Pro optimální odhad stavu $z(t)$ se tedy hodí Rozšířený Kalmanův Filtr. Matice $A(t)$, $C(t)$ jsou:

$$A(t) = \begin{bmatrix} z_2(t) & z_1(t) & u(t) & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$C(t) = [z_4(t) \quad 0 \quad 0 \quad z_1(t) \quad u(t)]$$

Kapitola 7

TEORIE CHAOSU

Dodatek A

Přehled maticové algebry

A.1 Blokové matice

Necht matice A je zadána blokově

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad (\text{A.1})$$

Pak definujeme *Schurův komplement* A_{22} jako

$$D_{22} = A_{22} - A_{21}A_{11}^{-1}A_{12} \quad (\text{A.2})$$

a *Schurův komplement* A_{11} jako

$$D_{11} = A_{11} - A_{12}A_{22}^{-1}A_{21} \quad (\text{A.3})$$

Inverze k A pak může být psána ekvivalentně jedním ze tří vzorců

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}D_{22}^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}D_{22}^{-1} \\ -D_{22}^{-1}A_{21}A_{11}^{-1} & D_{22}^{-1} \end{bmatrix} \quad (\text{A.4})$$

$$A^{-1} = \begin{bmatrix} D_{11}^{-1} & -D_{11}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}D_{11}^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{21}D_{11}^{-1}A_{12}A_{22}^{-1} \end{bmatrix} \quad (\text{A.5})$$

$$A^{-1} = \begin{bmatrix} D_{11}^{-1} & -D_{11}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}D_{11}^{-1} & D_{22}^{-1} \end{bmatrix} \quad (\text{A.6})$$

Porovnáním těchto tří tvarů se dá odvodit *lemma o inverzní matici*

$$(A_{11}^{-1} - A_{12}A_{22}A_{21})^{-1} = A_{11} - A_{11}A_{12}(A_{21}A_{11}A_{12} + A_{22}^{-1})^{-1}A_{21}A_{11} \quad (\text{A.7})$$

Dále platí

$$|A| = |A_{11}||D_{22}| \quad (\text{A.8})$$

$$|A| = |A_{22}||D_{11}| \quad (\text{A.9})$$

A.2 Maticový kalkulus

Bud' $f : \mathbb{R}^n \rightarrow \mathbb{R}$ funkce, $x_0 \in \mathbb{R}^n$ je reálný vektor. Definujeme *gradient* funkce f v bodě x_0 jako vektor

$$\frac{\partial f}{\partial x}(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \dots, \frac{\partial f}{\partial x_n}(x_0) \right)^T$$

Bud' $F : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ funkce, $X_0 \in \mathbb{R}^{m \times n}$ je reálná matice. *Gradient* funkce F v bodě X_0 je matice

$$\frac{\partial F}{\partial X}(X_0) = \left[\frac{\partial F}{\partial X_{ij}}(X_0) \right]_{i,j}$$

Některé užitečné gradienty:

$$\frac{\partial}{\partial x}(y^T x) = y \tag{A.10}$$

$$\frac{\partial}{\partial x}(x^T y) = y \tag{A.11}$$

$$\frac{\partial}{\partial x}(y^T A x) = A y \tag{A.12}$$

$$\frac{\partial}{\partial x}(x^T A y) = A y \tag{A.13}$$

$$\frac{\partial}{\partial x}(x^T A x) = (A + A^T)x \tag{A.14}$$

$$\frac{\partial}{\partial A} \text{tr}(A) = I \tag{A.15}$$

$$\frac{\partial}{\partial A} \text{tr}(BAD) = B^T D^T \tag{A.16}$$

$$\frac{\partial}{\partial A} \text{tr}(ABA^T) = 2AB \tag{A.17}$$

$$\frac{\partial}{\partial A} |BAD| = |BAD| A^{-T} \tag{A.18}$$

kde jsme položili $A^{-T} = (A^{-1})^T$.

Bud' $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ funkce, $x_0 \in \mathbb{R}^n$ je reálný vektor. *Jakobián* funkce g v bodě x_0 je matice

$$\frac{\partial g}{\partial x}(x_0) = \left[\frac{\partial g_i}{\partial x_j}(x_0) \right]_{i,j}$$

Některé užitečné Jakobiány:

$$\frac{\partial}{\partial x}(Ax) = A \tag{A.19}$$

Stopa matice je součet jejích diagonálních prvků. Značíme $\text{tr } A$.

Platí

$$\text{tr}(A + B) = \text{tr } A + \text{tr } B \tag{A.20}$$

$$y^T x = \text{tr}(yx^T) \tag{A.21}$$

Dodatek B

Pravděpodobnost

B.1 Základní pojmy

Označme $(\Omega, \mathcal{A}, \mathcal{P})$ jako pravděpodobnostní prostor, kde Ω značí základní prostor, \mathcal{A} jevé pole na Ω a P je pravděpodobnostní míra (pravděpodobnost) na \mathcal{A} .

Náhodná veličina (náhodný vektor) je zobrazení $X : \Omega \rightarrow \mathbb{R}^n$.

Distribuční funkce náhodného vektoru $X \in \mathbb{R}^n$ je zobrazení $F : \mathbb{R}^n \rightarrow \mathbb{R}$ definované vztahem

$$F(x) = P(\{\omega \in \Omega : X(\omega) \leq x\})$$

Hustota pravděpodobnosti náhodného vektoru $X \in \mathbb{R}^n$ je zobrazení $F : \mathbb{R}^n \rightarrow \mathbb{R}$ a je definována jako

$$p(x) = \frac{\partial^n F}{\partial x_1 \cdots \partial x_n}(x)$$

Hustota pravděpodobnosti má smysl pouze pro diferencovatelné distribuční funkce (spojité náhodné vektory).

Tam, kde nemůže dojít k nedorozumění, budeme všechny distribuční funkce označovat symbolem F a všechny hustoty symbolem p . $p(x)$ bude tedy hustota příslušná náhodné veličině X a $p(y)$ bude hustota příslušná náhodné veličině Y .

Spojení $Z = (X^T, Y^T)^T$ je náhodný vektor. Známe-li hustotu $p(z)$, pak hustoty $p(x)$ a $p(y)$ získáme marginalizací

$$p(x) = \int p(z) dy \tag{B.1}$$

$$p(y) = \int p(z) dx \tag{B.2}$$

B.2 Číselné charakteristiky (momenty)

Střední hodnota náhodného vektoru x (první moment) je vektor, definovaný jako

$$E(x) = \int_{\mathbb{R}^n} x dF(x)$$

Druhý moment náhodného vektoru x je matice

$$m_2(x) = \int_{\mathbb{R}^n} xx^T dF(x)$$

Rozptyl náhodného vektoru x je definován jako

$$\text{var}(x) = \text{E}(x - \text{E}(x))(x - \text{E}(x))^T$$

Platí, že $\text{var}(x) = m_2(x) - [\text{E}(x)]^2$.

Zápisem $X \in \mathcal{L}n$ rozumíme, že $X : \Omega \rightarrow \mathbb{R}^n$ je náhodný vektor dimenze n s konečným druhým momentem. Takový vektor má i konečný první moment.

$\{x_j\}_{j=-\infty}^{\infty}$ je spočetná posloupnost náhodných vektorů (náhodný proces).

B.3 Normální rozdělení

Náhodný vektor $X \in \mathcal{L}n$ má normální rozdělení, právě když hustota pravděpodobnosti je dána vzorcem

$$p(x) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} e^{\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)} \quad (\text{B.3})$$

kde $\mu \in \mathbb{R}^n$ je reálný vektor a $\Sigma \in \mathbb{R}^{n \times n}$ je symetrická pozitivně definitní matice. Pak píšeme $X \sim N(\mu, \Sigma)$.

Lze ukázat, že pro parametry μ, Σ platí

$$\text{E} X = \mu \quad (\text{B.4})$$

$$\text{var} X = \Sigma \quad (\text{B.5})$$

Normálně rozdělené náhodné vektory X a Y jsou *nezávislé* právě tehdy, jsou-li *nekorelované*, tj. $\text{cov}(X, Y) = 0$.

Mějme náhodný vektor $X \in \mathcal{L}n$, $X \sim N(\mu_x, \Sigma_x)$, vektor $a \in \mathbb{R}^m$ a matici $B \in \mathbb{R}^{m \times n}$. Pak náhodný vektor $Y = a + BX$ má *normální* rozdělení a platí $Y \sim N(a + B\mu_x, B\Sigma_x B^T)$. Zkráceně

$$X \sim N(\mu_x, \Sigma_x) \implies a + BX \sim N(a + B\mu_x, B\Sigma_x B^T) \quad (\text{B.6})$$

B.4 Podmíněná hustota pravděpodobnosti

Podmíněná hustota pravděpodobnosti $p(x|y)$ je definována následujícím předpisem

$$p(x|y) = \frac{p(x, y)}{p(y)} \quad (\text{B.7})$$

Na tuto funkci nahlížíme jako na funkci proměnné x při fixované hodnotě y . Je to hustota pravděpodobnosti náhodné veličiny x při dané hodnotě náhodné veličiny y , přičemž závislost obou náhodných veličin je vyjádřena simultánní hustotou pravděpodobnosti $p(x, y)$. Jako hustota pravděpodobnosti splňuje podmínku

$$\int p(x|y) dx = 1 \quad (\text{B.8})$$

Z definice podměné hustoty se jednoduše odvodí *řetězové pravidlo*

$$p(x, y) = p(x|y)p(y) = p(y|x)p(x) \quad (\text{B.9})$$

a také často používaný *Bayesův vzorec*

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)} = \frac{p(x|y)p(y)}{\int p(x|y)p(y) dy} \quad (\text{B.10})$$

Jsou-li náhodné vektory X a Y nezávislé, pak platí

$$p(x|y) = p(x) \quad (\text{B.11})$$

B.5 Podmíněné číselné charakteristiky

Pro náhodnou veličinu $X|z = (X|Z = z)$ zavádíme podmíněnou střední hodnotu $E(X|z)$ jako

$$E(X|z) = \int xp(x|z) dx \quad (\text{B.12})$$

z pravidel pro počítání s podmíněnými hustotami vyplývají ekvivalentní vztahy

$$E(X|Z) = \int x \frac{p(x, z)}{p(z)} dx \quad (\text{B.13})$$

$$E(X|Z) = \frac{\int xp(x, z) dx}{\int p(x, z) dx} \quad (\text{B.14})$$

$$E(X|Z) = \frac{\int xp(x|z)p(z) dx}{\int p(x|z)p(z) dx} \quad (\text{B.15})$$

B.6 Lineární regrese

Nechť $y, e \in \mathcal{L}1$ jsou náhodné veličiny, x_1, \dots, x_p jsou reálné proměnné, b_1, \dots, b_p jsou reálné konstanty (parametry, regresní koeficienty). *Lineární regresní model* je rovnice

$$Y = b_1x_1 + \dots + b_px_p + e \quad (\text{B.16})$$

popisující lineární závislost *výstupu* Y na *regresech* x_j . Člen e je náhodná chyba, nebo tzv. *šum měření* s vlastnostmi

$$\begin{aligned} E(e) &= 0 \\ \text{var}(e) &= \sigma^2 \end{aligned}$$

kde σ^2 je známá konstanta.

Lineární regresní model je přístupný pozorování. Provedeme n měření (pozorování) a dostaneme vektor výstupů y a *matici plánu* X :

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix}$$

Při měření předpokládáme, že sumy e_i tvoří posloupnost stejně rozdělených a *stochasticky nezávislých* náhodných veličin.

Na základě dat X, y máme odhadnout neznámé regresní koeficienty $b = (b_1, \dots, b_p)^T$ tak aby výraz

$$q(b) = \sum_{i=1}^n (y_i - \sum_{j=1}^p b_j x_{ij})^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (\text{B.17})$$

byl minimální. Funkce $q(b)$ se nazývá *reziduální součet čtverců*. Je to kvadratická forma v proměnných b a její globální minimum se nachází v bodě \hat{b} , který se vypočte jako

$$\hat{b} = (X^T X)^{-1} X^T y \quad (\text{B.18})$$

Je vidět, že \hat{b} je *lineární* funkcí y . Výraz

$$s^2 = q(\hat{b}) = \frac{1}{n-p} \sum_{i=1}^n (y_i - \sum_{j=1}^p \hat{b}_j x_{ij})^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (\text{B.19})$$

se nazývá *reziduální rozptyl* a je nestranným odhadem hodnoty σ^2 .