Univerzita Hradec Králové

Přírodovědecká fakulta

Katedra matematiky

Metoda nejmenších čtverců

Bakalářská práce

Autor: Michaela Zavřelová
Studijní program: B 1103 Aplikovaná matematika
Studijní obor: Finanční a pojistná matematika
Vedoucí práce: Mgr. Jitka Kühnová, Ph.D.

Hradec Králové květen 2016

Univerzita Hradec Králové

Přírodovědecká fakulta

Zadání bakalářské práce

Autor:	**Michaela Zavřelová**
Studijní program:	B 1103 Aplikovaná matematika
Název práce:	Metoda nejmenších čtverců
Název práce v AJ:	The least squares method
Cíl a metody práce:	Tématem této bakalářské práce je metoda nejmenších čtverců a její využití. V práci budou popsány základní matematické vlastnosti, které jsou využity v definicích a výpočtech. Dále se práce zaměří na samotnou metodu nejmenších čtverců a její použití na konkrétních příkladech.
Garantující pracoviště:	Katedra matematiky Přírodovědecké fakulty UHK
Vedoucí práce:	Mgr. Jitka Kühnová, Ph.D.
Konzultant:	
Oponent:	RNDr. Michal Čihák, Ph.D.
Datum zadání práce:	11. 4. 2015
Datum odevzdání práce:	16. 5. 2016

Prohlášení:

Prohlašuji, že jsem svou bakalářskou práci vypracovala samostatně pod vedením Mgr. Jitky Kühnové, Ph.D. a že jsem v seznamu použité literatury uvedla všechny prameny, ze kterých jsem vycházela.

V Hradci Králové dne 16. 5. 2016 Michaela Zavřelová

Poděkování

Tímto bych chtěla poděkovat své vedoucí bakalářské práce Mgr. Jitce Kühnové, Ph.D. za trpělivost, kterou ke mně projevila, za ochotu kdykoli se sejít a za cenné rady, bez kterých by tato práce nemohla vzniknout.

Anotace

ZAVŘELOVÁ, Michaela. *Metoda nejmenších čtverců*. Hradec Králové, 2016. Bakalářská práce na Přírodovědecké fakultě Univerzity Hradec Králové. Vedoucí bakalářské práce Mgr. Jitka Kühnová, Ph.D., 43 s.

Tématem této bakalářské práce je metoda nejmenších čtverců a její využití. Metoda nejmenších čtverců je jednou ze základních metod aproximace funkcí ve statistice. Tato metoda je základní technikou používanou k vyrovnání naměřených dat, odhadu parametrů aproximačních funkcí například v regresní analýze. V práci budou popsány základní matematické vlastnosti, které jsou využity v definicích a výpočtech. Cílem této práce je zaměření na samotnou metodu nejmenších čtverců a její použití na konkrétních příkladech.

Klíčová slova

Ortogonální polynomy, regresní modely, metoda nejmenších čtverců, rezidua, soustava normálních rovnic

Annotation

ZAVŘELOVÁ, Michaela. *The least squares method*. Hradec Králové, 2016. Bachelor Thesis at Faculty of Science University of Hradec Králové. Thesis Supervisor Mgr. Jitka Kühnová, Ph.D., 43 p.

The theme of this thesis is the least squares method and its application. The least squares method is one of the basic approximation methods in statistics. This method is a basic technique used to offset the measured data, parameter estimation approximation of functions, for example, in regression analysis. This thesis will describe the basic mathematical properties that are used in the definitions and calculations. The aim of this thesis is focused on the least squares method and its application to concrete examples.

Keywords

Orthogonal polynomials, regression models, the least squares method, residues, system of normal equations

Obsah

Úvod **7**

1 Základní pojmy **8**
1.1 Vektorový prostor 8
1.2 Ortogonální polynomy 10
1.2.1 Čebyševovy polynomy 11
1.2.2 Gramovy polynomy 11

2 Jednoduchá regrese **12**
2.1 Obecný lineární regresní model 12
2.1.1 Cíle regresní analýzy 12
2.1.2 Klasifikace regresních modelů 12
2.1.3 Sestavování odhadů 15
2.1.4 Formulace klasického lineárního regresního modelu 16
2.2 Použití metody nejmenších čtverců k odhadu regresní funkce 19
2.3 Odhad regresní funkce metodou nejmenších čtverců 22
2.3.1 Jednoduché ukázky příkladů v Octave 22
2.3.2 Příklad programování regresní analýzy v R 24

3 Obecná metoda nejmenších čtverců **26**
3.1 Motivace zavádění 26

4 Aplikace metody nejmenších čtverců **31**
4.1 Řešení úlohy regresní analýzy pomocí matic 31
4.2 Řešení současných ekonomických úloh 32

Závěr **41**

Literatura **43**

Úvod

Tato bakalářská práce popisuje téma metody nejmenších čtverců a její aplikace na současných ekonomických úlohách. První kapitola se věnuje zavedení souvisejících pojmů, které v této práci budeme dále využívat, je vypracována na základě zdrojů: [1], [2], [3], [4], [5], [6], [7], [8]. Druhá kapitola se podrobněji zabývá užitím metody nejmenších čtverců v regresní analýze, tato kapitola je vytvořena na základě zdrojů: [1], [9], [10], [11], [12], [13]. Třetí kapitola je věnována metodě nejmenších čtverců, která využívá ortogonální polynomy k odhadu aproximující funkce, je vypracována na základě zdrojů: [1], [6], [7], [10], [12], [14], [15]. Poslední čtvrtá kapitola je zaměřena na aplikace metody nejmenších čtverců na současných ekonomických úlohách a dále také na konstrukce aproximujících funkcí pomocí ortogonálních polynomů. K sestavení poslední kapitoly byly využity zdroje: [10], [16], [17]. V této bakalářské práci jsou využívány programy: Octave, R a ke konstrukci obrázků pak program Geogebra.

Metoda nejmenších čtverců je univerzální metodou, jak aproximovat naměřená data. Tím rozumíme, jak proložit naměřená data křivkou, nebo jak řešit přeurčené soustavy rovnic. Autorství metody nejmenších čtverců je sporné. Datuje se k roku 1805 a nejčastěji se přisuzuje německému matematikovi Carlu Fridrichu Gaussovi (1777–1855). Ovšem nezávisle na něm tuto metodu objevil a následně publikoval její použití francouzský matematik Adrien Marie Legendre (1752–1833).

Metodu nejmenších čtverců využívá jednoduchá regresní analýza, což vedlo k výběru samotného tématu této bakalářské práce. Díky regresní analýze můžeme vyrovnávat vztahy zejména v oblasti ekonomických, společenských věd. Mnohdy totiž v ekonomických vědách, které se zabývají popisováním vztahů poptávkových, produkčních funkcí, nejsme schopni zajistit dlouhodobou úspěšnost jiných než lineárních regresních modelů, které sestavujeme v souladu s metodou nejmenších čtverců.

Kapitola 1

Základní pojmy

1.1 Vektorový prostor

Definice 1.1. Mějme pole P jehož prvky budeme označovat jako skaláry. Vytvoříme množinu V s operacemi:

1. sčítání $V \times V \to V$; $(\mathbf{u}, \mathbf{v}) \mapsto \mathbf{u} + \mathbf{v}$; $\mathbf{u}, \mathbf{v} \in V$
2. násobení skalárem $P \times V \to V$; $(p, \mathbf{u}) \mapsto p \cdot \mathbf{u}$; $p \in P, \mathbf{u} \in V$.

Tak, že pro každé $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ a $p, q \in P$ platí:

$$\begin{aligned} & \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u} && \exists 1 \in P, \quad 1 \cdot \mathbf{u} = \mathbf{u} \\ & \mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w} && p \cdot (q \cdot \mathbf{u}) = (p \cdot q) \cdot \mathbf{u} \\ \exists \mathbf{0} \in V, \quad & \mathbf{u} + \mathbf{0} = \mathbf{0} && (p + q) \cdot \mathbf{u} = (p \cdot \mathbf{u}) + (q \cdot \mathbf{u}) \\ \exists -\mathbf{u} \in V, \quad & \mathbf{u} + (-\mathbf{u}) = \mathbf{0} && p \cdot (\mathbf{u} + \mathbf{v}) = (p \cdot \mathbf{u}) + (p \cdot \mathbf{v}) \end{aligned}$$

Prvek $\mathbf{0}$ nazýváme nulový vektor a prvek $-\mathbf{u}$ nazýváme opačný k prvku $\mathbf{u}$. Množinu s těmito vlastnostmi nazveme vektorový prostor nad polem P a její prvky vektory.

Lineární závislost

Definice 1.2. Nechť V je vektorový prostor nad polem P a $\mathbf{u}_1, \ldots, \mathbf{u}_n \in V$. Pak řekneme, že vektory $\mathbf{u}_1, \ldots, \mathbf{u}_n$ jsou lineárně nezávislé, jestliže existují prvky $c_1, c_2, \ldots, c_n \in P$ tak, že pouze pro $c_1 = c_2 = \ldots = c_n = 0$ platí:

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_n\mathbf{u}_n = \mathbf{0}.$$

Poznámka 1.1. Pokud nejsou lineárně nezávislé říkáme, že jsou lineárně závislé.

Podprostor vektorového prostoru V

Řekneme, že $M \subseteq V$ je podprostor vektorového prostoru V, pokud pro $\forall\, \mathbf{u}, \mathbf{v} \in M$ a $p \in P$ platí:

$$\mathbf{u} + \mathbf{v} \in M \quad \text{a} \quad p \cdot \mathbf{u} \in M.$$

Lineární obal množiny

Nechť V je vektorový prostor nad polem P. Lineárním obalem podmnožiny M prostoru V rozumíme průnik všech podprostorů prostoru V, které množinu M obsahují, značíme $[M]$.

Poznámka 1.2. Pro každý vektor $\mathbf{v} \in [M]$ platí: $\exists\, \mathbf{u}_1, \ldots, \mathbf{u}_n \in M$ tak, že $\mathbf{v} = a_1\mathbf{u}_1 + + a_2\mathbf{u}_2 + \ldots + a_n\mathbf{u}_n$.

Generátor vektorového prostoru

Systém generátorů prostoru V je každá podmnožina M prostoru V, jejíž lineární obal je celý prostor $V([M] = V)$. Říkáme, že M generuje V.

Báze vektorového prostoru

Vektory $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n \in V$ tvoří bázi vektorového prostoru V, pokud jsou lineárně nezávislé a platí:

$$[\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n] = V.$$

Vektory $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ generují vektorový prostor V.

Dimenze vektorového prostoru

Dimenze je rovna počtu prvků báze vektorového prostoru.

Skalární součin

Definice 1.3. Nechť V je vektorový prostor nad polem P. Skalárním součinem rozumíme zobrazení $(\cdot\,,\cdot) : V \times V \to \mathbb{R}$, která ke každé dvojici vektorů $\mathbf{u}, \mathbf{v} \in V$ přiřazuje reálné číslo (skalár) tak, že platí následující vztahy:

$$\begin{aligned} (\mathbf{u}\,,\mathbf{v}) &= (\mathbf{v}\,,\mathbf{u}),\\ (\mathbf{u}\,,\mathbf{v}+\mathbf{w}) &= (\mathbf{u}\,,\mathbf{v}) + (\mathbf{u}\,,\mathbf{w}),\\ \forall\, k \in P,\ (k\cdot\mathbf{u}\,,\mathbf{v}) &= k\cdot(\mathbf{u}\,,\mathbf{v}),\\ (\mathbf{u}\,,\mathbf{v}) &\geqq 0 \wedge [(\mathbf{u}\,,\mathbf{u}) = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}]. \end{aligned}$$

Vybrané příklady skalárních součinů

1. *Euklidovský skalární součin*
   Nechť $V = \mathbb{R}^n$, pak
   $$(\mathbf{u}\,,\mathbf{v}) = u_1v_1 + u_2v_2 + \ldots + u_nv_n = \mathbf{u}^T\mathbf{v},$$
   kde $\mathbf{u} = (u_1, u_2, \ldots, u_n)^T$ a $\mathbf{v} = (v_1, v_2, \ldots, v_n)^T$.
2. *Skalární součin spojitých funkcí*
   Nechť je dán prostor spojitých reálných funkcí na uzavřeném intervalu $\langle a, b\rangle$ a váhová funkce $\omega(x)$, pak
   $$(f(x)\,,g(x)) = \int_a^b \omega(x)\cdot f(x)\cdot g(x)\,\mathrm{d}(x), \tag{1.1}$$
   je skalární součin.

Poznámka 1.3. Důkazy toho, že se jedná o skalární součiny nalezneme v [7].

Euklidovský prostor

Vektorový prostor ve kterém je definovaný skalární součin se nazývá Euklidovský prostor.

Ortogonalita

Nechť je dán Euklidovský prostor V se skalárním součinem $(\mathbf{u}\,,\mathbf{v})$. Pak, pokud pro vektory $\mathbf{u}$, $\mathbf{v} \in V$ platí:

$$(\mathbf{u}\,,\mathbf{v}) = 0,$$

řekneme, že vektory $\mathbf{u}$ a $\mathbf{v}$ jsou ortogonální (kolmé). Značíme $\mathbf{u} \perp \mathbf{v}$.

Norma vektoru

Nechť je dán Euklidovský prostor V, pak výraz

$$\|\mathbf{u}\| = \sqrt{(\mathbf{u}\,,\mathbf{u})},$$

nazveme normou (velikostí) vektoru $\mathbf{u}$.

Poznámka 1.4. Tzv. Euklidovská norma je určená vztahem:

$$\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}.$$

1.2 Ortogonální polynomy

Polynomem stupně n budeme chápat výraz:

$$P_n = P_n(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots + a_n x^n, \quad a_0, a_1, \ldots, a_n \in \mathbb{R}.$$

Množina Π_n polynomů stupně nejvýše n se skalárním součinem na množině bodů $x_1, \ldots, x_n$

$$(P_1(x)\,, P_2(x)) = \sum_{i=1}^{n} \omega(x_i) \cdot P_1(x_i) \cdot P_2(x_i), \tag{1.2}$$

tvoří Euklidovský prostor.

Poznámka 1.5. V této práci uvažujeme jen diskrétní metodu nejmenších čtverců. Teorie ortogonálních polynomů je podrobněji popsána v [6]. Ortogonální systémy polynomů můžeme sestrojit pro každou váhovou funkci ω a každou tabulku n bodů.

Systém ortogonálních polynomů můžeme konstruovat pomocí Gram-Schmidtova ortogonalizačního procesu, nebo pomocí tříčlenného rekurentního vzorce:

$$P_{j+1}(x) = (x - \beta_j)P_j(x) - \gamma_j P_{j-1}(x), \quad j = 0, 1, \ldots,$$

kde klademe $\gamma_0 = 0$, $P_{-1}(x) \equiv 0$ a $P_0(x) = 1$. Koeficienty jsou dány vztahy:

$$\beta_j = \frac{(xP_j\,, P_j)}{\|P_j\|^2}, \qquad j = 0, 1, \ldots,$$

$$\gamma_j = \frac{\|P_j\|^2}{\|P_{j-1}\|^2}, \qquad j = 1, 2, \ldots .$$

Při užití diskrétní metody nejmenších čtverců můžeme použít zejména dva systémy ortogonálních polynomů: *Čebyševovy polynomy* a *Gramovy polynomy*, které mají významné minimalizační vlastnosti.

1.2.1 Čebyševovy polynomy

Čebyševovy polynomy T_n jsou ortogonální polynomy na uzavřeném intervalu $\langle -1, 1\rangle$ s váhovou funkcí $\omega(x) = \frac{1}{\sqrt{1-x^2}}$. Ortogonalita je určená vztahem:

$$\int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}}\,\mathrm{d}(x) = 0, \quad n, m \in \mathbb{N}_0.$$

Platí zde rekurentní vztah:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n \in \mathbb{N}_0,$$

tedy

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_2(x) = 2x^2 - 1.$$

1.2.2 Gramovy polynomy

Tvoří soustavu diskrétně ortogonálních polynomů $G_j(x), j = 0, 1, \ldots, m$ na ekvidistantní množině bodů $x_j = j - \frac{1}{2}m$, $j = 0, 1, \ldots, m$ s váhovou funkcí $\omega_j \equiv 1$. Tyto Gramovy polynomy jsou dány rekurentním vztahem:

$$G_{j-1}(x) = xG_j(x) - \frac{j^2[(m+1)^2 - j^2]}{4(4j^2-1)}G_{j-1}(x), \quad j \geqq 1, \tag{1.3}$$

kde $G_0(x) \equiv 1$, $G_1(x) = x$.

Poznámka 1.6. Gramovy polynomy G_j splňující vztah (1.3) mají u nejvyšší mocniny koeficient 1. Teorie Gramových polynomů je podrobněji popsána v [1].

Metoda nejmenších čtverců používá i další ortogonální polynomy mezi nimi jsou například *Hermitovy polynomy* nebo *Legendrovy polynomy*, tyto polynomy jsou podrobněji popsány např. v [1], [7]. V této práci se jimi už nebudeme zabývat, protože tato problematika by přesahovala rámec této práce.

Kapitola 2

Jednoduchá regrese

2.1 Obecný lineární regresní model

2.1.1 Cíle regresní analýzy

Motivací pro zavádění regresní analýzy je popsání charakteristických vztahů mezi veličinami pomocí konstrukce vhodně zvolených regresních modelů. Vhodný regresní model chápeme jako optimálně zvolené matematické vyjádření změn hodnot popisované proměnné pomocí regresní analýzy. Snahou při hledání optimální regresní funkce, která by co nejlépe zachycovala vysvětlovanou proměnnou, je dosáhnout maximální shody mezi skutečnými a vyrovnanými hodnotami. Kvalita regresní funkce se hodnotí stupněm shody naměřených hodnot s modelovými odhady. Regresní model je velmi citlivý na změny dat, často malá změna naměřených hodnot má velký dopad na změny hodnot odhadnutých. Tyto změny odhadnutých hodnot mohou vést až k samotné změně zvolené regresní funkce. Mějme $X_1, X_2, \ldots, X_k$ nezávislé vysvětlující proměnné, Y vysvětlovanou závislou veličinu, pak výše uvedený vztah lze zapsat:

$$Y = \varphi(X_1, X_2, \ldots, X_k),$$

kde φ je hledaná regresní funkce.

Poznámka 2.1. Regresí se rozumí jednostranné závislosti, které pozorujeme u konkrétních dat v systematických změnách podmíněných průměrů (nejjednodušší způsob vyjádření průběhu závislosti) hodnot popisované závislé proměnné Y při systematických změnách kombinací hodnot vysvětlujících, nezávislých proměnných $X_1, X_2, \ldots, X_k$.

V ideálním případě, by měla regresní funkce věrně zobrazovat vztahy mezi jednotlivými veličinami na základě zatimního poznání. Překážkami přesnosti zobrazování se zde stávají chyby v měření veličin a samozřejmě i volba příslušné regresní funkce. V úlohách, kde pracujeme s regresními modely, bývají často používané lineární regresní funkce, jako funkce, kterými popisujeme závislosti mezi veličinami X_i a Y. Důvodem je nejen snazší interpretace parametrů, ale především okolnost, že v oblasti nedostatečně zdůvodněných teorií je lineární zobrazení často rozumným a užitečným zjednodušením.

2.1.2 Klasifikace regresních modelů

Pod pojmem regresní funkce si tedy představujeme funkci, která zohledňuje systematické změny jiných náhodných veličin. Odhad typu regresní funkce, speciálně jejích jednotlivých

parametrů, je základním úkolem regresní analýzy. Odhadování regresní funkce je dále ztíženo skutečností, že na veličinu Y působí kromě $X_1, X_2, \ldots, X_k$ i další veličiny, které můžeme označit např. $X_{k+1}, X_{k+2}, \ldots$. Teoreticky jich může být až nekonečně mnoho a nejde zajistit, aby měly tyto hodnoty na veličinu Y jen drobný dopad. Proto všechny ostatní, neuvažované, rušivé složky souhrnně označujeme ε. Musíme si udělat představu o pravděpodobnostním chování rušivé složky, abychom správně odhadli regresní funkci. Pro jednoduchost předpokládejme, že střední hodnota ε je nulová. Takto definovaná střední hodnota vyjadřuje požadavek, aby očekávaný dopad působení rušivé složky na regresní model byl nulový. To si můžeme dovolit, protože předpokládáme, že celkové (očekávané) působení rušivé složky v regresním modelu je zachyceno v parametru posunutí regresní funkce. Odchylky od učiněných předpokladů rušivé složky mají negativní dopad na užitečnost regresního modelu. Ukažme si zápis obecného regresního modelu:

$$Y = \eta + \varepsilon, \tag{2.1}$$

kde $\eta = \eta(X_1, X_2, \ldots, X_k)$ a $\varepsilon = \varepsilon(X_{k+1}, X_{k+2}, \ldots)$. Pro jednotlivá pozorování platí:

$$Y_i = \eta_i + \varepsilon_i, \quad i = 1, 2, \ldots, n.$$

Zvláštní postavení mezi regresními modely mají modely lineární:

$$Y = \beta_0 + \beta_1 f_1 + \ldots + \beta_r f_r + \varepsilon, \tag{2.2}$$

kde jednotlivé funkce $f_1 = f_1(X_1, X_2, \ldots, X_k), f_2 = f_2(X_1, X_2, \ldots, X_k), \ldots$ $f_r(X_1, X_2, \ldots, X_k)$ nazýváme regresory. V lineárním regresním modelu předpokládáme, že mezi složkami η a ε platí součtový vztah plynoucí ze zápisu (2.1). Nespornou výhodou lineárního modelu je, že získáváme lepší představu o grafickém průběhu regresní funkce a dále snadnější interpretace výsledků. Vysvětlení, proč se nejčastěji v praxi setkáme s lineárním regresním modelem jsou následující:

- Zejména v oblasti společenských věd, kde se zabýváme větším počtem vysvětlujících proměnných X_i, používáme lineární regresní model z důvodu zjednodušení. Jeho výhodou je, že se s ním dobře početně pracuje.
- V ekonomických vědách, které se zabývají popisováním vztahů poptávkových, produkčních funkcí, nejsme schopni zajistit dlouhodobou úspěšnost jiných než lineárních regresních modelů.
- Při předpokladu normálního rozdělení všech proměnných a podmíněné střední hodnotě, která je ke všem kombinacím náhodných hodnot X_i lineární funkcí, je vhodné použít lineární regresní model.
- V některých případech je vhodné nelineární model převést transformací na lineární regresní model. Ovšem toto řešení nemusí vždy vést k požadovaným výsledkům. Uspokojivá shoda linearizovaných hodnot, totiž nemusí znamenat dobrou shodu s hodnotami původními.
- Zavedení lineárního modelu je mnohdy jediným rozumným řešením nelineárních úloh.

Při konstrukci kvalitního regresního modelu musíme zohledňovat vlastnosti vysvětlujících proměnných X_i a pravděpodobnostní chování náhodné složky ε. K odhadu neznámých parametrů v regresním modelu lze použít tzv. *metodu nejmenších čtverců*.

Mezi nejpoužívanější regresní modely se zaměřením na lineární modely patří:

Model lineární v parametrech

Jde o základní model, který je lineární z hlediska všech parametrů, odpovídající vztahu (2.2). Parametr β_0 je absolutní člen a parametry $\beta_1, \beta_2, \ldots, \beta_r$ jsou dílčí regresní koeficienty. Funkce $f_1, \ldots f_r$ představují libovolné známé funkce původních proměnných X_1, $X_2, \ldots, X_k$, pro které platí, že neobsahují žádné další neznámé parametry. Předpokládá se, že každá z k vysvětlujících proměnných je v regresním modelu zastoupená alespoň jedním z r regresorů $f_j, j = 1, 2, \ldots, r$. Pojem regresor užíváme z důvodu odlišení souboru původních hodnot proměnných od nově vytvořených hodnot regresorů. Například pro zcela lineární model platí $r = k$ a pro racionální celistvou nebo lomenou funkci stupně s, kde $k = 1$, platí $r = s$. Zařazením absolutního členu do rovnice je numericky výhodné a symbolizuje existenci neuvažovaných vlivů.
Ukažme si to na konkrétním zadání

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1^2 + \beta_4 X_2^2 + \varepsilon.$$

Vidíme, že předpis rovnice obsahuje $k = 2$ původních proměnných X_1, X_2 a $r = 4$ regresory $f_1, \ldots, f_4$. Při znalosti hodnot X_1, X_2 není těžké dopočítat hodnoty nových kvadratických proměnných.

Zcela lineární model

Ve zcela lineárním modelu se předpokládá součtový vliv všech činitelů, je zcela správné použít ho v úlohách s větším počtem vysvětlujících proměnných a v případech vícerozměrného normálního rozdělení uvažovaných náhodných veličin. Regresní funkce je tvaru:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon. \tag{2.3}$$

Tento lineární model obsahuje k proměnných, popisovanou proměnnou Y, náhodné složky chyb ε a neznámé parametry β_j. Zařazení dílčích hodnot $\beta_j, j = 1, 2, \ldots, k$ představuje očekávanou změnu Y při růstu proměnné X_i. Vyjadřuje o kolik se změní Y, když hodnota proměnné X_i vzroste o jednotku. Tento vliv na veličinu Y je splněn za podmínky, že jednotlivé proměnné X_i jsou vzájemně nezávislé.
Ukážeme si základní zcela lineární modely.

1. *Model konstanty*:
$$Y = \beta_0 + \varepsilon,$$
je takový model, kde neuvažujeme žádnou z proměnných $X_1, X_2, \ldots, X_k$. Popisovaná veličina Y tedy náhodně kolísá okolo konstanty β_0. Tato situace bývá uváděna jako výjimečný, limitní případ některých složitějších modelů. Ovšem v převážné většině případů předpokládáme, že regresní model obsahuje alespoň jednu veličinu.
2. *Model regresní přímky*:
Model s $k = 1$ vysvětlující proměnnou:
$$Y = \beta_0 + \beta_1 X + \varepsilon.$$

3. *Model regresní roviny*:
   Model s $k = 2$ vysvětlujícími proměnnými zapíšeme ve tvaru:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon.$$

   Zde můžeme interpretovat β_1 jako očekávanou změnu proměnné X_1 za předpokladu, že hodnota X_2 zůstává nezměněna, můžeme sledovat o kolik se změní veličina Y v závislosti na jednotkové změně X_1. Kdybychom předpokládali, že nezměněnou zůstane hodnota X_1, mohli bychom považovat β_2 za očekávanou změnu, která bude mít dopad na veličinu Y.

Modely racionální celistvé a lomené

Nejznámější v této skupině je model regresního polynomu s-tého stupně, který je lineární z hlediska parametrů, ale nelineární z hlediska proměnné X

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \ldots + \beta_s X^s + \varepsilon.$$

Častým modelem regresního polynomu je tzv. regresní parabola, která je polynomem druhého stupně $s = 2$.

Modely převoditelné transformací na lineární model

Někdy je výhodné předpokládat součinový typ regresního modelu.
Obecný předpis regresního modelu pro exponenciální, mocninné a další regresní funkce je ve tvaru:

$$Y = \eta\varepsilon,$$

kde η zastupuje regresní funkci a ε je složka rušivých hodnot. Časté je použití obecného lineárně-exponenciálního regresního modelu s $X_1, X_2, \ldots, X_k$, který lze zapsat ve tvaru:

$$\eta = \exp(\beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k + \varepsilon).$$

Modely nelineární vzhledem k parametrům

Ve skutečnosti vztahy mezi veličinami bývají většinou nelineární, proto musíme uvažovat regresní modely, které jsou nelineární z hlediska parametrů. V různých vědních oborech lze tyto modely třídit podle odlišných kritérií, například podle stupně a formy nelinearity. S nelineárními modely se nejčastěji setkáváme v oblasti ekonomických procesů, kde uvažujeme průběh spotřeby, poptávky a nebo investic.

2.1.3 Sestavování odhadů

V této části se budeme zabývat sestavováním odhadů regresních modelů ve dvou i vícerozměrném prostoru, to znamená, že budeme nahrazovat hodnoty Y_i náležící náhodnému vektoru $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)^T$ odhadnutými (teoretickými, vypočítanými) hodnotami $\widehat{Y}_i$. Ty získáme využitím realizací regresorů proměnných X_i, $i = 1, 2, \ldots, n$ do odhadované regresní funkce.

Konstrukce odhadu regresního modelu vychází z podstaty stochastických vztahů mezi veličinami a z existence rušivé složky ε, ale také ze všech ostatních chyb, kterých jsme se dopustili. Tudíž nejlepším možným odhadem je odhad, při kterém je celková chybovost ta nejmenší. Obtížnost však spočívá v určení nejmenší chybovosti. Pro názornost popíšeme vztahy na

funkci regresní přímky. Označme odhad regresní přímky $\widehat{Y} = b_0 + b_1 X$, kde b_0, b_1 jsou odhady parametrů β_0, β_1. Pro jednotlivá měření by předpis vypadal takto: $\widehat{Y}_i = b_0 + b_1 X_i$. Chyby odhadu, též nazývána jako rezidua (na obrázku 2.1), můžeme vyjádřit jako rozdíl $e_i = Y_i - \widehat{Y}_i$. Rezidua dále považujeme za odhad rušivé složky ε. Naší snahou je sestavit kvalitní odhad regresního modelu, k tomu nám může pomoci, že jsme schopni zajistit, aby součet reziduí byl nulový:

$$\sum_{i=1}^{n} e_i = 0.$$

To platí, pokud realizace odhadnuté přímky prochází bodem $[\bar{X}, \bar{Y}]$ aritmetických průměrů $\bar{X}, \bar{Y}$, které odpovídají průměrům veličin X_i a Y_i, $i = 1, \ldots, n$. Tento vztah odvodili na konkrétních příkladech v [9, s. 31]. Ovšem nulový součet reziduí k sestavení kvalitní regresní funkce nestačí, proto bereme v úvahu i kritérium reziduálního součtu čtverců pro určení regresních parametrů $\beta_j, j = 1, \ldots, k$,

$$Q(\mathbf{e}) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \widehat{Y}_i)^2.$$

K tomuto se používají různé postupy založené na *metodě nejmenších čtverců*. Metoda nejmenších čtverců přiřazuje všem pozorováním stejnou váhu. Abychom se vyhli problémům s rozdělením náhodné složky ε, je výhodnější dávat větším reziduím menší váhu, a tím znevýhodňovat výjimečná pozorování. Toto vede k použití různých dalších metod používaných k minimalizaci ε, například *metoda minimální hodnoty rezidua*, nebo *metoda součtu absolutních hodnot reziduí*. Tyto metody jsou podrobněji popsány v [11], ale my je zde dále řešit nebudeme, protože tato problematika je složitá a přesahovala by rámec této práce.

2.1.4 Formulace klasického lineárního regresního modelu

Pro naměřené hodnoty proměnných $X_1, X_2, \ldots, X_k$, je veličina Y náhodnou veličinou, kterou lze popsat lineární funkcí $k+1$ neznámých parametrů. Náhodnou veličinu Y dále popisujeme náhodným výběrem $\mathbf{Y}$, kde $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)^T$. Jednotlivé složky Y_i mají stejné rozložení jako původní náhodná veličina Y a platí:

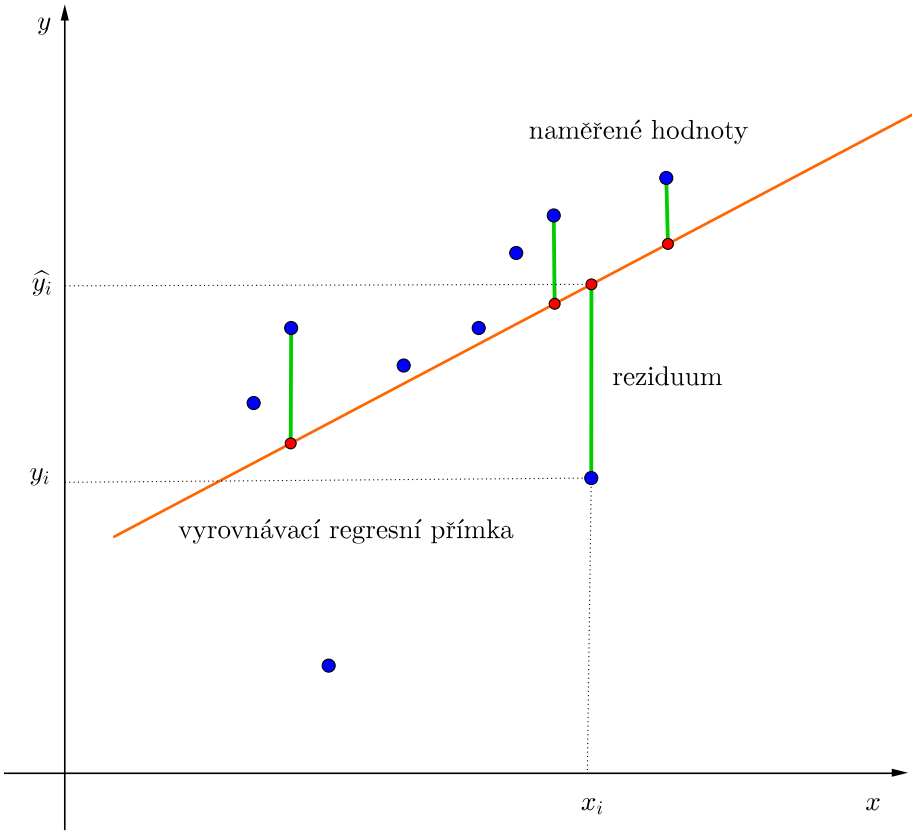
$$E(Y_i) = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik},$$

kde x_{ij} jsou nám známé realizace veličin $X_1, X_2, \ldots, X_k$. Dále, až na výjimky, předpokládáme pro každé i rozptyl $var(Y_i) = \sigma^2$, kde $\sigma > 0$. Rozepsání celého lineárního modelu se zapisuje ve formě lineárních rovnic:

$$\begin{aligned}
Y_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_k x_{1k} + \varepsilon_1 \\
Y_2 &= \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \beta_k x_{2k} + \varepsilon_2 \\
&\vdots \\
Y_n &= \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_k x_{nk} + \varepsilon_n,
\end{aligned}$$

nebo v maticovém vyjádření:

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \ldots & x_{1k} \\ 1 & x_{21} & \ldots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \ldots & x_{nk} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}, \tag{2.4}$$



Obrázek 2.1: Sestavení odhadu pomocí regresní přímky

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

$\mathbf{X}$ je nenáhodná matice o n řádcích a $k+1$ sloupcích (další přípustná označení: *regresní matice, matice plánu, matice modelu*), $\boldsymbol{\beta}$ je $(k+1)$členný vektor neznámých parametrů modelu a $\boldsymbol{\varepsilon}$ je vektor o n členech vyjadřující odchylky od modelu. Pro matici plánu:

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{nk} \end{pmatrix}$$

předpokládáme, že $n > k$ a o hodnosti matice $\mathbf{X}$ předpokládáme $h(\mathbf{X}) = k + 1$. Bude-li tento předpoklad splněn, říkáme, že jde o lineární regresní model plné hodnosti. V této práci předpokládáme modely plné hodnosti a dále s nimi pracujeme. V takovém případě jsou sloupce matice $\mathbf{X}$ nezávislé.

Ze zápisu (2.4) je vidět, že v n lineárních rovnicích je $p = k + 1$ neznámých regresních parametrů a n hodnot vektoru $\boldsymbol{\varepsilon}$. V případě lineární regresní funkce ve tvaru (2.2) dostáváme obdobnou situaci jen rovnic máme $p = r + 1$. Jediný rozdíl oproti modelu $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$

je, že matici plánu $\mathbf{X}$ o rozměrech $n \times (k+1)$, nahrazuje matice $\mathbf{F}$ o rozměrech $n \times (r+1)$, kde r je počet známých funkcí (regresorů) původních k vysvětlujících proměnných. Obecný tvar modelu $\mathbf{Y} = \mathbf{F}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ obsahuje $p = r + 1$ neznámých parametrů a n hodnot náhodné rušivé složky. Zcela lineární regresní model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, nebo obecný lineární regresní model $\mathbf{Y} = \mathbf{F}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ můžeme za určitých podmínek považovat za *klasický lineární regresní model*. Pro jednoduchost uvažujeme dále jen model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$.

Vlastnosti klasického lineárního modelu
Sloupce matice plánu $\mathbf{X}$ jsou lineárně nezávislé, $h(\mathbf{X}) = k + 1 = p \leq n$, kde n představuje počet pozorování, p je počet regresních parametrů:

$$h(\mathbf{X}) = p.$$

Matice $\mathbf{X}^T\mathbf{X}$ je čtvercová, symetrická, s hodností

$$h(\mathbf{X}^T\mathbf{X}) = k + 1 = p.$$

Je tedy regulární a existuje k ní inverzní matice $(\mathbf{X}^T\mathbf{X})^{-1}$.
Odvození 1.

- Matice je symetrická, pokud platí $\mathbf{A}^T = \mathbf{A}$. Pro matici $\mathbf{X}^T\mathbf{X}$ platí:

  $$(\mathbf{X}^T\mathbf{X})^T = \mathbf{X}^T(\mathbf{X}^T)^T = \mathbf{X}^T\mathbf{X},$$

  je tedy symetrická.
- Důkaz $h(\mathbf{X}^T\mathbf{X}) = h(\mathbf{X}) = p$ je např. v [10, s. 62].

Při odhadování hodnot parametrů vycházíme ze zkušenosti, nebo z předpokladů o hodnotách jednotlivých složek vektoru $\boldsymbol{\beta}$. Pro zlepšování regresních modelů se doporučuje použití dalších doplňkových informací. Budeme uvažovat hodnoty rušivé složky $\boldsymbol{\varepsilon}$:

$$\boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix},$$

které jsou navzájem nezávislé, mají normální rozdělení a jsou rozptýleny okolo nulové střední hodnoty:

$$E(\varepsilon_i) = 0, \quad i = 1, 2, \ldots, n.$$

Protože jsou hodnoty ε_i a ε_j navzájem nezávislé (nekorelované), je zajištěna nulová kovariance $cov(\varepsilon_i, \varepsilon_j) = 0$, pro $i \neq j = 1, 2, \ldots, n$. Pro všechna ε_i předpokládáme rozptyl $D(\varepsilon_i) = \sigma^2$. Obecně má vektor $\boldsymbol{\varepsilon}$ n-rozměrné normální rozdělení $N_n(\mathbf{0}, \sigma^2\mathbf{I})$. Tyto předpoklady umožňují zápis kovarianční matice náhodného vektoru $\boldsymbol{\varepsilon}$ ve tvaru:

$$cov(\boldsymbol{\varepsilon}) = \begin{pmatrix} \sigma^2 & 0 & \ldots & 0 \\ 0 & \sigma^2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma^2 \end{pmatrix} = \sigma^2\mathbf{I}.$$

Náhodné veličiny $Y_i, i = 1, 2, \ldots, n$, jsou nezávislé a mají normální rozdělení se střední hodnotou η_i a rozptylem σ^2. Kovarianční matice je tedy ve tvaru:

$$cov(\mathbf{Y}) = cov(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}.$$

Střední hodnotu vektoru $\mathbf{Y}$ lze jednoduše odvodit:

Odvození 2.

$$E(\mathbf{Y}) = E(\mathbf{X}\boldsymbol{\beta}) + E(\boldsymbol{\varepsilon}) = E(\mathbf{X}\boldsymbol{\beta}) = \mathbf{X}\boldsymbol{\beta} = \eta.$$

Kovarianční matici vektoru $\mathbf{Y}$ také jednoduše odvodíme:

Odvození 3.

$$cov(\mathbf{Y}) = cov(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) = cov(\varepsilon) = \sigma^2\mathbf{I}.$$

Z výše uvedeného plyne, že vektor $\mathbf{Y}$ má n-rozměrné normální rozdělení $N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$.

2.2 Použití metody nejmenších čtverců k odhadu regresní funkce

Mějme regresní model:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}.$$

Vektor $\mathbf{b}$ odhadů parametrů $\boldsymbol{\beta}$ určujeme tak, aby byla co nejmenší délka vektoru reziduí $\mathbf{e} = \mathbf{Y} - \mathbf{X}\mathbf{b}$. Tedy naší snahou je minimalizovat délku všech odchylek $\boldsymbol{\varepsilon} = \mathbf{Y} - \mathbf{X}\boldsymbol{\beta}$. Ukazuje se jako výhodnější místo určování minima normy vektoru $\boldsymbol{\varepsilon}$, minimalizovat druhou mocninu normy, tedy výraz:

$$Q(\boldsymbol{\varepsilon}) = \sum_{i=1}^{n} \varepsilon_i^2 = \boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} = \| \boldsymbol{\varepsilon} \|^2 = \| \mathbf{Y} - \mathbf{X}\boldsymbol{\beta} \|^2,$$

minimalizujeme v bodě $\mathbf{b} = \mathbf{b}(\mathbf{Y})$. Funkce $Q(\boldsymbol{\varepsilon})$, kterou nazýváme reziduální součet čtverců, nabývá svého minima pro:

$$\| \mathbf{Y} - \mathbf{X}\mathbf{b} \|^2 = \min_{\boldsymbol{\beta} \in \mathbb{R}^{k+1}} \| \mathbf{Y} - \mathbf{X}\boldsymbol{\beta} \|^2 .$$

Ukázku reziduálního součtu čtverců vidíme na obrázku (2.2).

Věta 2.1. Najít minimum funkce $Q(\varepsilon)$ znamená řešit soustavu tzv. normálních rovnic

$$\mathbf{X}^T\mathbf{Y} = \mathbf{X}^T\mathbf{X}\boldsymbol{\beta},$$

kde

$$\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}, \tag{2.5}$$

je odhad (konkrétní jedno řešení) $\boldsymbol{\beta}$ pořízený metodou nejmenších čtverců.

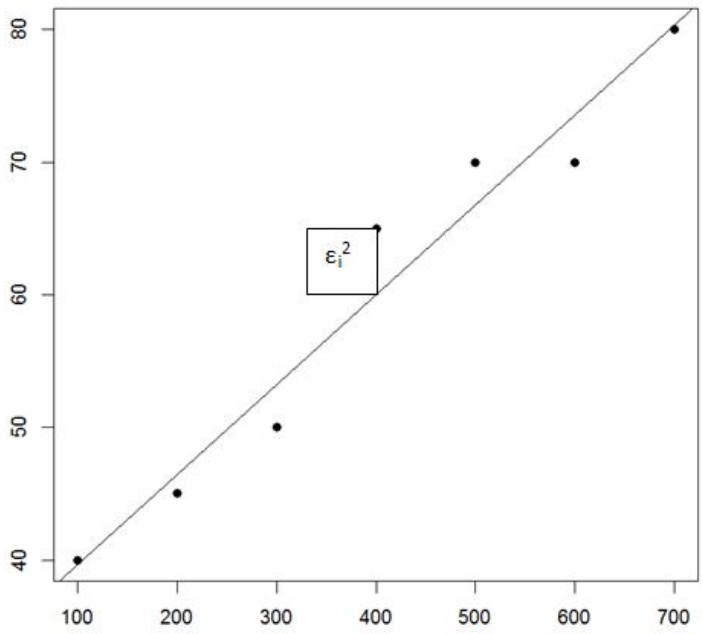
Poznámka 2.2. Pro řešení $\mathbf{b}$ soustavy normálních rovnic:

$$\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{Y},$$

platí

$$\mathbf{X}^T(\mathbf{X}\mathbf{b} - \mathbf{Y}) = \mathbf{0} \quad \text{resp.} \quad (\mathbf{Y} - \mathbf{X}\mathbf{b})^T\mathbf{X} = \mathbf{0}. \tag{2.6}$$

Této vlastnosti využijeme v důkazu věty (2.1).



Obrázek 2.2: Ukázka reziduálního součtu čtverců ve dvourozměrném prostoru

Důkaz.

1. Důkaz implikace $(\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}) \Rightarrow (\mathbf{b} = \arg\min Q(\boldsymbol{\varepsilon}))$: „*Pokud získáme odhad* $\boldsymbol{b}$ *jako řešení normálních rovnic, pak je minimem metody nejmenších čtverců.*“

$$Q(\boldsymbol{\varepsilon}) = \| \mathbf{Y} - \mathbf{X}\boldsymbol{\beta} \|^2 = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) = (\mathbf{Y} - \mathbf{X}\mathbf{b} + \mathbf{X}\mathbf{b} - \mathbf{X}\boldsymbol{\beta})^T$$
$$(\mathbf{Y} - \mathbf{X}\mathbf{b} + \mathbf{X}\mathbf{b} - \mathbf{X}\boldsymbol{\beta}) = (\mathbf{Y} - \mathbf{X}\mathbf{b})^T(\mathbf{Y} - \mathbf{X}\mathbf{b}) + (\mathbf{Y} - \mathbf{X}\mathbf{b})^T(\mathbf{X}\mathbf{b} - \mathbf{X}\boldsymbol{\beta}) +$$
$$+ (\mathbf{X}\mathbf{b} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{Y} - \mathbf{X}\mathbf{b}) + (\mathbf{X}\mathbf{b} - \mathbf{X}\boldsymbol{\beta})^T)(\mathbf{X}\mathbf{b} - \mathbf{X}\boldsymbol{\beta}).$$

Je nám známo, že $\mathbf{b}$ je řešením normálních rovnic a splňuje (2.6), pak když výraz:

$$(\mathbf{Y} - \mathbf{X}\mathbf{b})^T(\mathbf{X}\mathbf{b} - \mathbf{X}\boldsymbol{\beta})$$

upravíme na:

$$(\mathbf{Y} - \mathbf{X}\mathbf{b})^T\mathbf{X}(\mathbf{b} - \boldsymbol{\beta})$$

vidíme, že je roven nule. Stejně pro výraz:

$$(\mathbf{X}\mathbf{b} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{Y} - \mathbf{X}\mathbf{b}),$$

platí:

$$(\mathbf{b} - \boldsymbol{\beta})^T\mathbf{X}^T(\mathbf{Y} - \mathbf{X}\mathbf{b}) = 0.$$

Po upravení pro zbylý sčítanec platí:

$$(\mathbf{b} - \boldsymbol{\beta})^T\mathbf{X}^T\mathbf{X}(\mathbf{b} - \boldsymbol{\beta}) = \mathbf{z}^T\mathbf{z} = \sum_{i=1}^{n} z_i^2 \geqq 0.$$

Pro funkci $Q(\boldsymbol{\varepsilon})$ platí:

$$\begin{aligned} Q(\boldsymbol{\varepsilon}) =& \| \mathbf{Y} - \mathbf{X}\boldsymbol{\beta} \|^2 = \\ &= (\mathbf{Y} - \mathbf{X}\mathbf{b}^T)(\mathbf{Y} - \mathbf{X}\mathbf{b}) + (\mathbf{b} - \boldsymbol{\beta})^T\mathbf{X}^T\mathbf{X}(\mathbf{b} - \boldsymbol{\beta}) \geqq \\ &\geqq (\mathbf{Y} - \mathbf{X}\mathbf{b}^T)(\mathbf{Y} - \mathbf{X}\mathbf{b}) = Q(\mathbf{e}). \end{aligned}$$

A tím je tedy dokázáno, že $Q(\boldsymbol{\varepsilon})$ nabývá svého minima pro odhad $\mathbf{b}$,

$$Q(\boldsymbol{\varepsilon}) \ \geqq \ Q(\mathbf{e}).$$

2. Nyní musíme dokázat opačnou implikaci $(\mathbf{b} = \arg\min Q(\boldsymbol{\varepsilon})) \ \Rightarrow \ (\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y})$: „*Nechť $Q(\boldsymbol{\varepsilon})$ nabývá svého minima v bodě $\boldsymbol{b}$, pak je $\boldsymbol{b}$ řešením normálních rovnic.*“ Vycházíme ze skutečnosti, že $\mathbf{b}$ je minimem funkce $Q(\boldsymbol{\varepsilon})$, pak musí platit:

$$\begin{aligned} \left.\frac{\partial Q(\boldsymbol{\varepsilon})}{\partial \beta_0}\right|_{\mathbf{b}} &= 2\sum_{i=1}^{n}(Y_i - \sum_{j=1}^{k+1} x_{ij}b_j)\cdot(-x_{i1}) = 0, \\ &\vdots \\ \left.\frac{\partial Q(\boldsymbol{\varepsilon})}{\partial \beta_k}\right|_{\mathbf{b}} &= 2\sum_{i=1}^{n}(Y_i - \sum_{j=1}^{k+1} x_{ij}b_j)\cdot(-x_{ik}) = 0, \end{aligned}$$

kde x_{ij} značí prvek v i-tém řádku a j-tém sloupci matice $\mathbf{X}$. Výše uvedený systém $k+1$ rovnic můžeme zapsat maticově:

$$\begin{aligned} (\mathbf{Y} - \mathbf{X}\mathbf{b})^T[\mathbf{X}]_{.1} &= 0, \\ &\vdots \\ (\mathbf{Y} - \mathbf{X}\mathbf{b})^T[\mathbf{X}]_{.k+1} &= 0, \end{aligned}$$

kde $[\mathbf{X}]_{.j}$ značí j-tý sloupec matice $\mathbf{X}$. Pak

$$\begin{aligned} (\mathbf{Y} - \mathbf{X}\mathbf{b})^T\mathbf{X} &= \mathbf{0}, \\ \mathbf{X}^T\mathbf{X}\mathbf{b} &= \mathbf{X}^T\mathbf{Y}. \end{aligned}$$

□

Při sestavování regresních modelů nás dále zajímá, jak velkou část variability závislé (vysvětlované) proměnné objasňuje regresní model.
Označme:

$$S(\mathbf{Y}) = \sum_{i=1}^{n}(Y_i - \overline{Y})^2 = \sum_{i=1}^{n} Y_i^2 - n\overline{Y}^2 \quad \text{celkový součet čtverců,}$$

pak výraz:

$$R^2 = 1 - \frac{Q(\mathbf{e})}{S(\mathbf{Y})}$$

nazýváme koeficient determinace a udává, kolik procent rozptylu vysvětlované proměnné je vysvětleno modelem a kolik zůstalo nevysvětleno (jak těsná je regresní závislost). Nabývá hodnot od nuly do jedné (teoreticky i včetně těchto krajních mezí), přičemž čím blíže je hodnota R^2 rovna jedné, tím je regresní model kvalitnější. Je zřejmé, že při snaze sestavit kvalitní regresní model dbáme na to, aby hodnota $S(\mathbf{Y})$ nabývala větších hodnot než hodnota $Q(\mathbf{e})$, kterou se snažíme minimalizovat.

2.3 Odhad regresní funkce metodou nejmenších čtverců

2.3.1 Jednoduché ukázky příkladů v Octave

1. **Použití metody nejmenších čtverců na pořízení odhadu lineární regresní funkce v programu Octave.**

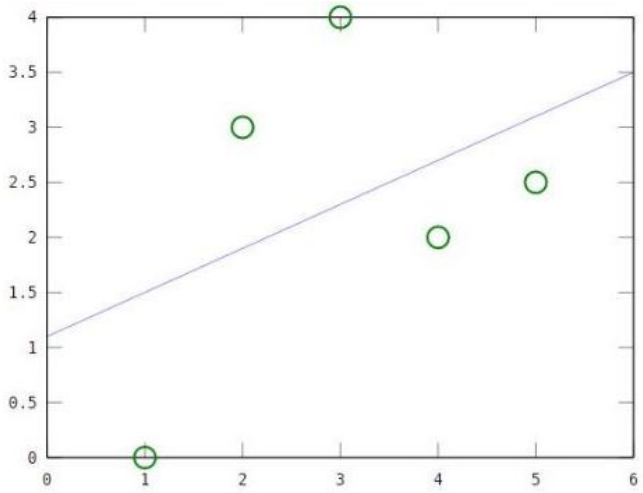
   Budeme vyšetřovat lineární závislost mezi veličinou Y a X, kde y_i a x_i jsou naměřené realizace náhodného výběru z těchto dvou náhodných veličin. Snažíme se tedy naměřená data proložit tak, aby reziduální součet čtverců $Q(\boldsymbol{\varepsilon})$ byl minimální. Experimentálním měřením jsme naměřili data $x_i = 1, \ldots, 5$ a jejich funkční hodnoty $y_i = 0, 3, 4, 2, 2, 5$. Začínáme příkazem `m = 1`, který představuje požadavek aproximace dat regresní přímkou a dále načteme data x_i, y_i. Pro zjednodušení použijeme příkaz: `p = polyfit (x,y,m)`, který najde koeficienty polynomu $p(x)$ stupně m, který odpovídá naměřeným hodnotám x_i, y_i ve smyslu metody nejmenších čtverců. Vzhledem k naměřeným hodnotám x_i použijeme měřítko pro vykreslování grafu příkazem: `xi = 0:.1:6`, je výhodné použít hodnoty 1 až 6 pro samotné vykreslení grafu, který nebude zalomen za pátou hodnotou. Pro vykreslení lineární regresní funkce zadáme příkaz:
   `plot(xi,polyval(p,xi),x,y,'o')`, který vykreslí graf na základě naměřených hodnot a vypočteného polynomu $p(x)$.

```
octave:> m = 1
octave:> x=[1 2 3 4 5]
octave:> y=[0 3 4 2 2.5]
octave:> p=polyfit(x,y,m)
p = 0.40000   1.10000
octave:> xi=0:.1:6
octave:> plot(xi,polyval(p,xi),x,y,'o')
```

   Tímto postupem jsme našli předpis regresní přímky $y = 0,4x + 1,1$, která je znázorněna na grafu (2.3).

2. **Použití metody nejmenších čtverců na pořízení odhadu kvadratické regresní funkce v programu Octave.**

   Budeme vyšetřovat kvadratickou závislost mezi veličinou Y a X z předchozího příkladu. Tentokrát začneme příkazem `m = 2`, který představuje požadavek aproximace dat regresní parabolou a dále načteme data x_i, y_i. Další postup je shodný s postupem vyšetřování výše uvedené lineární závislosti.



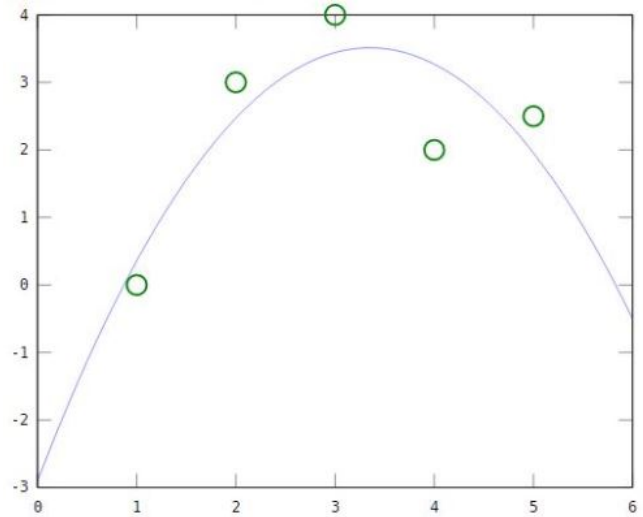
Obrázek 2.3: Ukázka lineární regrese metodou nejmenších čtverců v Octave

```
octave:> m=2

octave:> p=polyfit(x,y,m)
p = -0.57143   3.82857  -2.90000

octave:> plot(xi,polyval(p,xi),x,y,'o')
```

Našli jsme předpis kvadratické regresní funkce $y = -0,57x^2 + 3,83x - 2,90$, která je znázorněna na grafu (2.4).



Obrázek 2.4: Ukázka kvadratické regrese metodou nejmenších čtverců v Octave

2.3.2 Příklad programování regresní analýzy v R

Nyní si uvedeme příklad lineárního regresního modelu v ekonomických vědách, kdy je výhodné předpokládat lineární závislost ($y = b_1x + b_0$) mezi veličinami X a Y. V tabulce 2.1 jsou výchozí data dostupná z [13], která využijeme pro aplikování regresní analýzy v programu R. Zajímáme se o to, jestli existuje závislost mezi průměrnou výší úrokových měr a počtem uzavřených smluv o hypotečních úvěrech v České republice v letech 2009–2012. Načteme si tudíž tabulku 2.1 do programu R a zkusíme vyšetřit závislost uzavíraných smluv o hypotečních úvěrech na změně průměrné úrokové míry. Předpokládali jsme lineární závislost mezi počtem uzavřených smluv a průměrnou úrokovou mírou v jednotlivých letech. Hodnoty `Intercept` a `Coefficients` vyjadřují odhady parametrůb_0, b_1. Příkazem `summary(model)` dostaneme podrobnější informace o regresní závislosti. Hodnota `Pr(>|t|)` představuje porovnání s hladinou statistické významnosti b_0, b_1. Hodnota `Multiple R-squared` vyjadřuje koeficient determinace (čím blíže je číslu 1, tím je efektivnost lineární regrese větší). Hodnota `Adjusted R-squared` je upravená hodnota testového kritéria a hodnota `p-value: 0.007258` představuje porovnání s hladinou statistické významnosti, v tomto případě zamítáme hypotézu, že koeficient determinace je roven nule, tudíž koeficient determinace je různý od nuly a je tedy dokázána závislost mezi průměrnou výší úrokových měr a počtem uzavřených smluv o hypotečních úvěrech v České republice za období 2009–2012.

Rok	Prům. sazba	Počet uzav. hypoték	Objem h. v mld. Kč
Prosinec 2012	3,17 %	7 637	13,079
Prosinec 2011	3,56 %	7 293	12,847
Prosinec 2010	4,23 %	6 106	10,289
Prosinec 2009	5,61 %	3 575	5,953

Tabulka 2.1: Tabulka vývoje hypotečních úvěrů v období 2009 až 2012

```
> hypoteky
Prumer.ur.mira Objemhyp.uveruvmld. Pocetuzavrenychhypotek
1           3.17                7637                 13.079
2           3.56                7293                 12.847
3           4.23                6106                 10.289
4           5.61                3575                  5.953
> lm(Pocetuzavrenychhypotek ~ Prumer.ur.mira)
Call:
lm(formula = Pocetuzavrenychhypotek ~ Prumer.ur.mira)
Coefficients:
(Intercept)  Prumer.ur.mira
23.243          -3.066
> model <-lm(Pocetuzavrenychhypotek ~ Prumer.ur.mira)
> summary(model)
Call:
lm(formula = Pocetuzavrenychhypotek ~ Prumer.ur.mira)
Residuals:
1        2        3        4
-0.44472  0.51904  0.01528 -0.08960
```

```
Coefficients:   Estimate  Std. Error       t   value Pr(>|t|)
(Intercept)     23.2430     1.1149     20.85  0.00229 **
Prumer.ur.mira  -3.0660     0.2626    -11.67  0.00726 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual standard error: 0.4876 on 2 degrees of freedom
Multiple R-squared:  0.9855,    Adjusted R-squared:  0.9783
F-statistic: 136.3 on 1 and 2 DF,  p-value: 0.007258
```

Kapitola 3

Obecná metoda nejmenších čtverců

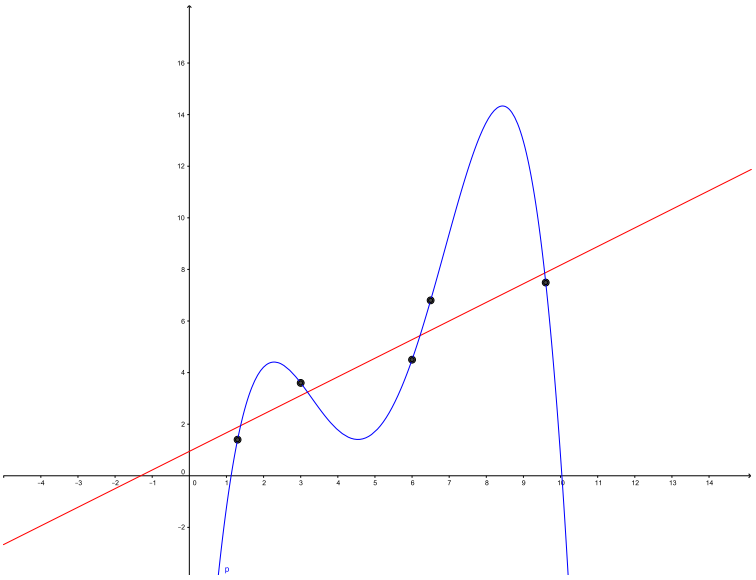
V návaznosti na předešlou kapitolu budeme uvažovat vztah:

$$y_i = \beta_0\varphi_0(x_i) + \beta_1\varphi_1(x_i) + \cdots + \beta_k\varphi_k(x_i), \quad i = 1, \ldots, n, \tag{3.1}$$

kde $\{\varphi_j(x)\}_{j=0}^{k}$ představují bázové vektory prostoru V. Naším cílem je určit koeficienty $\{\beta_j\}_{j=0}^{k}$.

3.1 Motivace zavádění

Metoda nejmenších čtverců je statistická (numerická) metoda, která se používá v případech, kdy máme naměřeny hodnoty $(x_i, y_i), i = 0, 1, \ldots, n$, které chceme aproximovat polynomem stupně nejvýše n.



Obrázek 3.1: Ukázka aproximace dat metodou nejmenších čtverců

V případě, že hledáme polynom:

$$P_n(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n,$$

který prochází body $(x_i, y_i), i = 0, 1, \ldots, n$, platí:

$$\begin{aligned} y_0 =& a_0 + a_1 x_0 + a_2 x_0^2 + \ldots + a_n x_0^n \\ y_1 =& a_0 + a_1 x_1 + a_2 x_1^2 + \ldots + a_n x_1^n \\ & \vdots \\ y_n =& a_0 + a_1 x_n + a_2 x_n^2 + \ldots + a_n x_n^n. \end{aligned}$$

Dostáváme $n + 1$ lineárních rovnic o $n + 1$ neznámých. Pokud budou jednotlivá měření nezávislá, pak má tato soustava řešení. Jsme tedy schopni určit koeficienty $a_0, \ldots, a_n$. Obecně ale vždy nepotřebujeme, aby hledaný polynom naměřenými hodnotami procházel, stačí nám pouze aproximace polynomem nižšího stupně ($k < n$) odpovídající vztahu (3.1), (viz obrázek 3.1). Chceme tedy data pouze přibližně aproximovat (z důvodu zjednodušení, víme, že jsou zatíženy chybou, atd.), proto obdržíme soustavu $n+1$ rovnic o $k+1$ neznámých. Pak získáváme tzv. *přeurčenou soustavu rovnic*. Při určování koeficientů β_j musíme dbát na to, abychom co nejlépe potlačili chyby v měření, které představuje soubor náhodných, rušivých složek ε, proto dodržujeme zásadu, že počet měření $(x_i, y_i)_{i=0}^n$ musí být větší než počet $\{\beta_j\}_{j=0}^k$ koeficientů, platí:

$$n \geq k + 1.$$

Abychom mohli vztah (3.1) zapsat maticově, zavedeme označení:

$$\boldsymbol{\varphi}_j = (\varphi_0(x), \varphi_1(x), \ldots, \varphi_k(x))^T, \quad j = 0, 1, \ldots, k,$$

$$\mathbf{A} = (\boldsymbol{\varphi}_0, \boldsymbol{\varphi}_1, \ldots, \boldsymbol{\varphi}_k), \quad \boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_k)^T, \quad \mathbf{y} = (y_1, y_2, \ldots, y_n)^T.$$

Přibližné splnění rovnic (3.1) lze zapsat ve tvaru $\mathbf{y} \approx \mathbf{A}\boldsymbol{\beta}$. Míru nesplnění rovnic (3.1) vyjádříme pomocí rezidua. Definujeme nové označení pro vektor reziduí:

$$\mathbf{r} = \mathbf{r}(\boldsymbol{\beta}) := \mathbf{y} - \mathbf{A}\boldsymbol{\beta}.$$

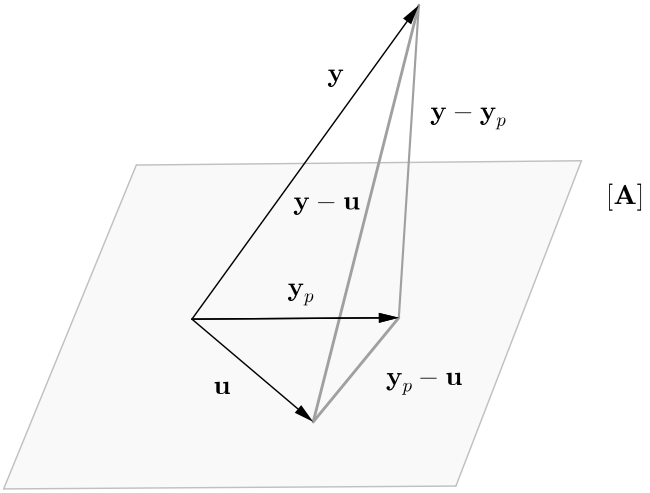
Naším cílem je minimalizovat délku vektoru rezidua. Tedy minimalizovat normu:

$$\|\mathbf{y} - \mathbf{A}\boldsymbol{\beta}\|,$$

což je ekvivalentní s minimalizací $\|\mathbf{y} - \mathbf{A}\boldsymbol{\beta}\|^2$.
Výraz $\|\mathbf{y} - \mathbf{A}\boldsymbol{\beta}\|^2$ je minimální, pokud je $\mathbf{A}\boldsymbol{\beta}$ ortogonální projekcí vektoru $\mathbf{y}$ do prostoru generovaného sloupci matice $\mathbf{A}$, tedy do prostoru $[\boldsymbol{\varphi}_0, \boldsymbol{\varphi}_1, \ldots, \boldsymbol{\varphi}_k]$. Platí $(\mathbf{y} - \mathbf{A}\boldsymbol{\beta}) \perp [\mathbf{A}]$, kde $[\mathbf{A}]$ značí $[\boldsymbol{\varphi}_0, \boldsymbol{\varphi}_1, \ldots, \boldsymbol{\varphi}_k]$.

Odvození 4. Označíme $\mathbf{y}_p$ kolmou projekcí vektoru $\mathbf{y}$ do $[\mathbf{A}]$ a $\mathbf{u}$ libovolnou projekcí $\mathbf{y}$ do prostoru $[\mathbf{A}]$. Pak vektor $\mathbf{y} - \mathbf{y}_p$ je ortogonální na libovolný vektor v prostoru $[\mathbf{A}]$, speciálně na vektor $\mathbf{y}_p - \mathbf{u}$ (viz obrázek 3.2), tedy $(\mathbf{y} - \mathbf{y}_p, \mathbf{y}_p - \mathbf{u}) = 0$. Pak platí:

$$\|\mathbf{y} - \mathbf{u}\|^2 = \|\mathbf{y} - \mathbf{y}_p + \mathbf{y}_p - \mathbf{u}\|^2 = \|\mathbf{y} - \mathbf{y}_p\|^2 + \|\mathbf{y}_p - \mathbf{u}\|^2 + 2(\mathbf{y} - \mathbf{y}_p, \mathbf{y}_p - \mathbf{u}) = \|\mathbf{y} - \mathbf{y}_p\|^2 + \|\mathbf{y}_p - \mathbf{u}\|^2.$$



Obrázek 3.2: Projekce vektorů do prostoru [**A**]

Pro libovolné $\mathbf{u}$ je výraz $\|\mathbf{y}-\mathbf{u}\|^2$ minimální, pokud $\|\mathbf{y}_p-\mathbf{u}\|^2=\mathbf{0}$:

$$\begin{aligned}\|\mathbf{y}_p-\mathbf{u}\|^2&=\mathbf{0}\\ \mathbf{y}_p-\mathbf{u}&=\mathbf{0}\\ \mathbf{u}&=\mathbf{y}_p,\end{aligned}$$

tedy v případě, že libovolná projekce je kolmou projekcí.
Pokud je tedy $\mathbf{A}\boldsymbol{\beta}$ kolmou projekcí $\mathbf{y}$ do [**A**], pak je vektor $(\mathbf{y}-\mathbf{A}\boldsymbol{\beta})$ ortogonální na každý sloupec matice $\mathbf{A}$, tedy dál postupujeme takto:

$$\begin{aligned}(\mathbf{y}-\mathbf{A}\boldsymbol{\beta},\boldsymbol{\varphi}_0)&=0\\ &\vdots\\ (\mathbf{y}-\mathbf{A}\boldsymbol{\beta},\boldsymbol{\varphi}_k)&=0.\end{aligned}$$

Postupně tento vztah upravíme pro $j = 0, 1, \ldots, k$

$$\begin{aligned}(\mathbf{y}, \boldsymbol{\varphi}_j) - (\mathbf{A}\boldsymbol{\beta}, \boldsymbol{\varphi}_j) &= 0\\ (\mathbf{A}\boldsymbol{\beta}, \boldsymbol{\varphi}_j) &= (\mathbf{y}, \boldsymbol{\varphi}_j)\\ \left(\sum_{m=0}^{k} \boldsymbol{\varphi}_m \beta_m, \boldsymbol{\varphi}_j\right) &= (\mathbf{y}, \boldsymbol{\varphi}_j)\\ \sum_{m=0}^{k} \beta_m (\boldsymbol{\varphi}_m, \boldsymbol{\varphi}_j) &= (\boldsymbol{\varphi}_j, \mathbf{y}).\end{aligned}$$

Z těchto vztahů můžeme soustavu rovnic přepsat v maticovém tvaru:

$$\begin{pmatrix} (\boldsymbol{\varphi}_0, \boldsymbol{\varphi}_0) & \cdots & (\boldsymbol{\varphi}_0, \boldsymbol{\varphi}_k) \\ \vdots & \ddots & \vdots \\ (\boldsymbol{\varphi}_k, \boldsymbol{\varphi}_0) & \cdots & (\boldsymbol{\varphi}_k, \boldsymbol{\varphi}_k) \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} = \begin{pmatrix} (\boldsymbol{\varphi}_0, \mathbf{y}) \\ (\boldsymbol{\varphi}_1, \mathbf{y}) \\ \vdots \\ (\boldsymbol{\varphi}_k, \mathbf{y}) \end{pmatrix}, \tag{3.2}$$

$$\mathbf{G}\boldsymbol{\beta} = \mathbf{d}. \tag{3.3}$$

Definice 3.1. Maticový zápis (3.2) představuje soustavu normálních rovnic. Determinant matice $\mathbf{G}$ se nazývá Gramův determinant (gramián) příslušný k funkcím $\boldsymbol{\varphi}_j(x)$, $j = 0, \ldots, k$.

Poznámka 3.1. V případě, že je vektorový prostor $V = \mathbb{R}^n$ s klasickým euklidovským skalárním součinem, pak platí:

$$\mathbf{G} = \mathbf{A}^T\mathbf{A}, \quad \mathbf{d} = \mathbf{A}^T\mathbf{y},$$

dostáváme stejný systém jako v kapitole o lineární regresi.

Odvození 5.

$(\boldsymbol{\varphi}_m, \boldsymbol{\varphi}_j) = \boldsymbol{\varphi}_m^T \cdot \boldsymbol{\varphi}_j$:

$$\mathbf{G} = \begin{pmatrix} \boldsymbol{\varphi}_0^T\boldsymbol{\varphi}_0 & \cdots & \boldsymbol{\varphi}_0^T\boldsymbol{\varphi}_k \\ \vdots & \ddots & \vdots \\ \boldsymbol{\varphi}_k^T\boldsymbol{\varphi}_0 & \cdots & \boldsymbol{\varphi}_k^T\boldsymbol{\varphi}_k \end{pmatrix} = \begin{pmatrix} \boldsymbol{\varphi}_0^T \\ \vdots \\ \boldsymbol{\varphi}_k^T \end{pmatrix} \cdot (\boldsymbol{\varphi}_0, \ldots, \boldsymbol{\varphi}_k) = \mathbf{A}^T \cdot \mathbf{A},$$

$(\boldsymbol{\varphi}_j, \mathbf{y}) = \boldsymbol{\varphi}_j^T \cdot \mathbf{y}$:

$$\mathbf{d} = \begin{pmatrix} \boldsymbol{\varphi}_0^T\mathbf{y} \\ \vdots \\ \boldsymbol{\varphi}_k^T\mathbf{y} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\varphi}_0^T \\ \vdots \\ \boldsymbol{\varphi}_k^T \end{pmatrix} \cdot \mathbf{y} = \mathbf{A}^T \cdot \mathbf{y}.$$

Řešení soustavy normálních rovnic $\mathbf{G}\boldsymbol{\beta} = \mathbf{d}$ je $\mathbf{b} = \mathbf{G}^{-1}\mathbf{d}$, protože matice $\mathbf{G}$ je obecně regulární.

V praxi ale může být tato matice špatně podmíněná, a tedy téměř singulární, což způsobuje problémy při výpočtu inverze $\mathbf{G}^{-1}$. Pokud ale místo obecné báze prostoru V volíme bázi ortogonální, získáváme diagonální Gramovu matici:

$$\begin{pmatrix} (\boldsymbol{\varphi}_0, \boldsymbol{\varphi}_0) & 0 & \dots & 0 \\ 0 & (\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_1) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (\boldsymbol{\varphi}_k, \boldsymbol{\varphi}_k) \end{pmatrix}.$$

Pokud tedy odhadujeme polynom ve tvaru:

$$y = P_k(x) = \beta_o + \beta_1 x + \beta_2 x^2 + \ldots + \beta_k x^k,$$

pak místo běžné (kanonické báze) $1, x, x^2, \ldots, x^k$, volíme jako bázi systém ortogonálních polynomů.

Kapitola 4

Aplikace metody nejmenších čtverců

4.1 Řešení úlohy regresní analýzy pomocí matic

1. Metodou nejmenších čtverců nalezneme polynom prvního stupně, který aproximuje vztah mezi veličinami X a Y daný tabulkou 4.1

i	1	2	3	4
x_i	2	4	6	8
y_i	2	5	10	14

Tabulka 4.1: Tabulka naměřených hodnot

Řešení
Hledáme tedy rovnici vyrovnávací přímky ve tvaru: $y = b_1x + b_0$. Předpokládáme model: $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} = \begin{pmatrix} X_1 & 1 \\ X_2 & 1 \\ X_3 & 1 \\ X_4 & 1 \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_0 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix},$$

$$\mathbf{X} = \begin{pmatrix} 2 & 1 \\ 4 & 1 \\ 6 & 1 \\ 8 & 1 \end{pmatrix}, \mathbf{Y} = \begin{pmatrix} 2 \\ 5 \\ 10 \\ 14 \end{pmatrix}$$

Najdeme řešení $\mathbf{b} = \begin{pmatrix} b_1 \\ b_0 \end{pmatrix}$, kde

$$\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}, \tag{4.1}$$

$$(\mathbf{X}^T\mathbf{X}) = \begin{pmatrix} 2 & 4 & 6 & 8 \\ 1 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \\ 4 & 1 \\ 6 & 1 \\ 8 & 1 \end{pmatrix} = \begin{pmatrix} 120 & 20 \\ 20 & 4 \end{pmatrix},$$

$$(\mathbf{X}^T\mathbf{X})^-1 = \begin{pmatrix} 0,0083 & 0,05 \\ 0,05 & 0,25 \end{pmatrix},$$

$$\mathbf{X}^T\mathbf{Y} = \begin{pmatrix} 2 & 4 & 6 & 8 \\ 1 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 5 \\ 10 \\ 14 \end{pmatrix} = \begin{pmatrix} 198 \\ 31 \end{pmatrix},$$

$$\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y} = \begin{pmatrix} 0,0083 & 0,05 \\ 0,05 & 0,25 \end{pmatrix} \cdot \begin{pmatrix} 198 \\ 31 \end{pmatrix} = \begin{pmatrix} 3,1934 \\ 17,65 \end{pmatrix}.$$

Tudíž dostáváme řešení, určené zaokrouhleným odhadem:

$$\mathbf{b} = \begin{pmatrix} b_1 \\ b_0 \end{pmatrix} = \begin{pmatrix} 3,19 \\ 17,65 \end{pmatrix}.$$

Takto určenou přímku zapisujeme ve tvaru $y = 3,19x + 17,65$.

4.2 Řešení současných ekonomických úloh

1. Určete vyrovnávací funkci popisující dlouhodobý vývoj cen akcií ČEZu na RM-Systému. Údaje podle oficiálních stránek RM-Systému, české burzy cenných papírů [16] ve dnech, kdy se akcie obchodovaly. Tyto údaje jsou zaznamenány v tabulce 4.4. Rozhodněte, jestli se na toto období hodí aproximace přímkou, nebo kvadratickou funkcí.

   Řešení

   (a) Hledání vyrovnávací přímky: $y = b_0 + b_1x$. Najdeme řešení $\mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \end{pmatrix}$, kde $\mathbf{b}$ splňuje rovnici (4.1). Maticové výpočty provádíme pro zjednodušení v programu Octave.

Datum	Cena za jednu akcii v Kč
31. 12. 2015	449,00
4. 1. 2016	431,70
5. 1. 2016	426,80
6. 1. 2016	422,60
7. 1. 2016	410,00
8. 1. 2016	415,30
11. 1 .2016	414,70
12. 1. 2016	409,80
13. 1. 2016	410,90
14. 1. 2016	406,90
15. 1. 2016	392,70
18. 1. 2016	385,90
19. 1. 2016	386,10
21. 1. 2016	375,00
22. 1. 2016	385,90
25. 1. 2016	383,00
26. 1. 2016	382,40
27. 1. 2016	395,00
28. 1. 2016	407,50
29. 1. 2016	416,00
1. 2. 2016	402,60
2. 2. 2016	398,00
3. 2. 2016	390,10
4. 2. 2016	400,00
5. 2. 2016	397,10
8. 2. 2016	386,20
9. 2. 2016	378,60
10. 2. 2016	377,00
11. 2. 2016	375,00
12. 2. 2016	372,00
15. 2. 2016	380,20
17. 2. 2016	372,60
18. 2. 2016	377,00
19. 2. 2016	375,00
22. 2. 2016	378,90
23. 2. 2016	376,00
24. 2. 2016	367,00
25. 2. 2016	366,00
26. 2. 2016	368,60
29. 2. 2016	368,70
1. 3. 2016	365,00

Tabulka 4.2: Tabulka vývoje akcií ČEZu

```
octave:> Y=[449;431.10;426.80; 422.60; 410; 415.30;
414.70; 409.80; 410.90; 406.90; 392.70; 385.90; 386.10;
375; 385.90; 383; 382.40; 395; 407.50; 416; 402.60; 398;
390.10; 400; 397.10; 386.20; 378.60; 377; 375; 372;
380.20; 372.60; 377; 375; 378.90; 376; 367; 366; 368.60;
368.70; 365]

octave:> X= (1:41)'

octave:> X=[ones(41,1) X]

octave:> b=((X' * X)^-1)*(X' * Y)
b =
422.2328
-1.4325
```

Tudíž dostáváme řešení:

$$\mathbf{b} = \begin{pmatrix} 422,2328 \\ -1,4325 \end{pmatrix}.$$

Předpis vyrovnávací přímky je tedy: $y = 422,2328 - 1,4325x$.
Reziduální součet čtverců:

```
octave:> Y'*Y - Y' * X * b
ans =  4768.4
```

rozhoduje o vhodnosti použitého modelu, značíme $Q(\mathbf{e}) = 4768,4$. Tuto hodnotu dostáváme pro případ, kdy popisujeme závislost přímkou.

(b) Hledání kvadratické vyrovnávací funkce $y = b_2x^2 + b_1x + b_0$.

Najdeme řešení $\mathbf{b} = \begin{pmatrix} b_2 \\ b_1 \\ b_0 \end{pmatrix}$, kde $\mathbf{b}$ splňuje rovnici (4.1). Maticové výpočty provádíme pro zjednodušení v programu Octave.

```
octave:> X= (1:41)'

octave:> X=[ones(41,1) X X.^2]

octave:> b=((X' * X)^-1)*(X' * Y)
b =
2.5750e-02
-2.5140e+00
4.2998e+02
```

Dostáváme řešení:

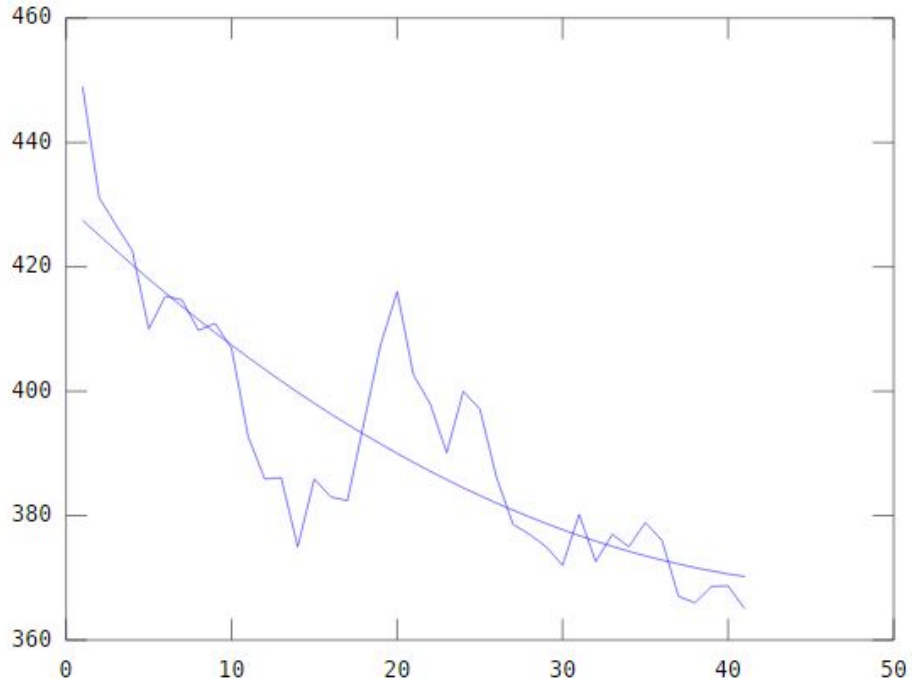
$$\mathbf{b} = \begin{pmatrix} 0,02575 \\ -2,5140 \\ 429,98 \end{pmatrix}.$$

Reziduální součet čtverců:

```
octave:> Y'*Y - Y' * X * b
ans =  4342.9,
```

tedy $Q(\mathbf{e}) = 4342,9$ pro případ, kdy aproximujeme výchozí data z tabulky 4.4 kvadratickou funkcí.

K posouzení vhodnosti modelu vyrovnávací funkce používáme kritéria metody nejmenších čtverců, která stanovují kvalitu modelu podle velikosti reziduálního součtu čtverců $Q(\mathbf{e})$. Tudíž vhodnějším modelem je v naší situaci odhad, že vyrovnávací křivkou je kvadratická funkce. Údaje o akciích jsou velmi nestabilní, mění se velmi často, proto nemůžeme dobře odhadovat budoucí vývoj. Přesvědčíme se o našem závěru z obrázku měření vývoje cen akcií.



Obrázek 4.1: Aproximace vývoje ceny akcií ČEZu za delší období

2. Určete křivku, která by vyrovnala růst (pokles) indexu spotřebitelských cen (životních nákladů). Data čerpáme z oficiálních stránek Českého statistického úřadu [17]. Cenové indexy poměřují úroveň cen vybraného spotřebního koše reprezentativních výrobků a služeb. Do spotřebního koše je zařazeno potravinářské zboží, nepotravinářské zboží (odívání, nábytek, potřeby pro domácnost, atd.) a služby (zdravotnictví, sociální péče, vzdělávání, stravování, ubytování, atd.). Tyto ukazatele jsou důležité pro výpočet inflace, která vychází z měření čistých cenových změn pomocí indexů spotřebitelských cen.

V letech	Úhrn indexu spotřeb. cen
1994	59,1
1995	64,5
1996	70,2
1997	76,2
1998	84,4
1999	86,2
2000	89,4
2001	93,6
2002	95,4
2003	95,5
2004	98,1
2005	100,0
2006	102,5
2007	105,4
2008	112,1
2009	113,3
2010	114,9
2011	117,1
2012	121,0
2013	122,7
2014	123,2
2015	123,6

Tabulka 4.3: Tabulka vývoje indexu spotřeb. cen v letech 1994 až 2015

Řešení

Hledání vyrovnávací přímky: $y = b_0 + b_1 x$. Najdeme řešení $\mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \end{pmatrix}$. Maticové výpočty provádíme pro zjednodušení v programu Octave.

```
octave:> X= (1:22)'

octave:> X=[ones(22,1) X]

octave:> Y = [59.1; 64.5; 70.2; 76.2; 84.4; 86.2; 89.4;
93.6; 95.4; 95.5; 98.1; 100.0; 102.5; 105.4; 112.1;
113.3; 114.9; 117.1; 121.0; 122.7; 123.2; 123.6]

octave:>  b=((X' * X)^-1)*(X' * Y)
b =
64.8455
2.9320
```

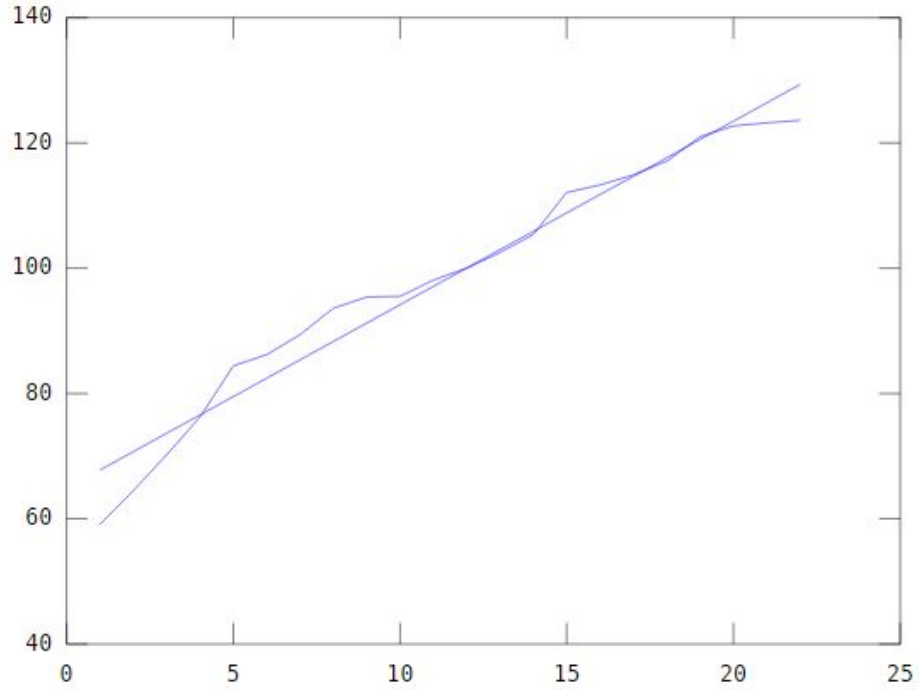
Dostáváme řešení:

$$\mathbf{b} = \begin{pmatrix} 64,8455 \\ 2,9320 \end{pmatrix}.$$

Předpis vyrovnávací přímky je tedy: $y = 64,8455 + 2,9320x$.
Reziduální součet čtverců:

```
octave:> Y'*Y - Y' * X * b
ans =  286.52
```

rozhoduje o vhodnosti použitého modelu, tedy $Q(\mathbf{e}) = 286,52$. Díky této hodnotě můžeme předpokládat, že budoucí vývoj spotřebitelských cen lze popsat pomocí rostoucí vyrovnávací přímky.



Obrázek 4.2: Aproximace vývoje indexu spotřebitelských cen

3. Na příkladu 2 z kapitoly „Jednoduchá regrese“ si nyní ukážeme, jak lze zadané hodnoty $(x_1, y_1), \ldots, (x_n, y_n)$ aproximovat polynomem druhého stupně za použití vědomostí z předchozí kapitoly o ortogonálních polynomech:

$$(P_j(x), P_k(x)) = \sum_{i=1}^{n} \omega(x_i) P_j(x_i) P_k(x_i).$$

K sestavení aproximační funkce potřebujeme vytvořit systém ortogonálních polynomů

ve tvaru:

$$P_{j+1}(x) = (x - \beta_j)P_j(x) - \gamma_j P_{j-1}(x),$$

$$\beta_j = \frac{(xP_j(x), P_j(x))}{(P_j(x), P_j(x))},$$

$$\gamma_i = \frac{(P_j(x), P_j(x))}{(P_{j-1}(x), P_{j-1}(x))}.$$

Při zadaných:

$$P_{-1}(x) \equiv 0, \quad P_0(x) = 1.$$

Hledaný předpis aproximující funkce je řešením obecného vztahu:

$$Y = \sum_{j=0}^{k} \varphi_j(x) \cdot \beta_j. \tag{4.2}$$

Řešení
Z tabulky naměřených hodnot sestavíme postupně ortogonální polynom prvního stupně P_1 a následně ortogonální polynom druhého stupně P_2.

x_i	1	2	3	4	5
y_i	0	3	4	2	2,5

Tabulka 4.4: Tabulka naměřených hodnot

Volíme váhovou funkci:

$$\omega(x) \equiv 1,$$

$P_{-1}(x)$ a $P_0(x)$ je dané, budeme určovat $P_1(x)$.

- Konstrukce P_1:

$$P_1(x) = (x - \beta_0)P_0(x) - \gamma_0 P_{-1}(x) = (x - \beta_0)P_0(x),$$

$$\beta_0 = \frac{(xP_0(x), P_0(x))}{(P_0(x), P_0(x))} = \frac{\sum_{i=1}^{n} \omega(x_i) \cdot x_i P_0(x_i) \cdot P_0(x_i)}{\sum_{i=1}^{n} \omega(x_i) \cdot P_0(x_i) \cdot P_0(x_i)} =$$

$$= \frac{1 \cdot 1 \cdot 1 \cdot 1 + 1 \cdot 2 \cdot 1 \cdot 1 + 1 \cdot 3 \cdot 1 \cdot 1 + 1 \cdot 4 \cdot 1 \cdot 1 + 1 \cdot 5 \cdot 1 \cdot 1}{1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1} =$$

$$= \frac{15}{5} = 3,$$

$$P_1(x) = x - 3.$$

- Konstrukce P_2:

$$P_2(x) =(x - \beta_1)P_1(x) - \gamma_1 P_0(x).$$

Dopočítáme koeficienty β_1, γ_1:

$$\beta_1 = \frac{(xP_1(x)\,, P_1(x))}{(P_1(x)\,, P_1(x))} = \frac{\sum_{i=1}^{n} \omega(x_i) \cdot x_i P_1(x_i) \cdot P_1(x_i)}{\sum_{i=1}^{n} \omega(x_i) \cdot P_1(x_i) \cdot P_1(x_i)} =$$

$$= \frac{1 \cdot (-2) \cdot (-2) + 1 \cdot (-2) \cdot (-1) + 1 \cdot 0 \cdot 0 + 1 \cdot 4 \cdot 1 + 1 \cdot 2 \cdot 10}{1 \cdot (-2) \cdot (-2) + 1 \cdot (-1)(-1) + 1 \cdot 0 \cdot 0 + 1 \cdot 1 \cdot 1 + 1 \cdot 2 \cdot 2} =$$

$$= \frac{4 + 2 + 4 + 20}{4 + 1 + 1 + 4} = \frac{30}{10} = 3,$$

$$\gamma_1 = \frac{(P_1(x)\,, P_1(x))}{(P_0(x)\,, P_0(x))} = \frac{10}{5} = 2.$$

Nyní můžeme dopočítat P_2 :

$$P_2(x) = (x - \beta_1)P_1(x) - \gamma_1 P_0(x) = (x - 3) \cdot (x - 3) - 2 \cdot 1 =$$
$$= x^2 - 6x + 9 - 2 = x^2 - 6x + 7.$$

Dalším krokem je sestavení soustavy normálních rovnic a nalezení jejího řešení $\mathbf{b} = \mathbf{G}^{-1}\mathbf{d}$. Abychom mohli určit Gramovu matici:

$$\mathbf{G} = \begin{pmatrix} (P_0(x)\,, P_0(x)) & 0 & 0 \\ 0 & (P_1(x)\,, P_1(x)) & 0 \\ 0 & 0 & (P_2(x)\,, P_2(x)) \end{pmatrix},$$

musíme ještě dopočítat $(P_2(x),\, P_2(x))$:

$$(P_2(x)\,, P_2(x)) = \sum_{i=1}^{n} \omega(x_i) \cdot P_2(x_i) \cdot P_2(x_i) =$$
$$= 1 \cdot 2^2 + 1 \cdot (-1)^2 + 1 \cdot (-2)^2 + 1 \cdot (-1)^2 + 1 \cdot 2^2 = 14.$$

Diagonální Gramova matice je tedy ve tvaru:

$$\mathbf{G} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 14 \end{pmatrix}.$$

Dále potřebujeme určit vektor $\mathbf{b} = \mathbf{G}^{-1}\mathbf{d}$:

$$\mathbf{G}^{-1} = \begin{pmatrix} \frac{1}{5} & 0 & 0 \\ 0 & \frac{1}{10} & 0 \\ 0 & 0 & \frac{1}{14} \end{pmatrix}, \quad \mathbf{d} = \begin{pmatrix} (P_0(x)\,, \mathbf{y}) \\ (P_1(x)\,, \mathbf{y}) \\ (P_2(x)\,, \mathbf{y}) \end{pmatrix}.$$

Pro jednotlivé složky dostáváme:

$$\begin{aligned}(P_0(x), \mathbf{y}) = \sum_{i=1}^{n} \omega(x_i) \cdot P_0(x_i) \cdot y_i &= 1 \cdot 1 \cdot 0 + 1 \cdot 1 \cdot 3 + 1 \cdot 1 \cdot 4 + 1 \cdot 1 \cdot 2 + \\ &\quad + 1 \cdot 1 \cdot 2,5 = 3 + 4 + 2 + 2,5 = 11,5 \\ (P_1(x), \mathbf{y}) = \sum_{i=1}^{n} \omega(x_i) \cdot P_1(x_i) \cdot y_i &= 1 \cdot (-2) \cdot 0 + 1 \cdot (-1) \cdot 3 + 1 \cdot 0 \cdot 4 + 1 \cdot 1 \cdot 2 + \\ &\quad + 1 \cdot 2 \cdot 2,5 = 0 - 3 + 0 + 2 + 5 = 4 \\ (P_2(x), \mathbf{y}) = \sum_{i=1}^{n} \omega(x_i) \cdot P_2(x_i) \cdot y_i &= 1 \cdot 2 \cdot 0 + 1 \cdot (-1) \cdot 3 + 1 \cdot (-2) \cdot 4 + \\ &\quad + 1 \cdot (-1) \cdot 2 + 1 \cdot 2 \cdot 2,5 = -3 - 8 - 2 + 5 = -8\end{aligned}$$

a výsledný vektor $\mathbf{d}$ je tedy:

$$\mathbf{d} = \begin{pmatrix} 11,5 \\ 4 \\ -8 \end{pmatrix}.$$

Řešení soustavy normálních rovnic je tedy:

$$\mathbf{b} = \begin{pmatrix} \frac{1}{5} & 0 & 0 \\ 0 & \frac{1}{10} & 0 \\ 0 & 0 & \frac{1}{14} \end{pmatrix} \cdot \begin{pmatrix} 11,5 \\ 4 \\ -8 \end{pmatrix} = \begin{pmatrix} 2,3 \\ 0,4 \\ -0,57143 \end{pmatrix}.$$

Nyní dosadíme do aproximační funkce, která je určená vztahem (3):

$$\begin{aligned}y = 2,3 \cdot P_0(x) + 0,4 \cdot P_1(x) + (-0,57143) \cdot P_2(x) &= \\ = 2,3 \cdot 1 + 0,4 \cdot (x - 3) - 0,57143 \cdot (x^2 - 6x + 7) &= \\ = -0,57143x^2 + 3,82857x - 2,9.&\end{aligned}$$

Dostáváme předpis $y = -0,57143x^2 + 3,82857x - 2,9$, který je shodný s odhadem v příkladu 2 z kapitoly „Jednoduchá regrese“. Tentokrát jsme tento předpis určili pomocí ortogonálních polynomů v souladu s obecnou *metodou nejmenších čtverců*.

Závěr

Ve vypracované bakalářské práci jsme se zabývali aproximací dat metodou nejmenších čtverců. Ukázali jsme použití metody nejmenších čtverců v regresní analýze a dále jsme ukázali, jak tato metoda využívá ortogonálních polynomů k sestavení odhadu aproximující funkce. Věnovali jsme se řešitelnosti soustav ortogonálních rovnic a dokázali jsme jejich řešení pomocí této metody. V poslední kapitole této bakalářské práce jsme ukázali aplikace metody nejmenších čtverců na vybraných příkladech ze současné ekonomické situace. Řešení těchto příkladů jsme podrobně popsali a doložili jsme je odpovídajícími grafy. Nalezli bychom i další obory pro aplikace této metody, ale takové úvahy by převyšovaly rámec této bakalářské práce.

Literatura

[1] RALSTON Anthony. *Základy numerické matematiky.* 2. vydání, Praha: vydavatelství Academia, 635 s., 1978.

[2] HAŠEK. Učební texty: Skalární součin, lineární obal množiny. `http://home.pf.jcu.cz/~hasek/LAG/Pr6/Pr_6_SkalarniSoucin.pdfy,http://home.pf.jcu.cz/~hasek/LAG/Pr2/LinearniObal_Podprostor.pdf`. [online], [cit. 2016-04-05].

[3] ZELINKA Jiří HOROVÁ Ivana. *Numerické metody.* 2. vydání, Brno: Masarykova univerzita v Brně, 294 s., 2004. ISBN 80-210-3317.

[4] HASÍK Karel. Numerické metody. `http://www.slu.cz/math/cz/knihovna/ucebni-texty/Numericke-metody/Numericke-metody.pdf`. [online], [cit. 2016-03-26].

[5] MARVAN Michal. Učební text: Vektorové prostory. `http://www.slu.cz/math/cz/knihovna/docs/algebra1/9.-vektorove-prostory`. [online], [cit. 2016-04-05].

[6] PŘIKRYL Petr. *Numerické metody matematické analýzy.* 1. vydání, Praha: SNTL: Nakladatelství technické literatury, 192 s., 1985. ISBN 04-013-85.

[7] PŘIKRYL Petr. *Numerické metody: Aproximace funkcí a matematická analýza.* 1. vydání, Plzeň: vydavatelství ZČU, 187 s., 1996. ISBN 55-067-96.

[8] SLAVÍK. Učební text: Euklidovské prostory. `http://alex.izona.net/vsczu/Treti%20semestr/Matematika%203/Skripta%20Slavik%20z%20jeho%20webu/2%20Euklidovske_prostory,_podprostory_a_podmnoziny.pdf`. [online], [cit. 2016-04-05].

[9] HEBÁK Petr a kolektiv. *Vícerozměrné statistické metody 2.* 1. vydání, Praha: nakladatelství INFORMATORIUM, spol. s. r. o., 239 s., 2005. ISBN 80-7333-036-9.

[10] ANDĚL Jiří. *Matematická statistika.* 2. vydání, Praha: SNTL: Nakladatelství technické literatury jako společné vydání s Alfa, vydavatelství technické a ekonomické literatury, 352 s., 1985. ISBN 04-003-85.

[11] ZVÁRA Karel. *Regrese.* 1. vydání, Praha: vydavatelství Matfyzpress, 253 s., 2008. ISBN 978-80-7378-041-8.

[12] ČERMÁK Libor. *Numerické metody II.* 1. vydání, Brno: akademické nakladatelství CERM, s. r. o., 179 s., 2004,. ISBN 80-214-2722-1.

[13] MMR: Ministerstvo pro místní rozvoj. Hypoteční úvěry poskytnuté od počátku činnosti. `http://www.mmr.cz/getmedia/fd4e7d75-946a-4983-9d35-18eaa6853b8a/Hypotecni-uvery-poskytnute-od-pocatku-cinnosti-hypotecnich-bank.pdf`. [online], [cit. 2016-03-10].

[14] HAVRDA Jan a kolektiv. *Numerické metody a matematická statistika.* 1. vydání, Praha: Vydavatelství ČVUT, 307 s., 1980.

[15] HLAVIČKA Rudolf RŮŽIČKOVÁ Irena. Numerické metody. `http://physics.ujep.cz/~jskvor/NME/DalsiSkripta/Numerika.pdf`. [online], [cit. 2016-03-10].

[16] RM-SYSTÉM česká burza cenných papírů. Kurzy akcií. `http://www.rmsystem.cz/`. [online], [cit. 2016-04-02].

[17] ČSÚ: Český statistický úřad. Indexy spotřebitelských cen, inflace, časové řady. `https://www.czso.cz/csu/czso/isc_cr`. [online], [cit. 2016-04-20].