Domácí úkol OPT 1

by a very good student

22. října 2019

Úkol Metoda nejmenších čtverců

Predikce průměrné hrubé mzdy

Řešíme minimalizační úlohu

$$\vec{x}^* = \operatorname{argmin} \left\| A\vec{x} - \vec{b} \right\|^2$$

Jelikož prokládáme statistická data přímkou, definujeme matici A jako

$$A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \dots & \dots \\ 1 & t_n \end{bmatrix}$$

kde $\vec{t}$ je sloupcový vektor skládající se z číselných označení jednotlivých roků, a $\vec{b}$ je vektor obsahující odpovídající statistické teplotní hodnoty.
Poté vypočteme pseudoinverzi matice A jako

$$A^\dagger = (A^T \cdot A)^{-1} \cdot A^T$$

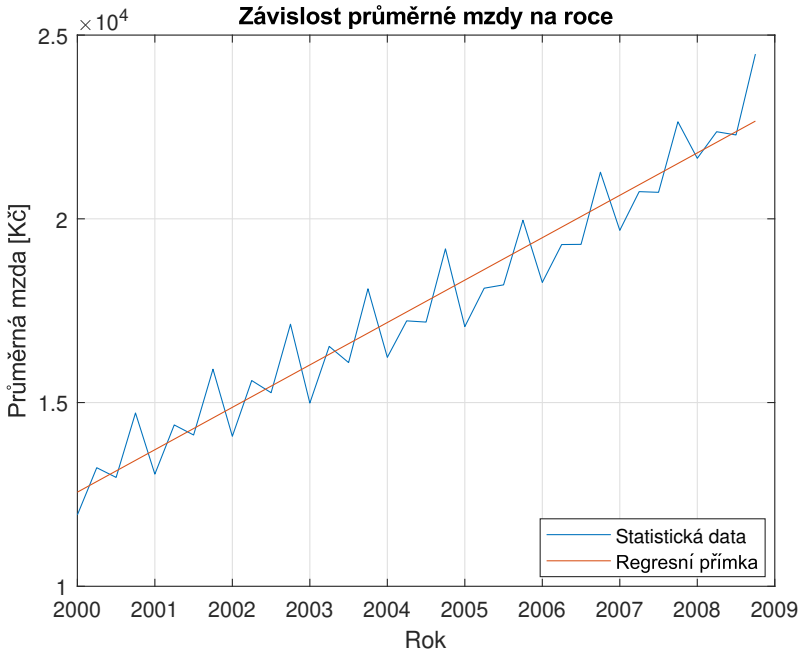
Hodnoty vektoru $\vec{x}^*$ dostaneme výpočtem

$$\vec{x}^* = A^\dagger \cdot \vec{b}$$

Výsledná regresní přímka má tvar $y(t) = x_0^* + x_1^* t$.
Pro zadaná data vychází

$$x_0^* = -2.2955 \cdot 10^6$$
$$x_1^* = 1.1540 \cdot 10^3$$

Odhad průměrné hrubé mzdy pro druhý kvartál roku 2009 je 23237 Kč.



Interpolace denní teploty ve Svatoňovicích

Jako regresní křivku použijeme funkci

$$\hat{T}(t) = x_0 + x_1 t + x_2 \sin(\omega t) + x_3 \cos(\omega t)$$

kde ω je kruhová frekvence odpovídající periodě sinusové funkce, $T = 365$ dní, tj. $\omega = \frac{2\pi}{365} \approx 0.0172$.
Matici A sestavíme jako

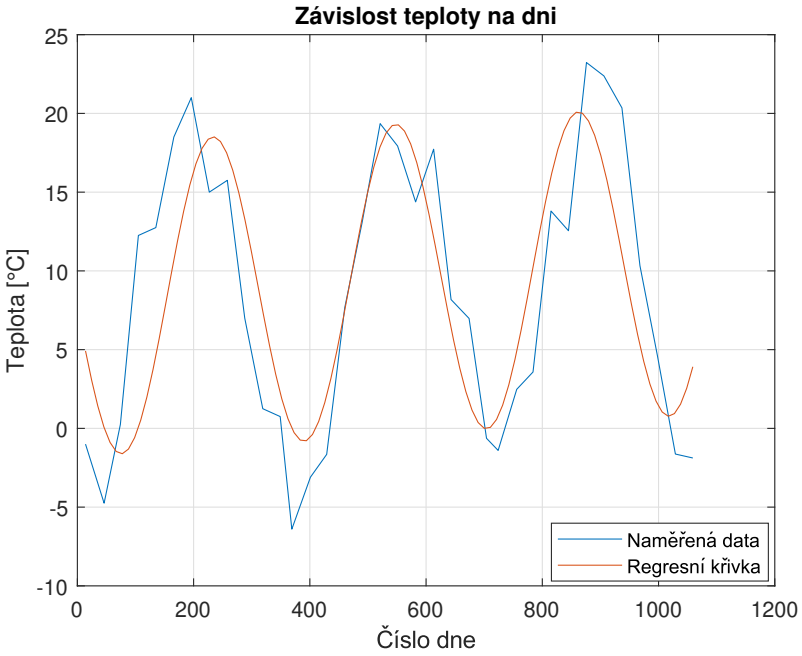
$$A = \begin{bmatrix} 1 & t_1 & \sin(\omega t_1) & \cos(\omega t_1) \\ 1 & t_2 & \sin(\omega t_2) & \cos(\omega t_2) \\ \dots & \dots & \dots & \dots \\ 1 & t_n & \sin(\omega t_n) & \cos(\omega t_n) \end{bmatrix}$$

Opět použijeme pseudoinverzi matice A a vypočteme vektor $\vec{x}$:

$$(x_0^*, x_1^*, x_2^*, x_3^*) = A^\dagger \cdot \vec{b}$$

Pro zadaná data vychází hodnoty

$$\begin{aligned} x_0^* &= 7.6238 \\ x_1^* &= 0.0014 \\ x_2^* &= -0.0174 \\ x_3^* &= -11.3054 \end{aligned}$$



Důkaz tvrzení

Máme dokázat, že namísto řešení nelineární optimalizační úlohy pro funkci

$$\hat{G}(t) = y_0 + y_1 t + A\sin(\omega t + \phi)$$

použijeme funkci

$$\hat{T}(t) = x_0 + x_1 t + x_2\sin(\omega t) + x_3\cos(\omega t)$$

protože pro každou čtveřici (y_0, y_1, A, ϕ) existuje čtveřice (x_0, x_1, x_2, x_3) tak, že $\hat{G}(t) = \hat{T}(t)$ pro každé $t \in \mathbb{R}$. Důkaz spočívá v jednoduché aplikace součtového goniometrického vzorce

$$\sin(x + y) = \sin(x)\cos(y) + \cos(x)\sin(y)$$

Potom dostaneme

$$A\sin(\omega t + \phi) = A\cos(\phi)\sin(\omega t) + A\sin(\phi)\cos(\omega t)$$

Tedy můžeme definovat $x_0 = y_0$, $x_1 = y_1$, $x_2 = A\cos(\phi)$ a $x_3 = A\sin(\phi)$, které existují pro každou reálnou hodnotu A a ϕ. Výrazy $\hat{G}$ a $\hat{T}$ poté nabývají ekvivalentních hodnot pro každé $t \in \mathbb{R}$.