Základy matematického modelování

17. května 2011

Obsah

0 Úvod - aneb co nás čeká **3**

1 Vyrovnávání dat **4**

1.1 Metoda nejmenších čtverců 4

1.2 Vážená metoda nejmenších čtverců* 9

1.3 Klouzavé průměry 9

1.4 Whittacker-Hendersonova metoda 12

2 Diferenciální rovnice **13**

2.1 Modely růstu 13

2.2 Doplňování rezerv v dávkově definovaném penzijním fondu 16

3 Teorie lineární regulace **17**

3.1 Stavový model 18

3.2 Obecný model 19

4 Markovovy řetězce **22**

4.1 Klasifikace stavů Markovova řetězce 23

4.2 Limitní a stacionární rozdělení Markovova řetězce 26

4.3 Systém bonus-malus v pojišťovnictví 27

4.4 * Markovovy řetězce se stavy z množiny reálných čísel 30

5 Časové řady **30**

5.1 Klouzavé průměry 31

5.2 Autoregresní modely řádu r (AR(r)) 32

5.3 Lineární predikce 36

6 Poissonův proces **38**

0 Úvod - aneb co nás čeká

Přednáška se zabývá základními matematickými metodami práce s daty, vzniklými nebo řízenými nějakým náhodným mechanismem. Nicméně na rozdíl od základní přednášky z pravděpodobnosti a matematické statistiky se zde budeme zabývat časovými daty, hlavně časovými řadami. Tedy daty, které mají podobu dvou posloupností $(t_1, t_2, \dots t_k)$ - časových okamžiků, ve kterých jsme prováděli pozorování a $(x_{t_1}, x_{t_2}, \dots, x_{t_k})$ - odpovídajících hodnot v těchto časech pozorovaných.

Začneme prostým vyrovnáváním dat, kdy se vhodnou úpravou dat snažíme docílit toho, aby upravená data lépe vystihovala pozorovaný jev (snažíme se různými metodami zbavit nepodstatných „náhodných fluktuací“ a získat vyrovnaná data ukazující hlavní trend). Při tomto způsobu analýzy nepoužíváme žádný stochastický model pro mechanismus generující pozorovaná data.

Složitější metodu pro růstová data si ukážeme v druhé kapitole, kdy už používáme model popisující vývoj celkové velikosti populace nebo celkového množství něčeho (např. rezervy penzijního fondu). Model je sice deterministický (daný řešením diferenciální rovnice popisující závislost mezi velikostí a rychlostí růstu populace), takže se v podstatě opět snažíme o proložení dat nějakou křivkou, ale nyní už máme tvar oné křivky daný předem (naším modelem - růstovým modelem daným diferenciální rovnicí), máme pro tu křivku interpretaci, máme interpretaci i pro parametry té křivky.

Ve třetí kapitole se podíváme na lineární soustavy a základy lineární regulace. Lineární soustava transformuje vstupní posloupnost stavů na výstupní posloupnost stavů a vztah mezi vstupní a výstupní posloupností je lineární (členy výstupní posloupnosti jsou lineární kombinací (části) vstupní posloupnosti). Výstup v čase t může záviset na části nebo na „celé minulosti“ vstupů do času t. Budeme zkoumat stabilitu lineárních soustav, tj. jestli se výstupní posloupnost vrací do rovnovážného stavu i při určitých výkyvech posloupnosti vstupní. Jako příklad si ukážeme zjednodušený model nabídky a poptávky. Tím zakončíme povídání o deterministických modelech a přesuneme se k modelům stochastickým čili náhodným.

Začneme Markovskými řetězci. Markovský řetězec je posloupnost náhodných veličin, tedy stav Markovského řetězce v daném čase je náhodná veličina. Nicméně struktura náhodnosti je tu velmi jednoduchá. Pravděpodobnostní rozdělení stavů Markovského řetězce v budoucnosti je totiž zcela určeno jeho stavem v přítomném okamžiku – tj. při znalosti současného stavu jsou minulost a budfile:///usr/share/ubuntu-artwork/home/index.htmloucnost Markovského řetězce (procesu, jenž Markovským řetězcem modelujeme) nezávislé. Ukážeme si jak lze takové modely využít v pojišťovnictví.

V páté kapitole se tak trochu vrátíme k analýze časových posloupností z kapitoly jedna, teď už ovšem vybaveni stochastickým modelem pro časové řady, který má také lineární strukturu obdobnou té v kapitole 3, a budeme se zabývat i predikcí budoucích pozorování v časových řadách.

Na závěr si v šesté kapitole ukážeme model, který se liší od všeho výše uvedeného. Protože všechny modely, které jsme zatím představili, vypadaly tak, že v pevných časových okamžicích jsme pozorovali realizaci nějaké náhodné veličiny, a zajímala nás velikost této veličiny. V poslední kapitole budeme zkoumat situaci, kdy v náhodných časových okamžicích se vyskytuje nějaká událost. Čili předmětem našeho zkoumání nejsou ony události jako takové, ale časy jejich výskytu, resp. počet takových událostí k nějakému pevnému datu (škodné události v pojištění). Takové situace se dají modelovat pomocí Poissonova procesu, případně složeného

Poissonova procesu a jako aplikaci si ukážeme analýzu systémů hromadné obsluhy.

1 Vyrovnávání dat

Naše pozorování jsou dána tabulkou čísel

x_1	x_2	$\cdots$	x_n
y_1	y_2	$\cdots$	y_n

(kde x_i často bývají časové údaje) a my chceme data proložit nějakou „hladkou funkcí“, která by vystihovala hlavní vlastnosti dat, ale ignorovala malé fluktuace a nepřesnosti.

1.1 Metoda nejmenších čtverců

Zde volíme tvar hladké funkce předem, například jako přímku $x \mapsto a + bx$ nebo parabolu $x \mapsto a + bx + cx^2$. Parametry křivky $a, b, c, \dots$ určíme jako ty, které minimalizují součet čtverců odchylek mezi křivkou a daty.

Prokládání přímkou $x \mapsto a + bx$

Vyrovnaná data budeme značit

$$\hat{y}_i := a + bx_i, \quad \text{pro } i = 1, \cdots, n \tag{1}$$

Reziduální součet čtverců budeme označovat file:///usr/share/ubuntu-artwork/home/index.html

$$S(a,b) := \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - a - bx_i)^2.$$

Tato funkce $(a,b) \mapsto S(a,b)$ je hladká a konvexní. Existuje-li její minimum, lze ho získat řešením rovnic $\frac{\partial S}{\partial a} = 0$, $\frac{\partial S}{\partial b} = 0$. Odpovídající rovnice lze zapsat ve tvaru

$$\sum_{i=1}^{n} (y_i - a - bx_i) = 0, \qquad \sum_{i=1}^{n} (y_i - a - bx_i)x_i = 0. \tag{2}$$

Pro další výpočty budeme používat značení

$$\overline{y} := \frac{1}{n} \sum_{i=1}^{n} y_i, \qquad \overline{x} := \frac{1}{n} \sum_{i=1}^{n} x_i.$$

První z rovnic v (2) lze tak přepsat do tvaru

$$\overline{y} = a + b\overline{x}. \tag{3}$$

Druhou rovnici v (2) lze s použitím (3) upravit takto

$$0 = \sum_{i=1}^{n} (y_i - \overline{y} + \overline{y} - a - bx_i + b\overline{x} - b\overline{x})(x_i - \overline{x} + \overline{x}) = \sum_{i=1}^{n} (y_i - \overline{y})(x_i - \overline{x}) - b \sum_{i=1}^{n} (x_i - \overline{x})^2.$$

Odtud plyne

$$b = \frac{\sum_{i=1}^{n} (y_i - \overline{y})(x_i - \overline{x})}{\sum_{i=1}^{n} (x_i - \overline{x})^2} = \frac{\sum_{i=1}^{n} y_i x_i - n\overline{y}\,\overline{x}}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}. \tag{4}$$

Přímka $x \mapsto a + bx$, kde parametry a, b splňují (3),(4), je pak hledanou přímkou. Vyrovnané hodnoty získáme dosazením do (1).

Prokládání funkcí tvaru $x \mapsto a_1 f_1(x) + \cdots + a_k f_k(x), \quad k \leq n$

V tomto případě dáme přednost vektorovému zápisu $y := (y_1, \cdots, y_n)^T \in \mathbb{R}^n$, podobně $\hat{y} := (\hat{y}_1, \cdots, \hat{y}_n)^T \in \mathbb{R}^n$. Definiční rovnost pro vyrovnaná data

$$\hat{y}_i = a_1 f_1(x_i) + \cdots + a_k f_k(x_i), \;\; i = 1, \cdots, n \tag{5}$$

tak lze přepsat do maticového tvaru

$$\hat{y} = Fa, \;\; \text{kde} \;\; F = (f_j(x_i))_{i,j=1}^{n,k} \;\; \text{a kde} \;\; a = (a_1, \cdots, a_k)^T \in \mathbb{R}^k. \tag{6}$$

Tento zápis lze rozepsat následujícím způsobem

$$\begin{pmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_n \end{pmatrix} = \begin{pmatrix} f_1(x_1) & \cdots & f_k(x_1) \\ \vdots & \ldots & \vdots \\ f_1(x_n) & \cdots & f_k(x_n) \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} \tag{7}$$

Reziduální součet čtverců je pak možné vyjádřit ve tvaru

$$S(a) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{k} a_j f_j(x_i) \right)^2 = (y - \hat{y})^T (y - \hat{y}) = (y - Fa)^T (y - Fa). \tag{8}$$

Věta 1 *Vektor $a \in \mathbb{R}^k$ je bodem globálního minima funkce $S(a)$ právě tehdy, když a je řešením soustavy normálních rovnic*

$$F^T F a = F^T y. \tag{9}$$

Hodnota vyrovnaných dat $\hat{y} := F\hat{a}$, kde $\hat{a}$ je řešením (9)*, nezávisí na volbě řešení $\hat{a}$. Rovnice* (9) *má vždy alespoň jedno řešení.*

Důkaz: Nechť $\hat{a}$ je řešením soustavy (9). Pak

$$\begin{aligned} S(a) &= (y - Fa)^T (y - Fa) = (y - F\hat{a} + F\hat{a} - Fa)^T (y - F\hat{a} + F\hat{a} - Fa) & (10) \\ &= S(\hat{a}) + (y - F\hat{a})^T F(\hat{a} - a) + (\hat{a} - a)^T F^T (y - F\hat{a}) + (\hat{a} - a)^T F^T F(\hat{a} - a). & (11) \end{aligned}$$

file:///usr/share/ubuntu-artwork/home/index.html Z předpokladu, že $\hat{a}$ je řešení (9), plyne, že

$$F^T (y - F\hat{a}) = F^T y - F^T F\hat{a} = 0 \tag{12}$$

a odtud pak plyne

$$S(a) = S(\hat{a}) + ||F(\hat{a} - a)||^2. \tag{13}$$

Bod $\hat{a}$ je tedy skutečně bodem minima funkce S.

Existence nějakého řešení: Hodnost rozšířené matice soustavy (9) (má rozměry $k \times (k+1)$)je podle věty o dimenzi jádra a obrazu rovna

$$\begin{aligned} h(F^T F, F^T y) &= k - \dim\{a \in \mathbb{R}^k : a^T F^T F = 0, a^T F^T y = 0\} & (14) \\ &= k - \dim\{a \in \mathbb{R}^k : F^T F a = 0, Fa = 0\} & (15) \\ &= k - \dim\{a \in \mathbb{R}^k : F^T F a = 0\} = h(F^T F), & (16) \end{aligned}$$

neboť hodnost $F^T F$ je stejná jako hodnost matice F a tedy $F^T F a = 0$ právě když $Fa = 0$. Existuje tedy alespoň jedno řešení $\hat{a}$ soustavy (9).

Buď nyní $\tilde{a}$ bodem globálního mimima funkce S. Pak z rovnice (13)

$$\min S = S(\tilde{a}) = S(\hat{a}) + ||F(\hat{a} - \tilde{a})||^2 = S(\hat{a}) = \min S.$$

Tedy $F\tilde{a} = F\hat{a}$ takže každý bod globálního minima funkce S splňuje rovnici (9) a vyrovnaná data $F\tilde{a}$ nezávisí na volbě $\tilde{a}$. □

Poznámka: Je-li matice F^TF regulární, pak vyrovnaná data obdržíme ze vztahu $\hat{y} := Hy$, kde $H := F(F^TF)^{-1}F^T$. Obecně není nutné předpokládat, že x_i jsou čísla, mohou to být např. dvojice či l-tice čísel.

Příklad: Prokládání přímkou $x \mapsto a_1 + a_2 x$

V tomto případě $f_1(x) = 1, f_2(x) = x$

$$F = (\mathbf{1}, x) = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}, \text{ kde } \mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \text{ a kde } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \tag{17}$$

Pak

$$F^TF = \begin{pmatrix} \mathbf{1}^T \\ x^T \end{pmatrix} (\mathbf{1}, x) = \begin{pmatrix} \mathbf{1}^T\mathbf{1} & \mathbf{1}^T x \\ x^T\mathbf{1} & x^T x \end{pmatrix} = \begin{pmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{pmatrix}. \tag{18}$$

Příslušná inverze je pak tvaru

$$(F^TF)^{-1} = \frac{1}{n\sum_{i=1}^n x_i^2 - n^2\overline{x}^2} \begin{pmatrix} \sum_{i=1}^n x_i^2 & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & n \end{pmatrix}. \tag{19}$$

Matice F^Ty je tvaru

$$\begin{pmatrix} \mathbf{1}^T \\ x^T \end{pmatrix} y = \begin{pmatrix} \mathbf{1}^T y \\ x^T y \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{pmatrix}. \tag{20}$$

Příslušné koeficienty a_1, a_2 pak získáme

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \frac{1}{\sum_{i=1}^n x_i^2 - n\overline{x}^2} \begin{pmatrix} \overline{y}\sum_{i=1}^n x_i^2 - \overline{x}\sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n x_i y_i - n\overline{x}\,\overline{y} \end{pmatrix} \tag{21}$$

Odtud plyne

$$a_2 = \frac{\sum_{i=1}^n x_i y_i - n\overline{x}\,\overline{y}}{\sum_{i=1}^n x_i^2 - n\overline{x}^2}, \tag{22}$$

a

$$a_1 = \frac{\overline{y}\left(\sum_{i=1}^n x_i^2 - n\overline{x}^2\right) - \overline{x}\left(\sum_{i=1}^n x_i y_i - n\overline{x}\,\overline{y}\right)}{\sum_{i=1}^n x_i^2 - n\overline{x}^2} = \overline{y} - a_2\overline{x}, \tag{23}$$

což přesně odpovídá vztahů m (3) a (4). □

Pravděpodobnostní interpretace: regresní model*

Předpokládáme, že posloupnost $y := (y_1, \cdots, y_n)^T$ je realizací náhodného vektoru $Y := (Y_1, \cdots, Y_n)^T$ vyhovující lineárnímu modelu

$$Y = Fa + e, \tag{24}$$

kde náhodný vektor $e = (e_1, \cdots, e_n)^T$ má složky s nulovou střední hodnotou a $\operatorname{var} e = \sigma^2 I$, kde I je $n \times n$-rozměrná jednotková matice a $\sigma^2 > 0$.

Poznámka: Zápis (24) říká, že vektor Y má střední hodnotu Fa a varianční matici $\sigma^2 I$. Právě tato lineární závislost[1] střední hodnoty $EY = (EY_1, \cdots, EY_n)^T$ na vektoru parametrů $a = (a_1, \cdots, a_k)'$ je důvodem, proč se modelu říká lineární.

Věta 2 *Nechť matice $F^T F$ je regulární a nechť $\tilde{H} \in \mathbb{R}^{n\times n}$ je matice taková, že $\tilde{H}F = F$. Nechť náhodný vektor Y vyhovuje modelu* (24)*. Označme*

$$\tilde{Y} := \tilde{H}Y \quad a \quad \hat{Y} := HY, \quad kde \quad H := F(F^T F)^{-1} F^T. \tag{25}$$

Pak $E\hat{Y} = E\tilde{Y} = EY = Fa$, navíc $E||\hat{Y} - E\hat{Y}||^2 \le E||\tilde{Y} - E\tilde{Y}||^2 < \infty$, přičemž rovnost nastává právě tehdy, když $\tilde{Y} = \hat{Y}$ skoro jistě.

Důkaz: Zřejmě $EY = Fa$. Dále $E\hat{Y} = \tilde{H}\,EY = \tilde{H}Fa = Fa$. Podobně $E\hat{Y} = H\,EY = Fa$, neboť $HF = F$. Navíc,

$$\begin{aligned} E||\tilde{Y} - Fa||^2 &= E(\tilde{Y} - \hat{Y} + \hat{Y} - Fa)^T(\tilde{Y} - \hat{Y} + \hat{Y} - Fa) \\ &= E||\tilde{Y} - \hat{Y}||^2 + E(\tilde{Y} - \hat{Y})^T(\hat{Y} - Fa) + E(\hat{Y} - Fa)^T(\tilde{Y} - \hat{Y}) + E||\hat{Y} - Fa||^2. \end{aligned}$$

První člen pravé strany je vždy nezáporný. Je nulový právě tehdy, když $\tilde{Y} = \hat{Y}$ skoro jistě. Zbývá tedy ukázat, že druhý a třetí člen jsou rovny nule. Oba jsou matice řádu 1×1, přičemž jeden z druhého (z formálního hlediska) získáme transponováním. Z faktického hlediska to znamená, že oba členy jsou si navzájem rovny.

Podle (25)

$$\tilde{Y} - \hat{Y} = (\tilde{H} - H)Y = (\tilde{H} - H)(Fa + e) = (\tilde{H} - H)e,$$

neboť $\tilde{H}F = HF = F$. Dále

$$\hat{Y} - Fa = H(Fa + e) - Fa = He.$$

Odtud

$$E(\tilde{Y} - \hat{Y})(\hat{Y} - Fa)^T = E(\tilde{H} - H)ee^T H^T = (\tilde{H} - H)(\sigma^2 I)H = \sigma^2(\tilde{H}H - H) = \mathbf{0} \in \mathbb{R}^{n\times n},$$

neboť $\tilde{H}H = H^2 = H = H^T$, kde $H = F(F^T F)^{-1}F$. $\square$

[1] Jiná lineární závislost na vektoru $a = (a_1, \cdots, a_k)^T$ než závislost tvaru $a \mapsto Fa$ kde F matice s k sloupci neexistuje (pokud má být výsledkem této závislosti konečně rozměrný vektor čísel).

Po částech parametrické vyrovnávání - splajny (spline)

Zatím jsme vyrovnávali data pomocí předem dané křivky, závislé na několika málo parametrech (např. parabola). Nyní budeme vyrovnávat pomocí křivek jež jsou flexibilnější – v jednotlivých předem zvolených intervalech je křivka definovaná různě, ale jako celek si stále zachovává určitou „hladkost“, vyjádřenou pomocí spojitosti derivací až do určeného řádu.

Definice 1 *Buď dána posloupnost bodů* $\{u_j\}_{j=1}^m$ *, ty nazveme uzly. Splajnem řádu* k *nazveme funkci* f*, která je v každém intervalu* $[u_j, u_{j+1}]$ *polynomem stupně* k *a která má v celém definičním oboru spojité derivace až do řádu* $k-1$*.*

Po částech parametrické vyrovnávání si ukážeme na případu kubického splajnu (splajnu řádu 3). Začneme případem dvouobloukového splajnu.

Nejprve zvolíme index $k \in \{1, \cdots, n\}$ a tím i odpovídající uzel x_k. Vyrovnané hodnoty $\hat{y}_i$ budeme definovat jako $f(x_i)$ pro $i \in \{1, \cdots, n\}$, kde

$$\begin{aligned} f(x) &= c_0 + c_1 x + c_2 x^2 + c_3 x^3, && \text{pokud } x \le x_k && (26) \\ &= c_0 + c_1 x + c_2 x^2 + c_3 x^3 + d(x - x_k)^3, && \text{pokud } x > x_k. && (27) \end{aligned}$$

Funkce f má tím pádem spojitou druhou derivaci (i v bodě x_k – ověření přenecháváme čtenáři). Hodnoty parametrů $c_0, \cdots, c_3, d$ volíme tak, abychom minimalizovali hodnotu

$$\sum_{i=1}^{k} \left(y_i - c_0 - c_1 x_i - c_2 x_i^2 - c_3 x_i^3\right)^2 + \sum_{i=k+1}^{n} \left(y_i - c_0 - c_1 x_i - c_2 x_i^2 - c_3 x_i^3 - d(x_i - x_k)^3\right)^2.$$

Zřejmě jde o metodu nejmenších čtverců s maticí

$$F = \begin{pmatrix} 1 & x_1 & x_1^2 & x_1^3 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{k+1} & x_{k+1}^2 & x_{k+1}^3 & (x_{k+1} - x_k)^3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & x_n^3 & (x_n - x_k)^3 \end{pmatrix}.$$

Zbývá tedy řešit soustavu normálních rovnic

$$\left(\begin{array}{ccccc|c} n & \sum x_i & \sum x_i^2 & \sum x_i^3 & \sum (x_i - x_k)_+^3 & \sum y_i \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \sum x_i^4 & \sum x_i (x_i - x_k)_+^3 & \sum x_i y_i \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 & \sum x_i^5 & \sum x_i^2 (x_i - x_k)_+^3 & \sum x_i^2 y_i \\ \sum x_i^3 & \sum x_i^4 & \sum x_i^5 & \sum x_i^6 & \sum x_i^3 (x_i - x_k)_+^3 & \sum x_i^3 y_i \\ \sum (x_i - x_k)_+^3 & \sum x_i (x_i - x_k)_+^3 & \sum x_i^2 (x_i - x_k)_+^3 & \sum x_i^3 (x_i - x_k)_+^3 & \sum ((x_i - x_k)_+)^6 & \sum (x_i - x_k)_+^3 y_i \end{array}\right),$$

kde $(x_i - x_k)_+ = I_{(x_i - x_k) > 0}(x_i - x_k)$ značí kladnou část čísla $(x_i - x_k)$.

Pro obecný p-obloukový kubický splajn postupujeme obdobně, volíme $p-1$ uzlů a $p+3$ parametrů $c_0, c_1, c_2, c_3, d_1, \ldots d_{p-1}$.

1.2 Vážená metoda nejmenších čtverců*

Je rozšířením metody nejmenších čtverců a umožňuje dát některým pozorováním větší či menší váhu, zohledňující například důležitost nebo důvěryhodnost jednotlivých pozorování.

Kritériem pro metodu vážených nejmenších čtverců je minimalizace

$$\sum_{i=1}^{n} w_i(y_i - \hat{y}_i)^2, \tag{28}$$

kde $w_i > 0$ jsou námi volené váhy a $\hat{y}_i$ jsou jako obvykle vyrovnané hodnoty. Označme $W = diag\,\{w_1, \cdots, w_n\}$ příslušnou diagonální matici sestavenou z jednotlivých vah. Hodnotu (28) lze zapsat ve tvaru

$$S_W(a) = (y - \hat{y})^T W(y - \hat{y}) = (y - Fa)^T W(y - Fa). \tag{29}$$

Doteď jsme používali značení $\hat{y} = Fa$, tj. vyrovnané hodnoty závisely na hodnotě parametru $a \in \mathbb{R}^k$. Od této chvíle budou vyrovnané hodnoty odpovídat hodnotě parametru $\hat{a}$ získané minimalizací vážených nejmenších čtverců (29).

Soustavu normálních rovnic získáme derivováním hodnoty

$$S_W(a) = y^T Wy - 2a^T F^T Wy + a^T F^T WFa \tag{30}$$

a porovnáním s nulou ve tvaru

$$0 = -2F^T Wy + 2F^T WFa \qquad \text{resp. ve tvaru} \qquad F^T WFa = F^T Wy. \tag{31}$$

Platí obdobná věta jako pro normální metodu nejmenších čtverců:

Věta 3 *Vektor $a \in \mathbb{R}^k$ je bodem globálního minima funkce $S_W(a)$ právě tehdy, když a je řešením soustavy normálních rovnic* (31). *Hodnota vyrovnaných dat $\hat{y} := F\hat{a}$, kde $\hat{a}$ je řešením* (31), *nezávisí na volbě řešení $\hat{a}$. Rovnice* (31) *má vždy alespoň jedno řešení.*

1.3 Klouzavé průměry

Klouzavé průměry (KP) jsou schopny postihnout trend v datech - tj. směr a míru pohybu pozorovaných hodnot (bez toho, že bychom měli pro trend nějaký specifický model). Data vyrovnávají pouze lokálně, tj. v daném bodě se vyrovnaná hodnota počítá pouze z několika okolních hodnot, nikoli z celé pozorované řady.

Na základě dat $y = (y_1, \cdots, y_n)^T$ získáme vyrovnané hodnoty $\hat{y}_i$ předpisem

$$\hat{y}_i := \sum_{j:=-r}^{r} a_j y_{i+j}, \ \text{pro } j = r+1, \cdots, n-r, \tag{32}$$

kde váhy $(a_{-r}, \cdots, a_r)$ splňují $\sum_{j=-r}^{r} a_j = 1$. Číslo $2r + 1$ nazveme délkou klouzavého průměru.

Klouzavé aritmetické průměry

Všechny váhy $a_{-r}, \cdots, a_r$ jsou stejné a jsou rovny hodnotě $\frac{1}{2r+1}$..

Vyrovnaná hodnota je tedy opravdu jen prostý průměr z $2r+1$ okolních hodnot. Výhodou klouzavých aritmetických průměrů je, že všechny váhy jsou nezáporné, nevýhodou, že není vyrovnán počáteční a koncový úsek dat.

Konstrukce KP vyrovnáváním úseků polynomy

Váhy klouzavých průměrů volíme tak, že aproximujeme $2r+1$ členů řady $y_{t-r}, \ldots, y_t, \ldots y_{t+r}$ vhodným polynomem (řádu k) a hodnotu tohoto polynomu v bodě t použijeme jako vyrovnanou hodnotu $\hat{y}_t$. Číslo k nazýváme řád klouzavého průměru.

K tomu, abychom dospěli k příslušným vahám pro klouzavé průměry délky $2r+1$ a řádu k budeme uvažovat model

$$y_{t+u} = c_0 + c_1 u + c_2 u^2 + \cdots + c_k u^k \qquad \text{pro} \qquad u = -r, \cdots, r.$$

Maticově to lze zapsat ve tvaru

$$\mathbf{y}_t = \begin{pmatrix} y_{t-r} \\ \vdots \\ y_{t+r} \end{pmatrix} = \begin{pmatrix} 1 & -r & (-r)^2 & \cdots & (-r)^k \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & r & r^2 & \cdots & r^k \end{pmatrix} \begin{pmatrix} c_0 \\ \vdots \\ c_k \end{pmatrix} \tag{33}$$

Skutečné vyrovnané hodnoty pak obdržíme dosazením $u := 0$

$$\hat{y}_t = c_0 + c_1 u + c_2 u^2 + \cdots + c_k u^k|_{u=0} = c_0. \tag{34}$$

Parametry $c_0, \cdots, c_k$ získáme metodou nejmenších čtverců.

Příklad: Kubický polynom $k = 3$, $r = 2$. Uvažovaný model je tvaru

$$y_{t+u} = c_0 + c_1 u + c_2 u^2 + \cdots + c_3 u^3. \tag{35}$$

K volbě parametrů c_0, c_1, c_2, c_3 použijeme metodu nejmenších čtverců. Nejprve sestavíme matici

$$F = \begin{pmatrix} 1 & -2 & 4 & -8 \\ 1 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \end{pmatrix}. \tag{36}$$

Parametr $c = (c_0, \cdots, c_3)'$ dostaneme jako řešení soustavy normálních rovnic

$$F^T F c = F^T \mathbf{y}_t, \text{ kde } \mathbf{y}_t = (y_{t-2}, y_{t-1}, y_t, y_{t+1}, y_{t+2})^T \tag{37}$$

pro $t = 3, \cdots, n-2$. Projekční matici H obdržíme pomocí vzorce

$$H = F(F^T F)^{-1} F^T = \begin{pmatrix} \frac{69}{70} & \frac{2}{35} & \frac{-3}{35} & \frac{2}{35} & \frac{-1}{70} \\ \frac{2}{35} & \frac{27}{35} & \frac{12}{35} & \frac{-8}{35} & \frac{2}{35} \\ \frac{-3}{35} & \frac{12}{35} & \frac{17}{35} & \frac{12}{35} & \frac{-3}{35} \\ \frac{2}{35} & \frac{-8}{35} & \frac{12}{35} & \frac{27}{35} & \frac{2}{35} \\ \frac{-1}{70} & \frac{2}{35} & \frac{-3}{35} & \frac{2}{35} & \frac{69}{70} \end{pmatrix} \tag{38}$$

Vyrovnaná hodnota $\hat{y}_t = c_0$ je pak rovna součinu prostředního řádku a sloupcového vektoru $\mathbf{y}_t$

$$\hat{y}_t = \frac{1}{35}(-3, 12, 17, 12, -3)\mathbf{y}_t, \tag{39}$$

pro $t = 3, \cdots, n-2$. □

Pro klouzavé průměry zkonstruované vyrovnáváním úseků polynomy obecně platí, že:

- součet vah je roven 1
- váhy jsou symetrické kolem prostřední hodnoty
- pro sudé k je klouzavý průměr řádu k a $k+1$ stejný.

Počáteční a koncové úseky řady (tedy $t \leq r$ a $t \geq n-r$) nemůžeme proložit pomocí vzorce (34), protože nemáme k dispozici odpovídající $\mathbf{y}_t$. Použijeme proto prokládání „prvním a posledním možným polynomem“

$$\hat{y}_{1+r+u} = c_0 + c_1 u + c_2 u^2 + \cdots + c_k u^k, \quad u = -r, \cdots, r, \tag{40}$$

a

$$\hat{y}_{n-r+u} = c_0 + c_1 u + c_2 u^2 + \cdots + c_k u^k, \quad u = -r, \cdots, r, \tag{41}$$

koeficienty spočítáme opět metodou nejmenších čtverců.

Dosazením $u = r+1$ do rovnice (41) s odhadnutými koeficienty dostaneme předpověď na jeden krok dopředu. Tento postup lze ale použít jen pro krátkodobé předpovědi.

Poznámka: Počáteční, koncové a předpovědní klouzavé průměry se liší pro řády k a $k+1$.

Příklad - pokračování: Hodnoty y_{n-1}, y_n lze vyrovnat prokládáním polynomu

$$\hat{y}_{n-2+u} = c_0 + c_1 u + c_2 u^2 + c_3 u^3, \quad u = -2, -1, 0, 1, 2. \tag{42}$$

Uvědomme si, že metodou nejmenších čtverců dojdeme ke stejné projekční matici H jako v (38) a tedy

$$\hat{y}_{n-1} = \frac{1}{35}(2, -8, 12, 27, 2)\mathbf{y}_{n-2}, \quad \hat{y}_n = \frac{1}{70}(-1, 4, -6, 4, 69)\mathbf{y}_{n-2}. \tag{43}$$

Obdobně dojdeme ke vzorcům pro začátek řady

$$\hat{y}_2 = \frac{1}{35}(2, 27, 12, -8, 2)\mathbf{y}_3, \quad \hat{y}_1 = \frac{1}{70}(69, 4, -6, 4, -1)\mathbf{y}_3. \tag{44}$$

Předpověď o jeden krok dopředu získáme použitím hodnoty $u = 3$. Příslušné koeficienty lze získat následujícím způsobem

$$(1, 3, 3^2, 3^3)(F^T F)^{-1} F^T = \left(-\frac{4}{5}, \frac{11}{5}, -\frac{4}{5}, -\frac{14}{5}, \frac{16}{5}\right). \tag{45}$$

Předpověď o jeden krok dopředu pak vychází ve tvaru

$$\hat{y}_{n+1} = \frac{1}{5}(-4, 11, -4, -14, 16)\mathbf{y}_{n-2}. \tag{46}$$

□

1.4 Whittacker-Hendersonova metoda

Při vyhlazování jsou většinou v protikladu dva požadavky:

- dosáhnout dobré shody mezi daty a vyrovnanými daty (tedy co nejmenší (průměrná kvadratická) odchylka vyrovnaných hodnot $\hat{y}_i$ od původních hodnot y_i.)
- vyrovnaná data by měla být co nejvíce „hladká“ (tedy co nejmenší variace vyrovnaných hodnot).

Whittacker-Hendersonova metoda umožňuje hledat kompromis mezi těmito dvěma požadavky.

Vyrovnané hodnoty totiž určíme tak, aby minimalizovaly hodnotu

$$M = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + h \sum_{i=r+1}^{n} (\Delta^r \hat{y}_i)^2, \tag{47}$$

kde $h > 0$ je parametr, který neodhadujeme, ale volíme. Symbol

$$\Delta^r \hat{y}_i = \Delta^{r-1} \hat{y}_i - \Delta^{r-1} \hat{y}_{i-1},$$

označuje rekurzivně definovanou r-tou zpětnou diferenci posloupnosti $\hat{y}_i$, kde

$$\Delta^0 \hat{y}_i = \hat{y}_i \qquad \text{a} \qquad \Delta^1 \hat{y}_i = \Delta \hat{y}_i = \hat{y}_i - \hat{y}_{i-1}.$$

Tuto zpětnou diferenci lze zapsat pomocí binomické formule

$$\Delta^r \hat{y}_i = \binom{r}{0} \hat{y}_i - \binom{r}{1} \hat{y}_{i-1} + \cdots + (-1)^r \binom{r}{r} \hat{y}_{i-r} \tag{48}$$

První člen ve vzorci (47) tedy měří shodu mezi y_i a $\hat{y}_i$ a druhý člen hladkost vyrovnání pomocí diferencí r-tého řádu (r se většinou volí 2,3 nebo 4). Parametr $h > 0$ volíme podle toho, jestli klademe větší důraz na shodu nebo na hladkost.

Abychom byli schopni vyjádřit hodnotu M maticově, definujeme matici K rozměru $(n-r) \times n$

$$K = \begin{pmatrix} (-1)^r & \cdots & -\binom{r}{1} & 1 & \cdots & 0 \\ \vdots & (-1)^r & \cdots & -\binom{r}{1} & \cdots & \vdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & (-1)^r & \cdots & -\binom{r}{1} & 1 \end{pmatrix} \tag{49}$$

Pak lze psát

$$M = (\hat{y} - y)^T (\hat{y} - y) + h \cdot \hat{y}^T K^T K \hat{y}. \tag{50}$$

Matice druhých derivací M podle $\hat{y}_i$ je $2(I + hK^TK)$, což je positivně definitní matice, tedy bod ve kterém $\nabla_{\hat{y}}M(\hat{y}) = 0$ je bod lokálního minima funkce M. Derivováním podle $\hat{y}$ obdržíme

$$\left(\frac{\partial M}{\partial \hat{y}_1}, \ldots, \frac{\partial M}{\partial \hat{y}_n}\right) = \nabla_{\hat{y}}M(\hat{y}) = 2\hat{y} - 2y + 2hK^TK\hat{y}. \tag{51}$$

Bod $\hat{y} = (I + hK^TK)^{-1}y$ je tak bodem minima M na $\mathbb{R}^n$.

Poznámka: Podobně jako u metody nejmenších čtverců bychom mohli jednotlivým hodnotám y_i přiřadit různou váhu. Místo M bychom pak minimalizovali hodnotu

$$M_W = (\hat{y} - y)^TW(\hat{y} - y) + \hat{y}^TK^TK\hat{y}, \tag{52}$$

kde W je diagonální matice obsahující na diagonále příslušné váhy. Odpovídající minimum je pak tvaru

$$\hat{y} = (W + hK^TK)^{-1}y. \tag{53}$$

2 Diferenciální rovnice

2.1 Modely růstu

V této kapitole budeme zabývat jednoduchými modely růstu (populací, objemů nějaké komodity atp.) Funkce $y(t) \geq 0$ bude označovat velikost populace v čase t. A model bude určen diferenciální rovnicí, která udává závislost rychlosti růstu v čase t (popsaný derivací $\frac{dy}{dt}$) na velikosti populace v čase t.

Diferenciální rovnice bude vždy tvaru

$$\frac{dy}{dt} = ay \cdot g\left(\frac{y}{k}\right), \tag{54}$$

kde g je daná funkce a $a, k > 0$ jsou konstanty resp. parametry. Naše modely budou tedy vždy obsahovat část odpovídající lineární závislosti rychlosti růstu na současné velikosti populace, ale tato závislost bude korigována zvoleným tvarem funkce g.

Exponenciální růst

$g(x) \equiv 1$. Příslušná diferenciální rovnice je pak tvaru

$$\frac{dy}{dt} = ay, \quad a > 0. \tag{55}$$

Její řešení je tvaru

$$y(t) = be^{at},$$

kde a, b jsou parametry. Toto je klasický model neomezeného růstu.

Logistický růst

$g(x) = 1 - x$. Diferenciální rovnice

$$\frac{dy}{dt} = ay(1 - \frac{y}{k}) \tag{56}$$

má řešení ve tvaru

$$y(t) = \frac{k}{1 + be^{-at}}, \tag{57}$$

kde $a, b, k > 0$ jsou kladné parametry zajišťující, že výsledná funkce $y(t)$ bude rostoucí. Povšimněme si, že platí

$$0 < y(t) < k, \quad t \in \mathbf{R} \qquad \lim_{t \to \infty} y(t) = k.$$

Růst podle logistické funkce je tedy omezen saturační hodnotou k a odpovídá situaci, kdy má populace k dispozici jen omezené zdroje.

Gompertzova křivka

$g(x) = -\ln x$. Diferenciální rovnice

$$\frac{dy}{dt} = -ay \ln(y/k) \tag{58}$$

má řešení ve tvaru

$$y(t) = k \exp\{be^{-at}\},$$

kde $a, k > 0$, $b < 0$. I tato funkce je rostoucí, má saturační úroveň rovnu k, ovšem funkce není symetrická kolem inflexního bodu, jako tomu bylo u logistického růstu. Tento model se používá k vyrovnávání tabulek úmrtnosti ($(1 - y(t)) \sim$ počet lidí dožívajících se věku t).

Určení parametrů logistické křivky metodou Pearl-Reeda

Dvoufázová metoda, nejdříve v první fázi určím hrubé odhady parametrů, které pak v druhé fázi „opravím“. Opravu mohu provést i několikrát iterativně.

Fáze (a) - hrubý odhad:
Data rozdělíme do tří skupin oddělených třetinovým a dvoutřetinovým kvantilem o přibližné velikosti l. Data v každé skupině zprůměrujeme (na hodnoty $\overline{y_0}, \overline{y}_1, \overline{y}_2$) a takto vzniklými třemi body proložíme logistickou křivku

$$y(t) = \frac{k_0}{1 + e^{a_0 - a_1 t}},$$

přičemž souřadnice bodů t volíme postupně $0, l, 2l$. Musí tedy platit rovnice

$$\ln \frac{k - \overline{y}_0}{\overline{y}_0} = a_0, \quad \ln \frac{k - \overline{y}_1}{\overline{y}_1} = a_0 - a_1 l, \quad \ln \frac{k - \overline{y}_2}{\overline{y}_2} = a_0 - 2a_1 l. \tag{59}$$

Pokud odhadneme k_0 pomocí kvadratické aproximace logistické funkce

$$k_0 = \frac{2\overline{y}_0\overline{y}_1\overline{y}_2 - \overline{y}_1^2(\overline{y}_0 + \overline{y}_2)}{\overline{y}_0\overline{y}_2 - \overline{y}_1^2}, \tag{60}$$

můžeme už snadno dopočítat a_1 z

$$a_1 = \frac{1}{l}\left(\ln \frac{k_0 - \overline{y}_0}{\overline{y}_0} - \ln \frac{k_0 - \overline{y}_1}{\overline{y}_1}\right), \tag{61}$$

a hodnotu a_0 obdržíme z první z rovnic (59).

Fáze (b) - Opravíme hodnoty pomocí metody nejmenších čtverců.
Zavedeme novou parametrizaci

$$y(t) = \frac{d}{c + e^{-(a_1+\varepsilon)t}},$$

kde

$$a = a_1 + \varepsilon, \quad b = c^{-1}, \quad k = d/c. \tag{62}$$

Pro malé hodnoty ε lze psát $e^{-\varepsilon t} \sim 1 - \varepsilon t$. Tímto způsobem přecházíme k modelu

$$y_{t_i} \sim \frac{d}{c + e^{-a_1 t_i}(1 - \varepsilon t_i)}, \qquad i = 1, \cdots, n. \tag{63}$$

Po úpravě dostaneme model v lineárním tvaru

$$y_{t_i} e^{-a_1 t_i} \sim d - c y_{t_i} + \varepsilon t_i y_{t_i} e^{-a_1 t_i}. \tag{64}$$

Metodou nejmenších čtverců dostaneme nové parametry

$$\begin{pmatrix} d \\ c \\ \varepsilon \end{pmatrix} = (F^T F)^{-1} F^T Y, \tag{65}$$

kde

$$F = \begin{pmatrix} 1 & -y_1 & t_1 y_{t_1} e^{-a t_1} \\ \vdots & \vdots & \vdots \\ 1 & -y_n & t_n y_{t_n} e^{-a t_n} \end{pmatrix} \quad \text{a} \quad Y = \begin{pmatrix} y_{t_1} e^{-a_1 t_1} \\ \vdots \\ y_{t_n} e^{-a_1 t_n} \end{pmatrix}. \tag{66}$$

Nyní dosazením do reparametrizace (62) získáme opravené hodnoty parametrů.

Určení parametrů logistické křivky Newton-Raphsonovou metodou

NR metoda spočívá v minimalizaci průměrné kvadratické odchylky pozorovaných hodnot od modelu. K této minimalizaci je použito klasické Newtonovy metody.

Úkolem je tedy nalézt minimum funkce

$$S(k, b, a) = \sum_{i=1}^{n} \left(y_i - \frac{k}{1 + b e^{-a t_i}} \right)^2 \tag{67}$$

Předpokládáme, že máme k dispozici počáteční hodnoty (k_0, b_0, a_0), které jsou relativně blízko skutečným parametrům. Tyto hodnoty můžeme získat např. pomocí předcházející metody.

Myšlenka Newtonovy metody ve zkratce: $S(k, b, a)$ si rozvineme v Taylorovu řadu druhého řádu kolem bodu (k_0, b_0, a_0)

$$S(k, b, a) = S(k_0, b_0, a_0) + \nabla S(k_0, b_0, a_0)(k - k_0, b - b_0, a - a_0)^T +$$
$$\frac{1}{2}(k - k_0, b - b_0, a - a_0) \nabla^2 S(k_0, b_0, a_0)(k - k_0, b - b_0, a - a_0)^T,$$

a hledáme bod, kde $\nabla S(k, b, a) = (0, 0, 0)^T$ (bod minima S). Derivací pravé strany dostaneme rovnici

$$\nabla^2 S(k_0, b_0, a_0)(k - k_0, b - b_0, a - a_0)^T = -\nabla S(k_0, b_0, a_0),$$

která nám určuje jak spočítat novou hodnotu iterovaných parametrů ze starých.

Konkrétně tedy opakujeme výpočet odchylek mezi hodnotami parametrů (k_n, b_n, a_n) a $(k_{n+1}, b_{n+1}, a_{n+1})$ podle vzorce

$$\nabla^2 S(k_n, b_n, a_n)(k_{n+1} - k_n, b_{n+1} - b_n, a_{n+1} - a_n)^T = -\nabla S(k_n, b_n, a_n), \tag{68}$$

tak dlouho dokud nedosáhneme požadované přesnosti.

2.2 Doplňování rezerv v dávkově definovaném penzijním fondu

Zaměstnavatel má penzijní fond pro své zaměstnance, kam zaměstnanci platí dávky a odkud se jim pak vyplácí důchod. Uvažujeme spojitý model pro výši fondu.

Teoretický model

$$\frac{dV(t)}{dt} = \delta V(t) + P(t) - {}^{T}\!P(t) \tag{69}$$

$V(t)$ výše rezervy v čase t

$P(t)$ intenzita příspěvku do fondu

${}^{T}\!P(t)$ intenzita vyplácení důchodu lidem odcházejícím v čase t do penze.

δ intenzita úroku

Skutečný stav

$$\frac{dF(t)}{dt} = \delta F(t) + C(t) - {}^{T}\!P(t), \tag{70}$$

přičemž $F(0) \neq V(0)$.

$F(t)$ skutečná výše fondu

$C(t)$ intenzita příspěvků do fondu (přepočítaná na roky)

Problém: Stanovení $C(t) = P(t) + \lambda(t)(V(t) - F(t))$ resp. stanovení $\lambda(t)$.

Odečtením rovnice teoretického modelu a rovnice popisující skutečný stav dojdeme k rovnici

$$\frac{d}{dt}(V(t) - F(t)) = \delta(V(t) - F(t)) + P(t) - C(t) = (\delta - \lambda(t))(V(t) - F(t)). \tag{71}$$

Řešení této rovnice je tvaru

$$V(t) - F(t) = (V(0) - F(0)) \exp\left\{ -\int_0^t (\lambda(u) - \delta)\, du \right\}. \tag{72}$$

Označme $\overline{a}_{n\rceil} = \int_0^n e^{-\delta t} dt = (1 - e^{-\delta n})/\delta$ počáteční hodnotu jednotkového finančního toku odpovídající časovém intervalu $[0, n]$. *Chceme, aby počáteční výchylka $V(0) - F(0)$ byla splacena za n let pevnými splátkami*

$$V(0) - F(0) = \overline{a}_{n\rceil}(C(t) - P(t)) \qquad \text{pro} \qquad 0 \le t \le n. \tag{73}$$

Tento požadavek se dá přepsat do tvaru

$$C(t) = P(t) + \frac{V(0) - F(0)}{\overline{a}_{n\rceil}} \qquad \text{pro} \qquad 0 \le t \le n. \tag{74}$$

Tvrzení Pro $t \geq 0$ platí

$$\exp\left\{-\int_0^t \left(\frac{1}{\overline{a}_{\overline{n-u}|}} - \delta\right) du\right\} = \frac{\overline{a}_{\overline{n-t}|}}{\overline{a}_{\overline{n}|}}. \tag{75}$$

Důkaz: Integrand je roven

$$\frac{1}{\overline{a}_{\overline{n-u}|}} - \delta = \frac{\delta}{1-e^{-\delta(n-u)}} - \delta = \frac{\delta e^{-\delta(n-u)}}{1-e^{-\delta(n-u)}} = -\frac{d}{du}\ln(1-e^{-\delta(n-u)}). \tag{76}$$

Po integraci dostáváme

$$-\int_0^t \left(\frac{1}{\overline{a}_{\overline{n-u}}} - \delta\right) du = \ln\frac{1-e^{-\delta(n-t)}}{1-e^{-\delta n}} = \ln\frac{\overline{a}_{\overline{n-t}|}}{\overline{a}_{\overline{n}|}} \tag{77}$$

□

Rovnost (74) je tak možné upravit do následující podoby

$$C(t) = P(t) + \frac{1}{\overline{a}_{\overline{n-t}|}}\frac{\overline{a}_{\overline{n-t}|}}{\overline{a}_{\overline{n}|}}(V(0) - F(0)) \tag{78}$$

$$= P(t) + \frac{1}{\overline{a}_{\overline{n-t}|}}\exp\left\{-\int_0^t \left(\frac{1}{\overline{a}_{\overline{n-u}|}} - \delta\right) du\right\}(V(0) - F(0)) \tag{79}$$

Nyní je vidět, že požadavek (74) resp. (79) je splněn, pokud volíme $\lambda(t) := 1/\overline{a}_{\overline{n-t}|}$.

3 Teorie lineární regulace

V této kapitole se budeme zabývat modelem, kde máme soustavu (černou skříňku), která převádí vstupní posloupnost $(u_0, u_1, \cdots)$ na výstupní posloupnost $(y_0, y_1, \cdots)$ na základě následujícího lineárního modelu regulace

$$y_t = \sum_{k=0}^{t} h_k u_{t-k} = \sum_{k=0}^{\infty} h_k u_{t-k}. \tag{80}$$

Abychom mohli mít $t \in \mathbf{Z}$, položíme $u_{-i} := 0 =: y_{-i}$ pro $i \geq 1$ (interpretace: soustava je před časem 0 v klidu). Na posloupnost (u_n) se budeme dívat jako na posloupnost odchylek vstupů od jakési rovnovážné polohy a podobně na výstupní posloupnost (y_n) se budeme dívat jako na posloupnost odchylek výstupu od určité hodnoty.

Zaveďme si základní pojmy.

Definice 2 *Vstupní posloupnost* (u_n) *ve tvaru*

$$u_0 = 1, \quad 0 = u_1 = u_2 = \cdots \tag{81}$$

budeme nazývat jednotkovým impulsem. Odezvou systému na tento vstup je

$$y_0 = h_0 u_0 = h_0, \quad y_1 = h_0 u_1 + h_1 u_0 = h_1, \quad \cdots \quad y_k = h_0 u_k + \cdots + h_k u_0 = h_k \quad \cdots \tag{82}$$

výstupní posloupnost ve tvaru $(h_0, h_1, \cdots)$. *Této posloupnost* $(h_0, h_1, h_2, \cdots)$ *budeme říkat impulsní charakteristika soustavy.*

Definice 3 *Vstupní posloupnost* (u_n) *ve tvaru*

$$1 = u_0 = u_1 = \cdots \tag{83}$$

nazveme jednotkovým skokem (v čase 0 *se změní vstup z* 0 *na* 1*). Odezvě soustavy na tento skok*

$$y_t = \sum_{k=0}^{t} h_k u_{t-k} = \sum_{k=0}^{t} h_k =: g_t \tag{84}$$

budeme říkat přechodová charakteristika soustavy.

K analýze soustav se nám bude hodit metoda vytvořujících funkcí pro posloupnosti, speciálně z-transformace.

Definice 4 *Je-li* $(a_k, k \in \mathbb{N}_0)$ *posloupnost (reálných) čísel, definujeme její z-transformaci následujícím způsobem*

$$A(z) = \sum_{k=0}^{\infty} a_k z^{-k}. \tag{85}$$

Tato řada konverguje pro $|z| > z_0$ a diverguje pro $|z| < z_0$ pro nějaké $z_0 \in [0, \infty]$.

Pro posloupnost vstupů (u_n) výstupů (y_n) a pro impulsní charakteristiku (h_k) definujeme $U(z), Y(z), H(z)$ jako příslušné z-transformace.

Definice 5 *Funkci* $H(z) = \sum_{k=0}^{\infty} h_k z^{-k}$ *budeme nazývat (impulsní) přenosovou funkcí soustavy.*

Proč se H nazývá přenosovou funkcí soustavy je vidět z následujícího pozorování:

Platí

$$Y(z) = \sum_{t=0}^{\infty} y_t z^{-t} = \sum_{t=0}^{\infty} \left(\sum_{k=0}^{\infty} h_k u_{t-k} \right) z^{-t} = \sum_{k=0}^{\infty} h_k z^{-k} \sum_{l=0}^{\infty} u_l z^{-l} = H(z)U(z) \tag{86}$$

pro taková z, že $H(z), U(z)$ absolutně konvergují, neboť $\cdots = u_{-2} = u_{-1} = 0$.

3.1 Stavový model

Příklad – bakaláři

Pro potřeby stavového modelu zavedeme posloupnost $(x_t \in \mathbb{R}^r, t \in \mathbb{N}_0)$ stavů systému (tedy x_t je obecně r-rozměrný vektor). Stav systému v čase 0 definujeme hodnotou $x_0 = \mathbf{0} \in \mathbb{R}^r$. Stavy systému v dalších okamžicích definujeme pomocí lineárního modelu

$$x_{t+1} = ax_t + bu_t, \quad \text{kde } a \in \mathbb{R}^{r \times r}, b \in \mathbb{R}^{r \times 1}. \tag{87}$$

Výstup systému pak budeme uvažovat ve tvaru

$$y_t = cx_t + du_t, \quad \text{kde } c \in \mathbb{R}^{1 \times r}, d \in \mathbb{R}^{1 \times 1}. \tag{88}$$

Ajak získáme ze stavového modelu přímou závislost výstupní posloupnosti na vstupní ve tvaru (80) ?

Vynásobením obou rovností faktorem z^{-t} a sečtením přes $t \in \mathbb{N}_0$ dojdeme k rovnostem

$$\sum_{t=0}^{\infty} x_{t+1} z^{-t} = a \sum_{t=0}^{\infty} x_t z^{-t} + b \sum_{t=0}^{\infty} u_t z^{-t} \tag{89}$$

$$\sum_{t=0}^{\infty} y_t z^{-t} = c \sum_{t=0}^{\infty} x_t z^{-t} + d \sum_{t=0}^{\infty} u_t z^{-t}. \tag{90}$$

Tyto rovnosti lze zapsat pomocí z-transformace $X(z) = \sum_{t=0}^{\infty} x_t z^{-t}$

$$zX(z) = aX(z) + bU(z) \tag{91}$$

$$Y(z) = cX(z) + dU(z). \tag{92}$$

Řešením těchto rovnic docházíme k rovnostem

$$X(z) = (zI - a)^{-1} bU(z) \tag{93}$$

$$Y(z) = cX(z) + dU(z) = [c(zI - a)^{-1} b + d]U(z), \tag{94}$$

kde $I \in \mathbb{R}^{r \times r}$ je r-rozměrná jednotková matice tedy $H(z) = [c(zI - a)^{-1} b + d]$.. Tuto rovnost můžeme dále upravit

$$H(z) = [c(zI - a)^{-1} b + d] = \frac{c \cdot \mathrm{adj}(zI - a) \cdot b + d \cdot \det(zI - a)}{\det(zI - a)} = \tag{95}$$

$$= \frac{N(z)}{D(z)} = \frac{n_0 z^r + \cdots + n_r}{d_0 z^r + \cdots + d_r} = \frac{n_0 + \cdots + n_r z^{-r}}{d_0 + \cdots + d_r z^{-r}} = \frac{n(z^{-1})}{d(z^{-1})}, \tag{96}$$

kde adj značí adjungovanou matici a $N(z)$ a $D(z)$ jsou z-transformace odpovídající polynomům $n(z^{-1})$ a $d(z^{-1})$.

3.2 Obecný model

Opět budeme předpokládat, že $u_{-i} = 0 = y_{-i}$ pro $i \in \mathbb{N}$. Dále budeme předpokládat, že výstup systému y_t v čase t je dán následující rovnicí

$$d_0 y_t + \cdots + d_r y_{t-r} = n_0 u_t + \cdots + n_r u_{t-r}, \quad \text{kde } d_0 \neq 0, \tag{97}$$

a budeme chtít opět najít vyjádření ve tvaru (80).

Vynásobením rovnosti (97) hodnotou z^{-t} a sečtením přes $t \in \mathbb{N}_0$ obdržíme rovnost

$$\sum_{t=0}^{\infty} (d_0 z^{-0} y_t z^{-t} + \cdots + d_r z^{-r} y_{t-r} z^{-(t-r)}) = \sum_{t=0}^{\infty} (n_0 z^{-0} u_t z^{-t} + \cdots + n_r z^{-r} u_{t-r} z^{-(t-r)}) \tag{98}$$

Využitím předpokladu $y_{-i} = 0 = u_{-i}$ pro $i \in \mathbb{N}$ docházíme k rovnosti

$$(d_0 z^0 + \cdots + d_r z^{-r}) \sum_{k=0}^{\infty} y_k z^{-k} = (n_0 z^0 + \cdots + n_r z^{-r}) \sum_{k=0}^{\infty} u_k z^{-k}, \tag{99}$$

to jest

$$d(z^{-1})Y(z) = n(z^{-1})U(z), \tag{100}$$

kde

$$d(x) = d_0x^0 + \cdots d_rx^r \quad \text{a } n(x) = n_0x^0 + \cdots n_rx^r. \tag{101}$$

Protože je tento obecný model speciálním případem lineárního modelu, ve kterém $y_t = \sum_{k=0}^{t} h_k u_{t-k}$ platí i tady vztah

$$Y(z) = H(z)U(z) \text{ pro } z \text{ taková, že } U, H \text{ konvergují absolutně.} \tag{102}$$

Docházíme tak k rovnosti

$$d(z^{-1})Y(z) = d(z^{-1})H(z)U(z) = n(z^{-1})U(z). \tag{103}$$

Tato rovnost platí pro každý vstup a z takové, že řady H, U konvergují absolutně.

Speciální volba vstupu $u_0 = 1$, $0 = u_1 = u_2 = \cdots$ odpovídající jednotkovému impulsu má z-transformaci ve tvaru $U(z) \equiv 1$. Dosazením této funkce do (103) přicházíme k rovnosti

$$d(z^{-1})H(z) = n(z^{-1}), \tag{104}$$

tedy

$$H(z) = \frac{n(z^{-1})}{d(z^{-1})} = \frac{n_0 + \cdots + n_rz^{-r}}{d_0 + \cdots + d_rz^{-r}}. \tag{105}$$

Poznámka: Rovnice $d_0y_t + \cdots + d_ry_{t-r} = n_0u_t + \cdots + n_ru_{t-r}$ se někdy zapisuje ve tvaru

$$(d_0z^0 + \cdots + d_rz^{-r})y_t = (n_0z^0 + \cdots + n_rz^r)u_t, \text{ zkráceně } d(z^{-1})y_t = n(z^{-1})u_t, \tag{106}$$

přičemž z se v této souvislosti interpretuje jako operátor posunutí, tj, $zx_t = x_{t+1}$. Inverzní operátor $z^{-1}x_t = x_{t-1}$ se pak interpretuje jako operátor zpětného posunutí.

Příklad: Nechť pro výstup platí rovnice

$$y_{t+1} - ay_t = u_{t+1} \quad \text{resp.} \quad (1 - az^{-1})y_t = u_t.$$

Odtud

$$H(z) = \frac{1}{1 - az^{-1}} = \frac{z}{z - a}$$

má pól v bodě $z = a$. Funkci H lze rozvinout následujícím způsobem

$$H(z) = \frac{1}{1 - az^{-1}} = \sum_{k=0}^{\infty} a^k z^{-k}, \tag{107}$$

což odpovídá impulsní charakteristice $h_k = a^k$. $\square$

Základní vlastností, která se u lineárních soustav zkoumá, je stabilita soustavy a vlastnosti se stabilitou související. Stabilita spočívá v tom, že soustava vychýlená z rovnovážného stavu (= nulového stavu) se dostane zpátky do tohoto stavu.

Definice 6 *Řekneme, že soustava je stabilní, pokud* $\lim_{t\to\infty} h_t = 0$. *(Tj. pokud odezva na jednotkový impuls se s rostoucím časem vytrácí.)*

Příklad - pokračování Je-li $h_k = a^k$ je soustava stabilní právě tehdy, když je $|a| < 1$. $\square$

Stabilita soustavy závisí na pólech přenosové funkce H.

Věta 4 *Soustava odpovídající obecnému modelu s přenosovou funkcí ve tvaru*

$$H(z) = \frac{n(z^{-1})}{d(z^{-1})} = \frac{N(z)}{D(z)}, \tag{108}$$

kde $N(z) = n_0 z^r + \cdots + n_r$ a $D(z) = d_0 z^r + \cdots + d_r$ jsou nesoudělné polynomy, je stabilní právě tehdy, když všechny kořeny polynomu $D(z)$ se nacházejí uvnitř jednotkového kruhu v komplexní rovině.

Důkaz: Dokážeme pouze jednu implikaci (že z kořenů uvnitř jednotkového kruhu plyne stabilita), druhá je příliš obtížná.

Podle základní věty algebry je možné polynom $D(z)$ rozložit na součin kořenových faktorů

$$D(z) = (z - z_1)^{m_1} \cdots (z - z_k)^{m_k}, \quad \text{kde } m_1 + \cdots + m_k = r \tag{109}$$

a kde $z_1, \cdots, z_k$ jsou kořeny polynomu $D(z)$ s násobnostmi $m_1, \cdots, m_k$. Na základě tohoto rozkladu jsme schopni racionální funkci H rozložit na součet konstanty a součet parciálních zlomků:

$$H(z) = h_0 + \sum_{i=1}^{k} \sum_{j=1}^{m_i} \frac{c_{ij}}{(z - z_i)^j} = h_0 + \sum_{i=1}^{k} \sum_{j=1}^{m_i} \frac{c_{ij}}{z^j} \sum_{t=0}^{\infty} \left(\begin{array}{c} t+j-1 \\ j-1 \end{array} \right) \left(\frac{z_i}{z} \right)^t \tag{110}$$

Podílovým kritériem pak zjistíme, že každá tato řada (uvažovaná jako mocninná řada s argumentem z^{-1}) konverguje absolutně pro $|z| = 1$. Stejným způsobem tedy konverguje i její součet $H(z) = \sum_{l=0}^{\infty} h_l z^{-l}$ pro $z = 1$. Nutnou podmínkou pro konvergenci řady $\sum_{l=0}^{\infty} h_l$ je ale podmínka $\lim_{l\to\infty} h_l = 0$, což je právě podmínka stability. $\square$

Poznámka: Pro stavový model máme z rovnice (95), že $D(z) = \det(zI - a)$ a tedy kořeny $D(z)$ jsou právě všechna vlastní čísla matice a. Soustava je tedy stabilní právě tehdy, když všechna vlastní čísla matice a jsou v absolutní hodnotě menší než 1.

Poznámka: Nechť je soustava stabilní, nechť vstupem je jednotkový skok $1 = u_0 = u_1 = \cdots$. Pak

$$\lim_{t\to\infty} y_t = \lim_{t\to\infty} \sum_{k=0}^{t} h_k u_{t-k} = \lim_{t\to\infty} h_k = \sum_{k=0}^{\infty} h_k = y_\infty. \tag{111}$$

Soustava se tak v nekonečném horizontu ustálí na nové úrovni. To vychází z předchozí věty, neboť, pokud impulsní charakteristika konverguje k nule (soustava je stabilní), činí tak geometricky rychle. Speciálně, v důkazu jedné implikace předcházející věty je ukázáno: pokud jsou kořeny $D(z)$ uvnitř jednotkového kruhu (podmínka stability podle věty), pak řada $\sum_{l=0}^{\infty} h_l$ konverguje absolutně.

4 Markovovy řetězce

Definice 7 *Posloupnost celočíselných náhodných veličin* $\{X_n\}_{n=0}^{\infty}$ *se nazývá Markovův řetězec, jestliže*

$$P[X_{n+1} = j | X_n = i_n, \cdots, X_0 = i_0] = P[X_{n+1} = j | X_n = i_n] \tag{112}$$

pro každé $n \geq 0$, *a* $i_0, \cdots, i_n \in \mathbb{Z}$ *takové, že* $P(X_n = i_n, \cdots, X_0 = i_0) > 0$. *Hodnoty* n *interpretujeme jako časové okamžiky, hodnoty, kterých řetězec* X_n *může nabývat, interpretujeme jako stavy.*

Vztah (112) vyjadřuje markovskou vlastnost, tj. fakt, že výsledek v čase $n+1$ závisí pouze na stavu řetězce v čase n (stav v přítomném čase) nikoli na celé minulosti - tj. posloupnosti realizovaných stavů před časem n.

Příklady – hody kostkou, náhodná procházka

Definice 8 *Hodnotu*

$$p_{ij}(n, n+1) := P[X_{n+1} = j | X_n = i]$$

budeme nazývat pravděpodobnost přechodu ze stavu i do stavu j v čase $n+1$. *Markovův řetězec se nazývá homogenní, pokud hodnoty* $p_{ij} := p_{ij}(n, n+1)$ *nezávisí na hodnotě* n.
Označme $p_i := P(X_0 = i)$, *pak* $(p_i, i \in \mathbb{Z})$ *označuje počáteční rozdělení řetězce.*

Nadále budeme předpokládat, že daný Markovův řetězec je homogenní.

Věta 5 *Nechť* $\{X_n\}$ *je homogenní Markovův řetězec s množinou stavů* $S \subseteq \mathbb{Z}$, *počátečním rozdělením* $(p_i, i \in S)$ *a maticí přechodu* $(p_{ij}, i, j \in S)$. *Pak*

$$P(X_0 = i_0, \cdots, X_n = i_n) = p_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}.$$

Důkaz: Indukcí podle n. Pro $n = 0, 1$ tvrzení zřejmě platí. Pro $n = 2$ máme podle věty o podmíněné pravděpodobnosti

$$\begin{aligned} P(X_0 = i_0, X_1 = i_1, X_2 = i_2) &= P(X_0 = i_0) P[X_1 = i_1 | X_0 = i_0] P[X_2 = i_2 | X_1 = i_1, X_0 = i_0] \\ &= p_{i_0} p_{i_0 i_1} p_{i_1 i_2}, \end{aligned}$$

protože z markovské vlastnosti platí

$$P[X_2 = i_2 | X_1 = i_1, X_0 = i_0] = P[X_2 = i_2 | X_1 = i_1] = p_{i_1 i_2}.$$

A pro $n = m$ obdobně

$$\begin{aligned} P(X_0 = i_0, \cdots, X_n = i_n) &= P[X_n = i_n | X_0 = i_0, \cdots, X_{n-1} = i_{n-1}] P(X_0 = i_0, \cdots, X_{n-1} = i_{n-1}) \\ &= p_{i_{n-1} i_n} P(X_0 = i_0, \cdots, X_{n-1} = i_{n-1}) = p_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}, \end{aligned}$$

kde jsme nejdříve využili markovské vlastnosti a poté indukčního předpokladu. $\square$

Definujme si také pravděpodobnosti přechodu vyšších řádů (tj. po více krocích)

$$p_{ij}^{(0)} = I\{i = j\}, \quad p_{ij}^{(1)} = p_{ij}, \quad p_{ij}^{(n+1)} = \sum_{k \in S} p_{ik}^{(n)} p_{kj},$$

a matice pravděpodobností přechodu po jednom či více krocích $\mathbf{P} = (p_{ij}, i, j \in S)$ a $\mathbf{P}^{(n)} = (p_{ij}^{(n)}, i, j \in S)$. Že je to definice smysluplná ukazuje následující tvrzení.

Věta 6 *Nechť $m \geq$ je celé číslo, $i, j \in S$, pak $P[X_{m+n} = j | X_n = i] = p_{ij}^{(m)}$.*

Důkaz: Indukcí podle m. □

Rovněž platí $\mathbf{P}^{(n)} = \mathbf{P}^n$ pro $n \geq 0$ (důkaz opět indukcí).

Poznámka: Uvědomme si, že jak $\mathbf{P}$ tak i $\mathbf{P}^{(n)}$ jsou tzv. stochastické matice, tj. obsahují pouze čísla mezi 0 a 1 a součet každého řádku se rovná 1.

A jak se pravděpodobnosti přechodu skládají ukazuje následující věta.

Věta 7 *(Chapmanova-Kolmogorova rovnost) Platí*

$$p_{ij}^{(m+n)} = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)},$$

pro všechna celá $m, n \geq 0$.

Důkaz: Použijeme rozklad uvažovaného jevu podle všech možných stavů v čase k a postupné podmiňování

$$\begin{aligned} p_{ij}^{(m+n)} &= P[X_{m+n} = j | X_0 = i] = \sum_{k \in S} P[X_{m+n} = j, X_m = k | X_0 = i] \\ &= \sum_{k \in S} P[X_{m+n} = j | X_m = k, X_0 = i] P[X_m = k | X_0 = i] = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}, \end{aligned}$$

a v poslední rovnosti markovskou vlastnost a homogenitu řetězce. □

4.1 Klasifikace stavů Markovova řetězce

Definice 9 *Řekneme, že stav $j \in S$ je dosažitelný ze stavu $i \in S$, pokud existuje $n \geq 0$ takové, že $p_{ij}^{(n)} > 0$. Píšeme $i \to j$. Markovův řetězec nazveme nerozložitelným, pokud je každý stav dosažitelný z každého stavu.*

Definice 10 *Řekneme, že $\emptyset \neq C \subset S$ je uzavřená množina, pokud žádný stav vně množiny C není dosažitelný ze žádného stavu uvnitř množiny C.*

Věta 8 *C je uzavřená množina právě tehdy, když platí $p_{jk} = 0$ pro každé $j \in C$ a $k \notin C$.*

Důkaz: Implikace zleva doprava je zřejmá. Opačnou implikaci dokážeme indukcí. Musíme ukázat, že $p_{jk}^{(l)} = 0$ pro každé $l \in \mathbf{N}$, $j \in C$ a $k \notin C$. Pro $l = 1$ to platí z předpokladu, pro $l > 1$ pišme

$$p_{jk}^{(l)} = \sum_{i \in S} p_{ji}^{(l-1)} p_{ik} = \sum_{i \in C} p_{ji}^{(l-1)} p_{ik} + \sum_{i \notin C} p_{ji}^{(l-1)} p_{ik} = \sum_{i \in C} p_{ji}^{(l-1)} 0 + \sum_{i \notin C} 0 p_{ik} = 0,$$

kde jsme v předposlední rovnosti využili indukční předpoklad. □

Věta 9 *Řetězec je rozložitelný právě tehdy, když existuje uzavřená množina $C \neq S$.*

Důkaz: Zřejmý.

Věta 10 *Markovův řetězec je rozložitelný právě tehdy, když po vhodném přečíslování stavů je matice přechodu tvaru* $\begin{pmatrix} P_1 & 0 \\ Q & R \end{pmatrix}$, *kde* P_1, R *jsou čtvercové matice.*

Důkaz: Přečíslujeme stavy tak, aby nejdříve byly očíslované stavy z C. Pak snadno nahlédneme, že nová matice pravděpodobností přechodu bude mít přesně žádaný tvar. $\square$

Definice 11 *Buď* $i \in S$, *potom*

$$\nu_i = \inf\{n > 0, X_n = i\}$$

se nazývá doba prvního návratu řetězce do stavu i. Stav $i \in S$ *se nazývá trvalý, pokud*

$$P_i(\nu_i < \infty) = P(\nu_i < \infty | X_0 = i) = 1,$$

tj. řetězec, který vychází ze stavu i se s pravděpodobností 1 vrátí do stavu i po konečně mnoha krocích. V opačném případě (tedy, když je kladná pravděpodobnost, že se řetězec do stavu i nikdy nevrátí), řekneme, že stav i je přechodný.
Trvalý stav $i \in S$ *se nazývá nulový, pokud* $E_i\nu_i = \infty$ *a nenulový, pokud* $E_i\nu_i < \infty$.

Poznámka: Do trvalého stavu se Markovův řetězec s pravděpodobností 1 vrátí nekonečně krát.

Označme $f_i^{(n)} = P_i(\nu_i = n)$, a $f_i^{(0)} = 0$. Zřejmě stav i je trvalý právě tehdy, když $\sum_{n=1}^{\infty} f_i^{(n)} = 1$, a $E_i\nu_i = \sum_{n=1}^{\infty} n f_i^{(n)}$.

Tedy střední hodnota doby prvního návratu Markovského řetězce do trvalého stavu i je konečná, právě když je stav nenulový.

Poznámka: $p_{ii}^{(n)} = P_i(X_n = i) = \sum_{k=1}^{n} f_i^{(k)} p_{ii}^{(n-k)}$ platí pro $n \geq 1$.

Definice 12 *Stav* $i \in S$ *homogenního Markovova řetězce se nazývá periodický s periodou* $d > 1$, *pokud*

$$d = NSD\{n > 0, p_{ii}^{(n)} > 0\} > 1.$$

Pokud vychází $d = 1$, *říkáme, že stav* $i \in S$ *je aperiodický/neperiodický.*
Řekneme, že dva stavy jsou stejného typu, pokud jsou oba dva trvalé/přechodné, nulové/nenulové, neperiodické/periodické se stejnou periodou.

Věta 11 *Stav* $i \in S$ *je trvalý právě tehdy, když* $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$.

Trvalý stav $i \in S$ *je nulový právě tehdy, když* $p_{ii}^{(n)} \to 0$ *pro* $n \to \infty$.

Důkaz: Pro $x \in (0,1)$ definujme vytvořující funkce $P(x)$, $F(x)$ posloupností $p_{ii}^{(n)}, f_i^{(n)}$

$$P(x) := \sum_{n=0}^{\infty} p_{ii}^{(n)} x^n, \quad \text{a} \quad F(x) := \sum_{n=0}^{\infty} f_i^{(n)} x^n.$$

Z poznámky plyne, že pro $n \geq 1$ platí

$$p_{ii}^{(n)} x^n = \sum_{k=1}^{n} f_i^{(k)} p_{ii}^{(n-k)} x^n = \sum_{k=1}^{n} f_i^{(k)} x^k p_{ii}^{(n-k)} x^{n-k}$$

Sečtením této rovnosti přes $n \geq 1$ dostaneme

$$P(x) - 1 = \sum_{n=1}^{\infty} p_{ii}^{(n)} x^n = \sum_{n=1}^{\infty} \sum_{k=1}^{n} f_i^{(k)} x^k p_{ii}^{(n-k)} x^{n-k} = \left(\sum_{k=1}^{\infty} f_i^{(k)} x^k \right) \left(\sum_{l=0}^{\infty} p_{ii}^{(l)} x^l \right) = F(x)P(x).$$

Platí tedy $P(x) = 1/(1 - F(x))$. Dále

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \lim_{x \to 1^-} P(x) = \lim_{x \to 1^-} \frac{1}{1 - F(x)} = \infty$$

nastává právě tehdy, když

$$1 = \lim_{x \to 1^-} F(x) = \sum_{n=0}^{\infty} f_i^{(n)} = P_i(\nu_i < \infty).$$

Odtud je vidět, že stav $i \in S$ je trvalý právě tehdy, když $\sum_{i=0}^{\infty} p_{ii}^{(n)} = \infty$.

Náznak důkazu druhé části* : Zbylou část důkazu pouze naznačíme. Zřejmě

$$\begin{aligned}\frac{1 - F(x)}{1 - x} &= \frac{1}{1 - x} \left(\sum_{n=0}^{\infty} f_i^{(n)} - \sum_{n=0}^{\infty} f_i^{(n)} x^n \right) = \sum_{n=0}^{\infty} f_i^{(n)} \frac{1 - x^n}{1 - x} \\ &= \sum_{n=0}^{\infty} f_i^{(n)} (1 + \cdots + x^{n-1}) \to \sum_{n=0}^{\infty} f_i^{(n)} n = E_i \nu_i, \end{aligned} \tag{113}$$

pro $x \to 1^-$. Limitní přechod v (113) dostaneme z Fatouova lemmatu

$$\liminf_{x \to 1^-} \sum_{n=0}^{\infty} f_i^{(n)} (1 + \cdots x^{n-1}) \geq \sum_{n=0}^{\infty} f_i^{(n)} n,$$

a z nerovnosti $f_i^{(n)}(1 + \cdots + x^{n-1}) \leq f_i^{(n)} n$ pro $x \in (0, 1)$. Víme tedy, že trvalý stav $i \in S$ je nulový právě tehdy, když

$$\lim_{x \to 1^-} (1 - x) P(x) = \lim_{x \to 1^-} \frac{1 - x}{1 - F(x)} = 0.$$

Zbývá tedy ukázat, že $(\lim_{x \to 1^-} (1 - x)P(x) = 0)$ je ekvivalentní $(p_{ii}^{(n)} \to 0)$ pro $n \to \infty$. Jednu implikaci lze obdržet z Tauberovy věty, druhá je o něco (hodně) složitější. Podrobnosti a odkazy na další literaturu lze najít ve skriptech Z. Práškové a P. Lachouta: Základy náhodných procesů. □

Věta 12 *Nechť $i \to j \to i$. Pak stavy i a j jsou stejného typu. Pokud $i \to j \not\to i$, pak stav i je přechodný. V nerozložitelném řetězci jsou všechny stavy stejného typu.*

Věta 13 *V řetězci s konečně mnoha stavy nemohou být všechny stavy přechodné. V řetězci s konečně mnoha stavy neexistují stavy trvalé nulové.*

Definice 13 *Stav, ze kterého není dosažitelný žádný jiný stav, nazveme stavem absorbčním. Stav, který je trvalý, nenulový a neperiodický, nazveme stavem ergodickým.*

4.2 Limitní a stacionární rozdělení Markovova řetězce

V této podkapitole se budeme zabývat otázkou, co se děje, když $n \to \infty$, neboli když necháme Markovský řetězec dlouho běžet, tak jaké jsou vlastnosti marginálního rozdělení veličiny X_n.

Definice 14 *Rozdělení $\pi = (\pi_i, i \in S)$ nazveme stacionárním, pokud platí $\pi_j = \sum_{i \in S} \pi_i p_{ij}$ pro všechna $j \in S$. Maticově lze tuto soustavu rovností zapsat ve tvaru $\pi^T = \pi^T \mathbf{P}$.*
Rozdělení $\mathbf{a} = (a_i, i \in S)$ na S nazveme limitním, pokud pro každé $i \in S$ a každé počáteční rozdělení platí $a_i = \lim_{n\to\infty} p_i(n)$, kde $p_i(n) = P(X_n = i)$.

Poznámka: Všimněte si, že na rozdíl od stacionárního rozdělení, které se vztahuje k pravděpodobnostem přechodu, se limitní rozdělení vztahuje k rozdělení celého Markovského řetězce (jež je ovlivněno i počátečním rozdělením). Nicméně snadno nahlédneme, že je-li počáteční rozdělení Markovského řetězce rovno stacionárnímu, pak i rozdělení řetězce ve všech následujících časech $(p_i(n), i \in \mathbf{Z}), n \in \mathbf{N}$ je rovno stacionárnímu rozdělení.

Věta 14 *Nechť existuje limitní rozdělení $\mathbf{a}$ řetězce $\{X_n\}$, pak toto rozdělení je jeho stacionárním rozdělením.*

Důkaz: Zřejmě

$$p_j(n+1) = \sum_{i \in S} p_i(n) p_{ij}.$$

Podle Fatouova lemmatu

$$a_j = \lim_{n\to\infty} p_j(n+1) = \liminf_{n\to\infty} \sum_{i\in S} p_i(n) p_{ij} \geq \sum_{i\in S} \liminf_{n\to\infty} p_i(n) p_{ij} = \sum_{i \in S} a_i p_{ij}. \tag{114}$$

Posčítáme-li předchozí rovnosti přes $j \in S$, dojdeme k nerovnosti

$$1 = \sum_{j\in S} a_j \geq \sum_{i,j \in S} a_i p_{ij} = \sum_{i \in S} a_i \sum_{j \in S} p_{ij} = \sum_{i\in S} a_i = 1.$$

Tato nerovnost je tak ve skutečnosti rovností, což ve skutečnosti znamená, že ve všech nerovnostech (114) musí nastávat rovnost. $\square$

Věta 15 *V nerozložitelném řetězci existuje stacionární rozdělení právě tehdy, když všechny stavy jsou trvalé nenulové. V tom případě je existuje právě jedno stacionární rozdělení. Jsou-li navíc stavy řetězce neperiodické je stacionární rozdělení i rozdělením limitním, dokonce*

$$\lim_{n\to\infty} p_j(n) = \lim_{n\to\infty} p_{ij}^{(n)} = \pi_j = \frac{1}{E_j \nu_j},$$

kde $\pi = (\pi_i, i \in S)'$ je stacionární rozdělení a $E_j \nu_j = \sum_{n=1}^{\infty} n f_j(n)$ je střední hodnota doby prvního návratu do stavu j. Jsou-li všechny stavy periodické, pak platí

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_j(k) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)} = \pi_j,$$

Poznámka: V konečném nerozložitelném řetězci stacionární rozdělení vždy existuje.

Stacionární rozdělení nerozložitelného řetězce ovšem necharakterizuje jen limity absolutních pravděpodobností jednotlivých stavů. Charakterizuje také četnosti návratu do jednotlivých stavů (resp. čas řetězcem v jednotlivých stavech strávený) – jak ukazuje následující věta.

Věta 16 *Označme*

$$N_j(n) = \sum_{k=1}^{n} I(X_n = j), \qquad j \in S,$$

náhodnou veličinu udávající, kolikrát v prvních n krocích projde řetězec stavem j. V nerozložitelném řetězci s trvalými nenulovými stavy platí

$$\lim_{n\to\infty} \frac{N_j(n)}{n} = \pi_j \quad \textit{s pravděpodobností 1,}$$

pro každé $j \in S$.

4.3 Systém bonus-malus v pojišťovnictví

Používá se v havarijním pojištění. Systém má určitý počet tříd výše pojistného (bonusové třídy), do kterých je pojištěnec přiřazován na základě počtu škod uplatněných za pojistné období. Cílem připoužití tohoto systému je snížení heterogenity pojistného kmene, odrazení pojištěnců od uplatnění nízkých nároků a umožnění stanovení pojistného, jehož výše lépe odpovídá individuálním rizikům.

Bonusová třída pojištěnce a počet škod v daném období určí bonusovou třídu a tím i výši pojistného v dalším období $\longrightarrow$ modelem je homogenní Markovský řetězec.

Označme $S = \{1, \cdots, k\}$ bonusové třídy a $a = (a_1, \cdots, a_k)'$ výše pojistného jako část základu. Řetězec $\{X_n\}$ bude v čase n ve stavu i, pokud bude pojištěnec zařazen do i-té bonusové třídy. k_0 budeme označovat třídu, do které je pojištěnec zařazen na začátku (tj. v době příchodu do systému).

Španělský systém
$k = 5, a = \frac{1}{100}(60, 70, 80, 90, 100)', k_0 = 5$. Pokud pojištěnec bude uplatňovat nárok na výplatu v předcházejícím období, přesune se v následujícím kroku do 5-té třídy, pokud nebude, posune se o třídu níže, je-li to možné. Nechť pravděpodobnost uplatnění nároku je 0.1. Pak matice přechodu mezi třídami je tvaru

$$P = \begin{pmatrix} 0.9 & 0 & 0 & 0 & 0.1 \\ 0.9 & 0 & 0 & 0 & 0.1 \\ 0 & 0.9 & 0 & 0 & 0.1 \\ 0 & 0 & 0.9 & 0 & 0.1 \\ 0 & 0 & 0 & 0.9 & 0.1 \end{pmatrix}$$

Předpokládáme, že individuální riziko pojištěnce je charakterizováno rizikovým parametrem λ tak, že počet pojistných událostí j má Poissonovo rozdělení $p_j = \frac{\lambda^j}{j!}e^{-\lambda}, j = 0, 1, 2, \cdots$.. O výši škod Y_l budeme předpokládat, že jsou to nezávislé, stejně rozdělené náhodné veličiny, nezávislé na hodnotě parametru λ.

Britský systém
$k = 7, a = \frac{1}{100}(33, 40, 45, 55, 65, 75, 100)', k_0 = 6$ Pokud v minulém období (roce) pojištěnec neuplatňoval nárok na výplatu, posune se o jednu třídu níže, je-li to možné.

Nechť naopak v minulém období uplatňoval jednou nárok na výplatu. Pokud byl v první třídě, posune se o 3 třídy výše. Pokud byl v druhé nebo třetí, posune se o dvě třídy. V ostatních případech se posune o jednu třídu výše, je-li to možné. Za každý další nárok na výplatu (druhý a výše) se pojištěnec posune ještě o dvě třídy výše. Pokud není možné, aby se pojištěnec posunul o předepsaný počet tříd výše, skončí v poslední třídě.

Matice přechodu tak bude tvaru

$$P = \begin{pmatrix} e^{-\lambda} & 0 & 0 & \lambda e^{-\lambda} & 0 & \frac{\lambda^2}{2}e^{-\lambda} & 1-e^{-\lambda}(1+\lambda+\frac{\lambda^2}{2}) \\ e^{-\lambda} & 0 & 0 & \lambda e^{-\lambda} & 0 & \frac{\lambda^2}{2}e^{-\lambda} & 1-e^{-\lambda}(1+\lambda+\frac{\lambda^2}{2}) \\ 0 & e^{-\lambda} & 0 & 0 & \lambda e^{-\lambda} & 0 & 1-e^{-\lambda}(1+\lambda) \\ 0 & 0 & e^{-\lambda} & 0 & \lambda e^{-\lambda} & 0 & 1-e^{-\lambda}(1+\lambda) \\ 0 & 0 & 0 & e^{-\lambda} & 0 & \lambda e^{-\lambda} & 1-e^{-\lambda}(1+\lambda) \\ 0 & 0 & 0 & 0 & e^{-\lambda} & 0 & 1-e^{-\lambda} \\ 0 & 0 & 0 & 0 & 0 & e^{-\lambda} & 1-e^{-\lambda} \end{pmatrix}$$

Poznámka: Nechť matice pravděpodobností přechodu má kladný sloupec, potom existuje stacionární rozdělení $\pi = (\pi_i, i \in S)$, přičemž $\pi_j = \lim_{n\to\infty} p_{ij}^{(n)}$.

Střední hodnotu pojistného EC určíme pomocí stacionárního rozdělení. EC je rovna (vzhledem k základnímu pojistnému odpovídajícímu číslu 1)

$$EC = \pi_1 a_1 + \cdots + \pi_k a_k$$

Příklad: Španělský model, pravděpodobnost, že nedošlo ke škodě je $a = e^{-\lambda}$, a tedy pravděpodobnost uplatnění nároku je $1 - a = 1 - e^{-\lambda}$. Matice pravděpodobností přechodu je tvaru

$$P = \begin{pmatrix} a & 0 & 0 & 0 & 1-a \\ a & 0 & 0 & 0 & 1-a \\ 0 & a & 0 & 0 & 1-a \\ 0 & 0 & a & 0 & 1-a \\ 0 & 0 & 0 & a & 1-a \end{pmatrix}$$

Stacionární rozdělení je potom rovno

$$\pi = (a^4, a^3(1-a), a^2(1-a), a(1-a), 1-a)^T.$$

V modelu můžeme spočítat i další charakteristiky. Např.

Pravděpodobnost uplatnění pojistné události
Předpokládejme, že k uplatnění nastalé pojistné události dochází, pokud je škoda Y vyšší, než je odpovídající navýšení pojistného c v dalším roce po uplatnění. Tedy

$$P(\text{uplatnění}) = P(\text{uplatnění}|\text{nastává})P(\text{nastává}) = P(Y > c)\left(1 - e^{-\lambda}\right),$$

kde $1 - e^{-\lambda}$ je pravděpodobnost výskytu pojistné události. Pokud předpokládáme, že výše škody má logaritmicko-normální rozdělení s parametry μ a $\sigma^2 > 0.$, tj. $\log Y \sim N(\mu, \sigma^2)$. Pak

$$P(Y > c) = P\left(\frac{\ln Y - \mu}{\sigma} > \frac{\ln c - \mu}{\sigma}\right) = 1 - \Phi\left(\frac{\ln c - \mu}{\sigma}\right).$$

Výsledná pravděpodobnost uplatnění je pak rovna

$$\left[1 - \Phi\left(\frac{\ln c - \mu}{\sigma}\right)\right]\left(1 - e^{-\lambda}\right).$$

Úplný model

Nebo můžeme model rozšířit tak, aby zahrnoval různou individuální rizikovost chování jednotlivých pojištěnců, tedy každému pojištěnci odpovídá obecně různý parametr $\lambda \in \Lambda$ a individuální matice pravděpodobností přechodu $P(\lambda)$ na něm závisí. Obvykle se předpokládá, že rizikové parametry λ mají Γ-rozdělení popsané hustotou

$$f(\lambda) = \frac{\tau^h}{\Gamma(h)}\lambda^{-\tau h}, \quad \lambda > 0,$$

kde h a τ jsou parametry. Ty se pro potřeby modelu odhadnou z dat o škodách pomocí momentové metody, protože střední hodnota Γ-rozdělení je rovna $\frac{h}{\tau}$ a rozptyl $\frac{h}{\tau^2}$.

Pojištěnce máme ještě rozvrstvené podle doby, po kterou jsou pojištění. Označme příslušné relativní četnosti $w_1(\lambda), w_2(\lambda), \ldots w_n(\lambda)$ a A základní ryzí pojistné. Potom průměrné ryzí pojistné pojištěnce s rizikovým parametrem λ je

$$C(\lambda) = A\sum_{n=1}^{\infty} w_n(\lambda)\sum_{j\in S} p_j^{(n)}(\lambda)a_j,$$

kde $a = (a_1, \ldots a_J)$ je stupnice bonusu.

Ve vyváženém modelu by mělo platit

$$EC(\lambda) = \int_0^{\infty} C(\lambda)f(\lambda)d\lambda = ENEY,$$

kde EN je průměrný počet škod a EY je průměrná výše škody. Tato rovnice slouží pro určení základního ryzího pojistného A. Řekněme, že rozdělení doby, po kterou jsou pojištěnci pojištěni je geometrické nezávislé na λ

$$w_n(\lambda) = q^{n-1}(1-q), \quad n = 1, 2, \ldots.$$

Průměrná doba pojištění je pak $\frac{1}{1-q}$ a platí

$$C(\lambda) = A(1-q){p^{(1)}}^T(\lambda)\sum_{n=1}^{\infty} q^{n-1}P^{n-1}(\lambda)a = A(1-q){p^{(1)}}^T(I - qP(\lambda))^{-1}a,$$

kde I je jednotková matice, neboli

$$C(\lambda) = A(1-q)r_{k_0}(\lambda),$$

kde $r(\lambda)$ je řešením rovnice $(I - qP(\lambda)r(\lambda) = a$ a $r_{k_0}(\lambda)$ je řádek $r(\lambda)$ odpovídající počáteční bonusové třídě.

4.4 * Markovovy řetězce se stavy z množiny reálných čísel

Uvažujme následující lineární model

$$X_{n+1} = aX_n + \varepsilon_{n+1},$$

kde $\{\varepsilon_n\}_{n\in\mathbb{N}}$ je posloupnost nezávislých stejně rozdělených náhodných veličin s rozdělením $N(0,\sigma^2)$, s hustotou

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2\sigma^2}\right\}.$$

Přechodová hustota (ze stavu x do stavu y) - je hustota normálního rozdělení s parametry ax a σ^2

$$f(x,y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(y-ax)^2}{2\sigma^2}\right\}.$$

Přechodová hustota po k-krocích pak odpovídá normálnímu rozdělení s parametry $a^k x$ a $\sum_{l=0}^{k-1} a^{2l}\sigma^2 = \sigma^2\frac{1-a^{2k}}{1-a^2}$ ve tvaru

$$f^{(k)}(x,y) = \frac{1}{\sigma\sqrt{2\pi\frac{1-a^{2k}}{1-a^2}}} \exp\left\{-\frac{(y-a^k x)^2}{2\sigma^2\frac{1-a^{2k}}{1-a^2}}\right\}.$$

Je-li $|a|<1$, existuje limitní hustota f^∞ ve tvaru

$$f^\infty(y) = \lim_{k\to\infty}(x,y) = \frac{\sqrt{1-a^2}}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(1-a^2)y^2}{2\sigma^2}\right\},$$

přičemž hodnota této funkce nezávisí na x. Tato hustota je hustotou normálního rozdělení $N(0,\sigma^2/(1-a^2))$.

Toto rozdělení je stacionární v následujícím smyslu: Pokud hodnota řetězce X_0 v čase 0 má rozdělení $N(0,\sigma^2/(1-a^2))$, pak také hodnota řetězce v čase 1 $X_1 \sim N(0,\sigma^2/(1-a^2))$. V tom případě pro každé $n\in\mathbb{N}$ platí $X_n \sim N(0,\sigma^2/(1-a^2))$.

5 Časové řady

Časovou řadou rozumíme posloupnost (reálných) náhodných veličin $X_1,\cdots,X_n$, přičemž index $t=1,\cdots,n$ interpretujeme jako čas. Někdy ovšem také jen jednu realizaci $(x_1,\ldots,x_n)$ této náhodné posloupnosti.

Časová řada s nejjednodušší strukturou závislosti je posloupnost nezávislých stejně rozdělených náhodných veličin – ta se nazývá (striktní) bílý šum. Nás ovšem budou zajímat náhodné posloupnosti, kde nějaká závislost mezi veličinami odpovídajími různým časům (složkami, marginály) je.

Dále budeme uvažovat pouze posloupnosti, které jsou "invariantní v čase" – v silném nebo slabém smyslu.

Definice 15 *Posloupnost náhodných veličin $(X_1,\cdots,X_n)$ nazveme silně stacionární, pokud*

$$\mathcal{L}(X_{k_1},\cdots,X_{k_l}) = \mathcal{L}(X_{k_1+h},\cdots,X_{k_l+h})$$

platí pro každé $1\le k_1\le\cdots\le k_l+h\le n$.

Definice 16 *Posloupnost náhodných veličin* $(X_1, \cdots, X_n)$ *nazveme slabě stacionární, pokud*
(a) $EX_1 = \cdots = EX_n =: EX =: \mu$. *tj.* EX_i *nezávisí na* i
(b) $E(X_{t+k} - EX)(X_t - EX) =: r(k)$ *nezávisí na* $t = 1, \cdots, n-k$, *pro každé* $k = 0, 1 \cdots$

Funkce $r(k), k = 0, 1, \cdots$ *se nazývá autokovarianční funkce posloupnosti. Hodnota* $r(0)$ *pak označuje rozptyl posloupnosti (který je konstantní v čase). Vydělením autokovarianční funkce rozptylem obdržíme tzv. autokorelační funkci* $\rho(k) := r(k)/r(0)$.

Autokorelační funkce $\rho(k)$ nám tedy udává, jak moc jsou závislé dvě složky náhodné posloupnosti, které odpovídají časům posunutým o k jednotek.

Příklad: Pro striktní bílý šum platí $EX_t = 0$ pro všechna t a $r(k) = I(k = 0)var(X_1)$. Takže pokud je $var(X_1) < \infty$ je to slabě stacionární posloupnost. A určitě je to silně stacionární posloupnost. □

Definice 17 *Posloupnost náhodných veličin* $(X_1, \cdots, X_n)$ *nazveme bílý šum, pokud platí* $EX_t = 0$ *pro všechna* t *a* $r(k) = I(k = 0)var(X_1)$.

Tedy pro obyčejný bílý šum nemusejí být jeho složky nezávislé či stejně rozdělené, stačí nám, že jsou centrované a nekorelované a mají stejný rozptyl.

Poznámka: Pokud má silně stacionární posloupnost konečné druhé momenty, pak je slabě stacionární. Naopak, pokud je posloupnost náhodných veličin Gausovská (tj. všechna konečně rozměrná rozdělení jsou normální) a pokud je posloupnost slabě stacionární, je také silně stacionární, neboť vícerozměrné normální rozdělení je jednoznačně určeno svými prvními dvěma momenty (vektorem středních hodnot a rozptylovou maticí).

5.1 Klouzavé průměry

Jsou nejjednodušším zobecněním bílého šumu.

Definice 18 *Náhodnou posloupnost* $\{X_t, t \in T\}$, *kde* $T \subset \mathbf{Z}$ *je množina diskrétních časů nazveme posloupností klouzavých průměrů řádu* q, *pokud platí*

$$X_t = n_0 \varepsilon_t + \ldots n_q \varepsilon_{t-q}, \qquad \text{pro každé } t \in T,$$

pro nějaké koeficienty $n_0, \ldots, n_q$, $n_0, n_q \neq 0$ *a bílý šum* $\{\varepsilon_t\}$.

Takové posloupnosti už nejsou nekorelované, ale složky X_t, X_{t+k} jsou nekorelované pro dostatečně $(k > q)$ velký rozdíl v časech k.

Příklad: MA(1): $X_t = \varepsilon_t + \beta \varepsilon_{t-1}$, kde $\{\varepsilon_t\}$ je bílý šum.
Budeme se ptát, za jakých okolností (podmínek kladených na β) může být výsledná posloupnost stacionární a spočítáme autokovarianční funkci.

$$EX_t = E\varepsilon_t + \beta E\varepsilon_{t-1} = 0.$$

Z předpokladu nekorelovanosti ε_t a ε_{t-1} spočteme rozptyl.

$$r_X(0) = \operatorname{var} X_t = \operatorname{var} \varepsilon_t + \beta^2 \operatorname{var} \varepsilon_{t-1} = (1 + \beta^2)\, \sigma_\varepsilon^2.$$

Vynásobením MA(1) rovnosti hodnotou X_{t-1} a spočtením střední hodnoty obdržíme rovnost pro autokovarianci resp. autokorelaci se zpožděním 1.

$$r_X(1) = EX_tX_{t-1} = E(\varepsilon_t + \beta\varepsilon_{t-1})(\varepsilon_{t-1} + \beta\varepsilon_{t-2}) = \beta \mathrm{var}\, \varepsilon_{t-1} = \beta\sigma_\varepsilon^2.$$

Pro $k > 1$ dostáváme

$$r_X(k) = E(\varepsilon_t + \beta\varepsilon_{t-1})(\varepsilon_{t-k} + \beta\varepsilon_{t-k-1}) = 0,$$

neboť $t-1 > t-k$. Odtud dostáváme

$$\rho_X(1) = \frac{r_X(1)}{r_X(0)} = \frac{\beta\sigma_\varepsilon^2}{(1+\beta^2)\,\sigma_\varepsilon^2} = \frac{\beta}{1+\beta^2}.$$

Pro $k > 1$ dostaneme zřejmě $\rho_X(k) = 0$.

Závěr: Pro každou hodnotu parameru β může být MA(1) posloupnost stacionární. Pokud je stacionární, tak má nulovou střední hodnotu, rozptyl $(1+\beta^2)\sigma_\varepsilon^2$, autokorelaci se zpožděním 1 má $\beta/(1+\beta^2)$, přičemž autokorelace s větším zpožděním jsou nulové. □

5.2 Autoregresní modely řádu r (AR(r))

Jsou modely časových řad, ve kterých už jsou obecně všechny autokorelace nenulové a tak se v těchto náhodných posloupnostech ovlivňují složky, které jsou libovolně daleko od sebe (i když závislost se vzrůstajícím rozdílem v časech samozřejmě slábne).

Definice 19 *Posloupnost n.v. $X_n, n \in T$ nazveme autoregresní posloupností řádu p, pokud existují koeficienty $d_0, \cdots, d_p$ ($d_0 \neq 0 \neq d_p$) takové, že platí rovnost*

$$d_0X_t + \cdots + d_pX_{t-p} =: \varepsilon_t \tag{115}$$

kde $\{\epsilon_t\}$ je bílý šum.

Poznámka: Je-li autoregresní posloupnost X_n vyhovující modelu (115) stacionární, pak nutně všechny kořeny polynomu $d_0 + \cdots + d_px^p$ leží vně jednotkového kruhu. V tom případě

$$0 = E\varepsilon_t = d_0EX_t + \cdots + d_pEX_{t-p} = \left(\sum_{t=1}^{p} d_t\right)\mu \quad \Rightarrow \quad \mu = 0,$$

kde $\mu := EX_t$ nezávisí na t podle předpokladu stacionarity, neboť $\sum_{t=0}^{p} d_r = 0$ by znamenalo, že 1 je kořenem polynomu $d(x) = d_0 + \cdots + d_pz^p$.

Autoregresní model 1.řádu AR(1)

Model

$$X_t = aX_{t-1} + \varepsilon_t.$$

odpovídá volbě $d_0 = 1, d_1 = -a$. Polynom

$$d(x) = d_0x^0 + d_1x^1 = 1 - ax$$

má kořen v bodě $x = 1/a$. Bod $1/a$ je vně jednotkového kruhu právě tehdy, když $|a| < 1$. V tom případě je možné psát

$$X_t = aX_{t-1} + \varepsilon_t = a(aX_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \cdots = a^k X_{t-k} + \sum_{l=0}^{k-1} a^l \varepsilon_{t-l}, \tag{116}$$

a dokonce

$$X_t = \sum_{k=0}^{\infty} a^k \varepsilon_{t-k}. \tag{117}$$

Této vlastnosti (že je možné vyjádřit X_t jako lineární kombinaci posloupnosti $\{\epsilon_t, \epsilon_{t-1}, \dots\}$, tedy pomocí "minulosti" náhodné posloupnosti $\{\epsilon_t\}$) se říká kauzalita. Pro model AR(1) plyne z toho, že $|a| < 1$.

Autokovarianční funkci pak můžeme snadno dostat z rovnice (117) a vlastností bílého šumu:

$$r(k) = E\left(\sum_{i=0}^{\infty} a^i \varepsilon_{t-i}\right)\left(\sum_{j=0}^{\infty} a^j \varepsilon_{t-k-j}\right) = \sum_{j=0}^{\infty} a^{k+j} a^j \sigma_\varepsilon^2 = \frac{a^k}{1-a^2}\sigma_\varepsilon^2.$$

Pokud bychom ale neměli vyjádření posloupnosti $\{X_t\}$ jako kauzálního procesu a zajímal by nás třeba jenom rozptyl, mohli bychom postupovat i následovně: Rovnici (116) postupně vynásobíme hodnotami X_{t-1} a ε_t a následně spočteme střední hodnotu obou stran. Tímto způsobem dostaneme následující rovnosti

$$\begin{aligned} EX_t X_{t-1} &= aEX_{t-1}^2 + E\varepsilon_t X_{t-1} \\ EX_t \varepsilon_t &= aEX_{t-1}\varepsilon_t + E\varepsilon_t^2. \end{aligned}$$

Druhou z předchozích rovnic lze přepsat do následujícího tvaru

$$\sigma^2 = E\varepsilon_t^2 = EX_t\varepsilon_t = EX_t(X_t - aX_{t-1}) = r(0) - ar(1).$$

Naopak první rovnici lze přepsat takto

$$r(1) = ar(0) \quad \Rightarrow \quad a = \frac{r(1)}{r(0)} = \rho(1).$$

Odsud dostáváme rovnost pro rozptyl

$$r(0) = \frac{\sigma^2}{1 - ar(1)/r(0)} = \frac{\sigma^2}{1 - \rho(1)^2} \quad \text{resp.} \quad r(0) = \frac{\sigma^2}{1-a^2}.$$

Yule-Walkerovy rovnice

Předchozí postup na určení autokovarianční funkce autoregresní posloupnosti lze použít i pro obecnou autoregresní posloupnost.

Uvažujme autoregresní model upravený do následujícího tvaru

$$X_t = a_1 X_{t-1} + \cdots + a_p X_{t-p} + \varepsilon_t. \tag{118}$$

Tuto rovnost postupně násobíme $X_{t-1}, \cdots, X_{t-p}$ a počítáme střední hodnotu

$$\begin{aligned} EX_tX_{t-1} &= a_1EX_{t-1}^2 + \cdots + a_pEX_{t-p}X_{t-1} + E\varepsilon_tX_{t-1} \\ \cdots &= \cdots \qquad\qquad\qquad\qquad \cdots \qquad \cdots \\ EX_tX_{t-p} &= a_1EX_{t-1}X_{t-p} + \cdots + a_rEX_{t-p}^2 + E\varepsilon_tX_{t-p}. \end{aligned}$$

Tuto soustavu rovností lze přepsat do maticového zápisu

$$\begin{pmatrix} r(0) & \cdots & r(p-1) \\ \cdots & \cdots & \cdots \\ r(p-1) & \cdots & r(0) \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} r(1) \\ \vdots \\ r(p) \end{pmatrix}. \tag{119}$$

Vydělením celé rovnosti hodnotou $r(0)$ dojdeme k rovnosti

$$\begin{pmatrix} \rho(0) & \cdots & \rho(p-1) \\ \cdots & \cdots & \cdots \\ \rho(p-1) & \cdots & \rho(0) \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} \rho(1) \\ \vdots \\ \rho(p) \end{pmatrix}.$$

Z této rovnosti můžeme spočítat hodnoty autokorelační funkce $\rho(1), \ldots, \rho(p)$ a hodnoty $\rho(k)$ pro $k \geq s$ pak dostaneme z diferenční rovnice

$$\rho(k) - a_1\rho(k-1) - \ldots a_p\rho(k-p) = 0.$$

Rovnost pro rozptyl obdržíme, když vynásobíme rovnost (118) hodnotou ε_t.

$$\sigma^2 = E\varepsilon_t^2 = EX_t\varepsilon_t = EX_t(X_t - a_1X_{t-1} - \cdots - a_pX_{t-p}) = r(0)[1 - a_1\rho(1) - \cdots - a_p\rho(p)].$$

Platí tedy

$$\operatorname{var} X_t = r(0) = \frac{\sigma^2}{1 - a_1\rho(1) - \cdots - a_p\rho(p)}.$$

Yule-Walkerovy rovnice lze ovšem použít i obráceným způsobem – k odhadu parametrů a_i, modelu AR(r) pro daná data. Z dat určíme odhady autokovarianční funkce $\hat{r}(0), \ldots \hat{r}(p)$, ty dosadíme do rovnice (119) a vyřešením lineární soustavy pro $(a_1, \ldots, a_p)$ spočítáme odhady koeficientů $\hat{a}_i$. Že takto získané odhady jsou konzistentní ukazuje následující věta.

Věta 17 *Buď $X_0, \cdots, X_n$ autoregresní posloupnost řádu r, tj.*

$$X_t = a_1X_{t-1} + \cdots + a_pX_{t-p} + \varepsilon_t,$$

kde ε_t je posloupnost nezávislých náhodných veličin se střední hodnotou 0 a jednotkovým rozptylem, která je nezávislá s $(X_{t-1}, \cdots, X_0)$. Nechť je posloupnost $(X_0, \cdots, X_n)$ stacionární. Označme $\hat{a}$ odhad metodou Yule-Walkerových rovnic, pak

$$\sqrt{n}(\hat{a} - a) \to N(0, \sigma^2\Gamma_p^{-1}) \tag{120}$$

v distribuci pro $n \to \infty$, kde $\Gamma_{ij} = r(i-j)$.

Poznámka: Limitní vztah (120) se dá ekvivalentně přepsat do následující formy

$$\sup_{x \in \mathbb{R}^p} | P(\sqrt{n}(\hat{a}_1 - a_1) < x_1, \cdots, \sqrt{n}(\hat{a}_p - a_p) < x_p) - F_p(a_1, \cdots, a_p)| \to 0$$

pro $n \to \infty$, kde F_r je sdružená distribuční funkce rozděelení $N(0, \sigma^2 \Gamma_p^{-1})$.

A jak odhadujeme $\hat{r}(0), \ldots \hat{r}(p)$? Autokovarianci odhadujeme výběrovou autokovariancí

$$\hat{r}(k) = \frac{1}{n-k} \sum_{i=1}^{n-k} \left(X_{i+k} - \frac{1}{n-k} \sum_{i=1}^{n-k} X_{i+k} \right) \left(X_i - \frac{1}{n-k} \sum_{i=1}^{n-k} X_i \right),$$

kterýžto odhad lze použít i pro necentrovanou posloupnost. Střední hodnotu μ pak odhadujeme průměrem řady $\hat{\mu} := \frac{1}{n} \sum_{i=1}^{n} X_i$.

* ARMA modely

Samozřejmě je možné zkombinovat MA a AR modely do obecnějšího modelu nazývaného ARMA model.

Definice 20 *Posloupnost n.v. $X_n, n \in T$ vyhovuje ARMA(p,q) modelu (p,q jsou parametry z* **N***), pokud existují posloupnosti koeficientů $n_0, \cdots, n_q$ ($n_0, n_q \neq 0$) a $d_0, \cdots, d_p$, ($d_0, d_p \neq 0$) takových, že platí rovnost*

$$d_0 X_t + \cdots + d_p X_{t-p} = n_0 \varepsilon_t + \cdots + n_q \varepsilon_{t-q}, \tag{121}$$

kde $\{\epsilon_t\}$ je bílý šum.

A obdobně jako u autoregresních posloupností rovnice (121) definuje stacionární posloupnost s nulovou střední hodnotou, pokud kořeny polynomu $d_0 + \cdots + d_p x^p$ leží vně jednotkového kruhu.

Výpočet hodnot autokovarianční funkce v ARMA modelech či odhady koeficientů z výběrových autokovariancí probíhá u ARMA modelů obdobně jako u AR modelů.

Zde je na místě připomenout souvislost ARMA modelů s lineárnímy modely z kapitoly 3. Rovnice (121) je totiž opravdu formálně ekvivalentní s obecnou rovnicí lineárního systému, nyní ale určuje lineární vztah nikoli mezi nenáhodnými posloupnostmi vstupu a výstupu, ale mezi náhodnými posloupnostmi bílého šumu a ARMA procesu $\{X_t\}$. A podmínka na stabilitu lineárního systému odpovídá podmínce nutné pro to, aby bylo možno náhodný proces $\{X_t\}$ vyjádřit jako

$$X_t = \sum_{k=0}^{\infty} c_k Y_{t-k}, \tag{122}$$

kde $\{c_k\}_{k=0}^{\infty}$ je posloupnost koeficientů splňující $\sum_{k=0}^{\infty} |c_k| < \infty$ (a z této zase plyne, že potom je řada náhodných veličin $\sum_{k=0}^{\infty} c_k Y_{t-k}$ sčítatelná (tj. posloupnost částečných součtů těch náhodných veličin konverguje k nějaké náhodné veličině) v L_2 a tedy i v pravděpodobnosti).

Rovnice (122) se používá ke konstrukci nových náhodných procesů i mimo ARMA modelů, takže pro úplnost uveďme ještě dvě definice pojmů, se kterými se můžete v budoucnu setkat.

Definice 21 *Buď $Y_t, t \geq 0$ slabě stacionární posloupnost a nechť existuje posloupnost koeficientů $(\psi_k, k \in \mathbb{Z})$ taková, že $\sum_{k\in\mathbb{Z}} |\psi_k| < \infty$, pak řekneme, že posloupnost*

$$X_t := \sum_{k=-\infty}^{\infty} \psi_k Y_{t-k} \tag{123}$$

vznikla lineární filtrací z posloupnosti $\{Y_k\}$ a posloupnosti $\{\psi_k\}$ říkáme lineární filtr. Je-li $\psi_k = 0$ pro $k < 0$, pak říkáme, že jde o filtr fyzikálně uskutečnitelný resp. kauzální.

Definice 22 *O posloupnosti n.v. $\{X_n\}$ řekneme, že je lineární proces, pokud existuje posloupnost nezávislých náhodných veličin $\varepsilon_t, t \geq 0$ se střední hodnotou nula a jednotkovým rozptylem a lineární filtr $\{\psi_k\}$ takový, že $X_t = \sum_{k=-\infty}^{\infty} \psi_k \varepsilon_{t-k}$.*

O posloupnosti X_k řekneme, že je kauzální, pokud je navíc filtr $\{\psi_k\}$ kauzální.

5.3 Lineární predikce

V souvislosti s časovými řadami se ještě krátce zmíníme o problému predikce(předpovědi) dalších (zatím nepozorovaných) hodnot v časové řadě na základě zatím pozorovaných. Na příkladě si ukážeme, jak se odvodí nejlepší lineární predikce o jeden krok dopředu (postup lze samozřejmě zobecnit na předpověď o více kroků dopředu nebo na předpověď na základě více než jednoho pozorování X_t).

Tedy úloha zní: předpovědět hodnotu náhodné veličiny X_{t+1} na základě náhodné veličiny X_t, přičemž se omezíme jen na lineární predikce tvaru $\hat{X}_{t+1} = a + bX_t$. O posloupnosti $X_1, \cdots, X_n$ budeme předpokládat, že je slabě stacionární a jako kritérium optimality budeme brát střední kvadratickou chybu mezi X_t a předpovědí $\hat{X}_{t+1}$. Řešíme tedy minimalizační problém najít

$$\min_{a,b} E(X_{t+1} - a - bX_t)^2 = \min_{a',b} E(X_{t+1} - \mu - a' - b(X_t - \mu))^2,$$

kde $a' = a + \mu(b - 1)$.

Derivováním podle parametrů a', b dojdeme k nutným podmínkám pro lokální extrém

$$\begin{aligned}
\frac{\partial}{\partial a'} : \qquad & -2E(X_{t+1} - \mu - a' - b(X_t - \mu)) = 0 \\
\frac{\partial}{\partial b} : \qquad & -2E(X_{t+1} - \mu - a' - b(X_t - \mu))(X_t - \mu) = 0.
\end{aligned}$$

Z druhé rovnice obdržíme podmínku na b ve tvaru

$$E(X_{t+1} - \mu)(X_t - \mu) = bE(X_t - \mu)^2,$$

což vede k rovnosti

$$b = \frac{r(1)}{r(0)} = \rho(1).$$

Naopak z první rovnice dostaneme

$$a' = E(X_{t+1} - \mu - b(X_t - \mu)) = 0.$$

Výsledná predikce je pak tvaru

$$\hat{X}_{t+1} = a + bX_t = \mu + b(X_t - \mu) = \mu + \frac{r(1)}{r(0)}(X_t - \mu) = \mu + \rho(1)(X_t - \mu), \tag{124}$$

kde za hodnoty $\mu, r(0), r(1)$ se dosadí příslušné odhady $\hat{\mu}, \hat{r}(0), \hat{r}(1)$. Kvalita tohoto odhadu se posuzuje pomocí střední kvadratické odchylky

$$\begin{aligned} MSE &= E(X_{t+1} - \hat{X}_{t+1})^2 = E(X_{t+1} - \mu - \rho(1)(X_t - \mu))^2 \\ &= r(0)[1 + \rho(1)^2 - 2\rho(1)^2] = r(0)[1 - \rho(1)^2]. \end{aligned}$$

Obdobně řešíme modifikovanou úlohu najít lineární předpověď n.v. X_{t+1} pomocí hodnot X_t, X_{t-1}, tedy najít

$$\hat{X}_{t+1} = a + bX_t + cX_{t-1} = \mu + a' + b(X_t - \mu) + c(X_{t-1} - \mu),$$

tak, aby

$$\min_{a,b,c} E(X_{t+1} - \hat{X}_{t+1})^2 = \min_{a',b,c} E((X_{t+1} - \mu) - a' - b(X_t - \mu) - c(X_{t-1} - \mu))^2,$$

kde a' je nyní rovno $a + \mu(b + c - 1)$.

Derivováním podle parametru a' dojdeme k podmínce $a' = 0$. Derivováním podle zbylých dvou parametrů dojdeme k rovnostem

$$\begin{aligned} \frac{\partial}{\partial b}: &\quad E[(X_{t+1} - \mu) - b(X_t - \mu) - c(X_{t-1} - \mu)](X_t - \mu) = 0 \\ \frac{\partial}{\partial c}: &\quad E[(X_{t+1} - \mu) - b(X_t - \mu) - c(X_{t-1} - \mu)](X_{t-1} - \mu) = 0. \end{aligned}$$

Tuto soustavu lze přepsat následujícím způ sobem

$$\begin{pmatrix} r(0) & r(1) \\ r(1) & r(0) \end{pmatrix} \begin{pmatrix} b \\ c \end{pmatrix} = \begin{pmatrix} r(1) \\ r(2) \end{pmatrix}.$$

Tato rovnice má řešení ve tvaru

$$\begin{pmatrix} b \\ c \end{pmatrix} = \frac{1}{1 - \rho(1)^2} \begin{pmatrix} 1 & -\rho(1) \\ -\rho(1) & 1 \end{pmatrix} \begin{pmatrix} \rho(1) \\ \rho(2) \end{pmatrix} \tag{125}$$

Střední kvadratickou odchylku spočítáme dosazením odhadů b, c do

$$MSE = E\left((X_{t+1} - \mu) - \begin{pmatrix} b \\ c \end{pmatrix}^T \begin{pmatrix} X_t - \mu \\ X_{t-1} - \mu \end{pmatrix} \right)^2$$

což je rovno

$$MSE = r(0) \begin{pmatrix} -1 \\ b \\ c \end{pmatrix}^T \begin{pmatrix} \rho(0) & \rho(1) & \rho(2) \\ \rho(1) & \rho(0) & \rho(1) \\ \rho(2) & \rho(1) & \rho(0) \end{pmatrix} \begin{pmatrix} -1 \\ b \\ c \end{pmatrix}, \tag{126}$$

kde za hodnoty b, c dosadíme výsledek (125).

6 Poissonův proces

Nechť $T \subseteq \mathbb{R}$ je neprázdná množina, pak sytém náhodných veličin $(X(t), t \in T)$ definovaných na témže pravděpodobnostním prostoru nazveme náhodným procesem.

Funkci $t \mapsto X(t)(\omega)$ pro daný elementární jev $\omega \in \Omega$ nazveme trajektorií procesu X.

Definice 23 *Řekneme, že náhodný proces $(N(t), t \geq 0)$ se spojitým časem je čítací proces, pokud existuje neklesající posloupnost nezávislých náhodných veličin $\xi_i, i \in \mathbf{N}$ (časů výskytu událostí) takových, že*

$$N(t) = \sum_{i=1}^{\infty} I\{\xi_i \leq t\}.$$

Tedy $N(t)$ interpretujeme jako počet událostí, které nastaly během intervalu $[0, t]$.
Ekvivalentně definujeme čítací proces jako proces $(N(t), t \geq 0)$,který má zprava spojité neklesající trajektorie, nabývá pouze celočíselných hodnot a startuje z nezáporné hodnoty, tj. $N(0) \geq 0$.

Definice 24 *Řekneme, že proces $(N(t), t \geq 0)$ má nezávislé přírůstky, pokud pro každé*

$0 \leq t_0 \leq \cdots \leq t_n, t_i \in T, i \leq n \in \mathbb{N}$ *jsou přírůstky*

$$N(t_n) - N(t_{n-1}), \cdots, N(t_1) - N(t_0)$$

procesu $(N(t), t \in T)$ nezávislé náhodné veličiny.
Poissonovým procesem $(N(t), t \geq 0)$ s intenzitou $\lambda > 0$ nazveme proces splňující:
1. $N(0) = 0$.
2. Proces má nezávislé přírůstky.
3. Přírůstky $N(t + h) - N(t)$ mají Poisonovo rozdělení s parametrem λh,.
(Tedy parametr Poissonova rozdělení je úměrný délce časového intervalu, na kterém přírůstek uvažujeme, a konstanta úměrnosti je rovna intenzitě λ.)

Poznámka: Protože Poissonův proces $(N(t), t \geq 0)$ startuje z nuly (tj. $N_0 = 0$), lze jeho hodnotu v čase t vyjádřit následujícím způsobem

$$N(t) = N(t) - N(0) + N(0) = N(t) - N(0),$$

což je podle předpokladu náhodná veličina s Poissonovým rozdělením s parametrem λt. Platí tedy

$$P(N(t) = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

Věta 18 *Nechť $(N(t), t \geq 0)$ je Poissonův proces s intenzitou λ. Definujme posloupnost náhodných veličin $\{\xi_k, k \in \mathbb{N}\}$ předpisem*

$$\xi_k := \inf\{t \geq 0, N(t) = k\},$$

pak doby mezi událostmi

$$T_k = \xi_k - \xi_{k-1}, \qquad kde \quad \xi_0 := 0,$$

jsou nezávislé náhodné veličiny s exponenciálním rozdělením s parametrem λ, tj. s hustotou

$$f(x) = \lambda e^{-\lambda x}, \qquad x > 0.$$

Důsledek: Náhodná veličina ξ_k (doba do k-té poruchy, pokud události interpretujeme jako poruchy) má Erlangovo (Gamma) rozdělení s parametry λ a k, tj. s hustotou

$$g_k(x) = \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x}, \qquad x > 0.$$

Poznámka: Exponenciální rozdělení je tzv. rozdělení bez paměti, tj. pro náhodnou veličinu Y s exponenciálním rozdělením platí

$$P(Y > s + h | Y > s) = P(Y > h)$$

Věta 19 *Nechť $T_1, T_2, \cdots$ je posloupnost nezávislých náhodných veličin s exponenciálním rozdělením s parametrem $\lambda > 0$, tj. s hustotou*

$$f(x) = \lambda e^{-\lambda x} I\{x > 0\}.$$

Definujme $\xi_k := \sum_{i=1}^k T_i$, pak

$$N(t) = \sum_{k=1}^{\infty} I\{\xi_k \le t\}. \tag{127}$$

definuje Poissonův proces s intenzitou λ.

Poznámka: Proces $(N(t), t \ge 0)$ definovaný v (127) je zřejmě čítací proces. Dále

$$P(N(t) = k) = P(\xi_k \le t < \xi_{k+1}) = P(\xi_k \le t) - P(\xi_{k+1} \le t) = \int_0^t g_k(x)\,dx - \int_0^t g_{k+1}(x)\,dx$$
$$= \int_0^t \frac{\lambda^k}{k!} k x^{k-1} e^{-\lambda x}\,dx - \int_0^t \frac{\lambda^{k+1}}{k!} x^k e^{-\lambda x}\,dx = \frac{(\lambda t)^k}{k!} e^{-\lambda t},$$

kde v poslední rovnosti jsme použili integraci per partes. Ověření nezávislosti přírůstků je trochu složitější a nebudeme ho tu provádět.

Poznámka: Pro Poissonův proces platí také následující rovnosti:

$$\begin{aligned} P(N(t+h) = k+1 \,|N(t) = k) &= \lambda h + o(h), \\ P(N(t+h) = k \,|N(t) = k) &= 1 - \lambda h + o(h), \\ P(N(t+h) > k+1 \,|N(t) = k) &= o(h), \end{aligned} \tag{128}$$

kde pro funkci $o(h)$ platí $\lim_{h \to 0} \frac{o(h)}{h} = 0$. Neboli pravděpodobnost výskytu události v krátkém intervalu je úměrná intenzitě procesu λ a délce intervalu. Pravděpodobnost výskytu dvou a více událostí klesá k nule s klesající délkou intervalu a to rychleji, než délka intervalu. Vztahy (128) dokonce charakterizují Poissonův proces.

Příklad: Škodné události v neživotním pojištění:
Předpokládáme, že okamžiky pojistných událostí tvoří Poissonův proces $(N(t), t \ge 0)$ s konstantní intenzitou λ a výše škod $Y_k, k \in \mathbf{N}$ jsou navzájem nezávislé stejně rozdělené náhodné veličiny. Buď

$$S(t) = \sum_{k=0}^{N(t)} Y_k,$$

úhrn všech škod do doby t. Spočítejme střední hodnotu $S(t)$:

$$
\begin{aligned}
ES(t) = E(E[S(t)|N(t)]) &= \sum_{n=0}^{\infty} P(N(t)=n)\, E[S(t)|N(t)=n] \\
&= \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t}\, E\left(\sum_{k=1}^{n} Y_k\right) = \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t} n\, E(Y_1) \\
&= \lambda t\, E(Y_1) = E(N(t))\, EY_1.
\end{aligned}
$$

Tedy k úhradě škodních nákladů by měla pojišťovna dostávat od pojištěných netto pojistné λEY_1 za jednotku času. □

Příklad: Obslužná linka.
Představme si zařízení poskytující určitou službu – např. pokladnu v obchodě, telefonní linku pomoci zákazníkům apod. Chceme modelovat její chování. Předpokládáme, že příchody zákazníků tvoří Poissonův proces s intenzitou λ. Je-li linka obsazena, tvoří se fronta. Buď $X(t)$ počet zákazníků, kteří jsou v systému (= obsluhovaný zákazník a zákazníci ve frontě) v čase t. Označme

$$p_k(t) = P(X(t) = k), \qquad k = 0, 1, 2, \dots,$$

a předpokládejme, že doba trvání obsluhy jednoho zákazníka je náhodná veličina T s exponenciálním rozdělením s parametrem $\mu > 0$ a doby obsluhy pro různé zákazníky jsou nezávislé a nezávislé na Poissonově procesu příchodů.

Pro chování systému platí:

$$
\begin{aligned}
&P(\text{během } (t, t+h] \text{ přijde právě jeden zákazník}) &&= \lambda h + o(h), \\
&P(\text{během } (t, t+h] \text{ nepřijde žádný zákazník}) &&= 1 - \lambda h + o(h), \\
&P(\text{během } (t, t+h] \text{ přijde více než jeden zákazník}) &&= o(h).
\end{aligned}
$$

a

$$
\begin{aligned}
P(\text{obsluha skončí v } (t, t+h] \mid \text{zákazník je obsluhován}) &= P(T \in (t, t+h] | T > t) \\
&= \frac{1 - e^{-\mu(t+h)} - (1 - e^{-\mu t})}{e^{-\mu t}} = \mu h + o(h),
\end{aligned}
$$

kde jsme použili Taylorův rozvoj exponenciely.

$$
\begin{aligned}
P(\text{obsluha neskončí v } (t, t+h] \mid \text{zákazník je obsluhován}) &= P(T > t+h | T > t) \\
&= \frac{e^{-\mu(t+h)}}{e^{-\mu t}} = e^{-\mu h} = 1 - \mu h + o(h), \\
P(\text{během } (t, t+h] \text{ bude obsloužen víc než jeden zákazník}) &= o(h).
\end{aligned}
$$

Tedy

$$P(\text{v systému nastane nějaká změna}) = \mu h + \lambda h + o(h),$$

neboť

$$P(\text{během } (t, t+h] \text{ přijde a odejde stejné nenulové množství zákazníků}) = o(h).$$

Chceme zjistit, jak vypadá limitní chování $p_k(t)$ pro $t \to \infty$. Tedy jak vypadá limitní rozdělení počtu zákazníků v systému. Pro $p_k(t)$ můžeme odvodit soustavu diferenciálních rovnic, neboť:

$$p_0(t+h) = p_0(t)(1 - \lambda h + o(h)) + p_1(t)(\mu h + o(h)) + \sum_{j=2}^{\infty} p_j(t) o(h),$$

tedy

$$\frac{p_0(t+h) - p_0(t)}{h} = -p_0(t)\lambda + p_1(t)\mu + \frac{o(h)}{h},$$

a

$$\frac{d}{dt} p_0(t) = -\lambda p_0(t) + \mu p_1(t).$$

Obdobně pro $k = 1, 2, \ldots$

$$\begin{aligned} p_k(t+h) = &\sum_{j=0}^{k-2} p_j(t) o(h) + p_{k-1}(t)(\lambda h + o(h)) + p_k(t)(1 - \lambda h - \mu h + o(h)) \\ &+ p_{k+1}(t)(\mu h + o(h)) + \sum_{j=k+2}^{\infty} p_j(t) o(h), \end{aligned}$$

takže

$$\frac{d}{dt} p_k(t) = \lambda p_{k-1}(t) - (\lambda + \mu) p_k(t) + \mu p_{k+1}(t), \qquad k = 1, 2, \ldots.$$

Pokud existuje $p_k(\infty) = \lim_{t \to \infty} p_k(t)$, pak pro ni platí

$$\begin{aligned} 0 &= -\lambda p_o(\infty) + \mu p_1(\infty), \\ 0 &= \lambda p_{k-1}(\infty) - (\lambda + \mu) p_k(\infty) + \ \mu p_{k+1}(\infty), \qquad k = 1, 2, \ldots. \end{aligned}$$

Označíme-li $\rho = \frac{\lambda}{\mu}$, pak řešení soustavy diferenčních rovnic je

$$p_k(\infty) = \rho^k p_0(\infty),$$

a z normovací podmínky

$$\sum_{k=0}^{\infty} p_k(\infty) = 1,$$

máme, že limitní rozdělení existuje pokud $\rho < 1$ tedy $\lambda < \mu$ a pak

$$p_k(\infty) = (1 - \rho)\rho^k.$$

Tedy limitní rozdělení je geometrické s parametrem $\rho = \frac{\lambda}{\mu}$. $\square$