Cvičení z optimalizace
Metoda nejmenších čtverců: Identifikace autoregresního modelu

Vojtěch Franc, 2014

1 Identifikace parametrů autoregresního modelu

Mějme zadanou posloupnost reálných čísel $y_0, y_1, \ldots, y_T$. Posloupnost může například popisovat signál digitalizovaný v diskrétních časových okamžicích $t = 0, 1, \ldots, T$. Naším cílem bude odhadnout parametry autoregresního modelu této posloupnosti, který předpokládá, že

$$y_t = a_0 + \sum_{i=1}^{p} a_i y_{t-i} + \varepsilon_t \,, \qquad t \geq p \,, \tag{1}$$

tj. t-tý prvek posloupnosti je lineární kombinací p předcházejících prvků a náhodné odchylky (šumu) ε_t. Číslo $p \in \mathcal{N}^+$ se nazývá řádem modelu. Autoregresní model p-tého řádu je určen vektorem parametrů $\boldsymbol{a} = (a_0, a_1, \ldots, a_p) \in \mathbb{R}^{p+1}$. Odhadování těchto parametrů z dat, tj. ze zadané posloupnosti $y_0, y_1, \ldots, y_T$, se v teorii řízení nazývá *identifikace systému*. Jedna z metod identifikace je založena na hledání takových parametrů, které zajistí, že model co nejlépe odpovídá zadaným datům ve smyslu minimalizace součtu kvadrátu odchylek ε_t pro $t = p, \ldots, T$. To znamená, že hledáme vektor parametrů

$$\hat{\boldsymbol{a}} = \operatorname*{argmin}_{\boldsymbol{a} \in \mathbb{R}^{p+1}} F(\boldsymbol{a}) \,, \tag{2}$$

kde

$$F(\boldsymbol{a}) = \sum_{t=p}^{T} \varepsilon_t^2 = \sum_{t=p}^{T} \left(\left(a_0 + \sum_{i=1}^{p} a_i y_{t-i} \right) - y_t \right)^2 .$$

Řešení problému (2) lze převést na řešení přeurčené soustavy lineárních rovnic $\mathbf{M}\boldsymbol{a} \approx \boldsymbol{b}$ metodou nejmenších čtverců.

Takto získaný model se dá použít například pro kompresi posloupnosti $y_0, y_1, \ldots, y_T$. V tomto případě postačí uchovat podsekvenci $y_0, \ldots, y_{p-1}$ a odhadnuté parametry $\hat{\boldsymbol{a}}$. Zbývající členy poslopnosti $y_p, \ldots, y_T$ vygenerujeme rekurzivním použitím (1). Koeficient komprese je tedy $T/(2p{+}1)$. Na podobné myšlence je postavena komprese řečového signálu v reálném čase, která se používá například v mobilních telefonech.

Úkoly k vypracování

1. Stáhněte si WAV soubor se zvukem gongu `http://cmp.felk.cvut.cz/cmp/courses/OPT/cviceni/07/gong.wav` a nahrajte si ho do Matlabu pomocí příkazu

```
[y,Fs]=wavread('gong.wav');
```

`y` je vektor obsahující hodnoty zvukové sekvence vzorkované s frekvencí `Fs`. Pokud máte počítač se zvukovou kartou, můžete si signál přehrát pomocí příkazu `sound(y,Fs)`.

2. Formulujte úlohu (2) jako problém řešení soustavy lineárních rovnic $\mathbf{M}\boldsymbol{a} \approx \boldsymbol{b}$ metodou nejmenších čtverců, tj. jako úlohu $\min_{\boldsymbol{a}} \|\mathbf{M}\boldsymbol{a}-\boldsymbol{b}\|^2$. Napište explicitní tvar matice $\mathbf{M}$ a vektoru $\boldsymbol{b}$. Pro nahranou zvukovou sekvenci odhadněte parametry autoregresního modelu s řádem $p = 300$ s využitím operátoru Matlabu `M\b`.

   Výstup: i. předpis pro $\mathbf{M}$ a $\boldsymbol{b}$, ii. hodnota $\min_{\boldsymbol{a}} \|\mathbf{M}\boldsymbol{a} - \boldsymbol{b}\|^2$.

3. Implementujte algoritmus řešící problém nejmenších čtverců pomocí QR rozkladu. Vstupem algoritmu nechť je dvojice $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, $m > n$ a výstupem vektor $\hat{\boldsymbol{x}}$, který je řešením $\min_{\boldsymbol{x}\in\mathbb{R}^n} \|\mathbf{A}\boldsymbol{x} - \boldsymbol{b}\|^2$. Předpokládejte, že matice $\mathbf{A}$ má plnou hodnost. Uvnitř algoritmu použijte funkci Matlabu `[Q,R] = qr(A,0)` pro získání matice $\mathbf{Q}$ a $\mathbf{R}$. Váš algoritmus může použít kromě funkce `qr` jen základní operace, tj. nesmí například používat operaci "zpětného lomítka" nebo inverzi. Použijte Vaší implementaci k řešení problému z úkolu 2 a porovnejte Váš výsledek s výsledkem získaným pomocí `A\b`.

   Výstup: i. funkce v Matlabu `x=solve_ls(A,b)`, ii. vzdálenost vektorů parametrů $\|\hat{\boldsymbol{a}}_1 - \hat{\boldsymbol{a}}_2\|$, kde $\hat{\boldsymbol{a}}_1$ je řešení získané Vaší funkcí a $\hat{\boldsymbol{a}}_2$ řešení pomocí `A\b`.

4. Použijte odhadnutý model k vygenerování syntetického zvuku gongu. Porovnejte synteticky vygenerovaný a originální zvuk gongu graficky, tj. zobrazte oba průběhy do jednoho grafu pomocí funkce `plot`. Pokud to Váš počítač umožňuje, přehrajte si oba zvukové signály pomocí příkazu `sound(y,Fs)`.

   Výstup: i. graf.