M5120 Lineární statistické modely

2. cvičení

$$ \xdef\mcal#1{\mathcal{#1}} \xdef\scal#1#2{\langle #1, #2 \rangle} \xdef\N{\mathbb N} \xdef\R{\mathbb R} \xdef\Q{\mathbb{Q}} \xdef\Z{\mathbb{Z}} \xdef\D{\mathbb{D}} \xdef\bm#1{\boldsymbol{#1}} \xdef\vv#1{\mathbf{#1}} \xdef\vvp#1{\pmb{#1}} \xdef\floor#1{\lfloor #1 \rfloor} \xdef\ceil#1{\lceil #1 \rceil} \xdef\grad#1{\mathrm{grad} , #1} \xdef\ve{\varepsilon} $$

Lineární model

Obecně má tvar $$ Y_i = \beta_0 + \beta_1 x_{i, 1} + \dots + \beta_k x_{i, k} + \ve_i, $$ kde $i = 1, \dots, n$. Y_i je naše "cílová proměnná" (regresand). Proměnné $x_{i,1}, \dots, x_{i,k}$ jsou kovariáty (regresor, prediktor) a jsou pevně dané. Dále máme regresní koeficienty $\beta_0, \dots, \beta_k$ a ε_i je náhodná proměnná chyby.

Také platí $$ \varepsilon_i \sim^{iid} (0, \sigma^2) $$ $$ E(\varepsilon_i) = 0 $$ $$ var(\varepsilon_i) = \sigma^2 $$ $$ cov(\varepsilon_i) = 0 $$

Celkem máme $$ \begin{pmatrix} Y_1 \ Y_2 \ \vdots \ Y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{1,1} & \dots & x_{1, k} \ 1 & x_{2,1} & \dots & x_{2, k} \ \vdots & \vdots & \ddots & \vdots\ 1 & x_{n,1} & \dots & x_{n, k} \ \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \ \beta_1 \ \vdots \ \beta_k \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \ \varepsilon_2 \ \vdots \ \varepsilon_n \end{pmatrix} \tag{LSM} $$ A vektorově $$ \vv{Y} = \underbrace{\vv{X}}_{\text{matice plánu}} \cdot \vvp{\beta} + \vvp{\varepsilon} $$

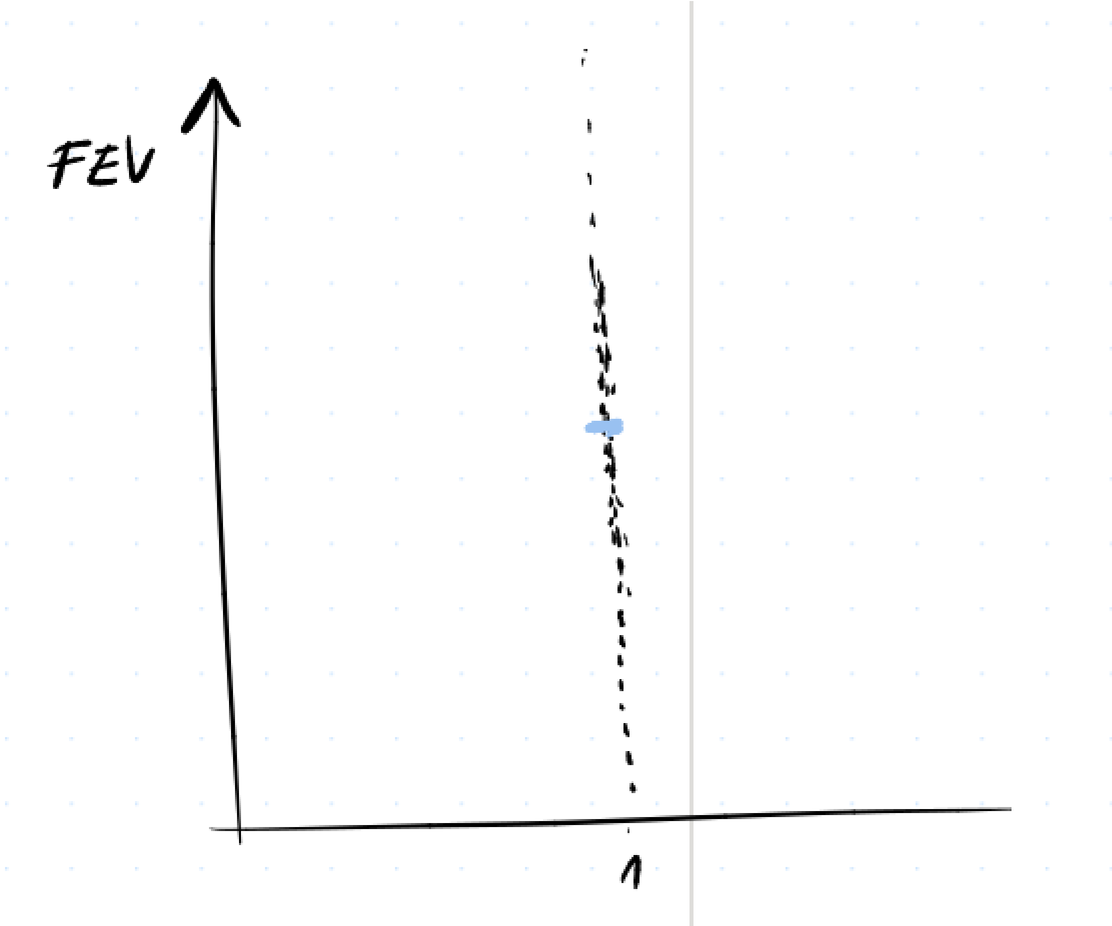
Tedy pro 2. cvičení

02 / a)

$$FEV_i = \beta_0 + \ve_i$$

Můžeme si představit jako funkci $y = \beta_0$

A matice plánu bude $$ \begin{pmatrix} 1 \ 1 \ \vdots \ 1 \end{pmatrix} $$ Tedy kapacita plic je podle tohoto modelu konstantní a vizuálně znázorněné jako

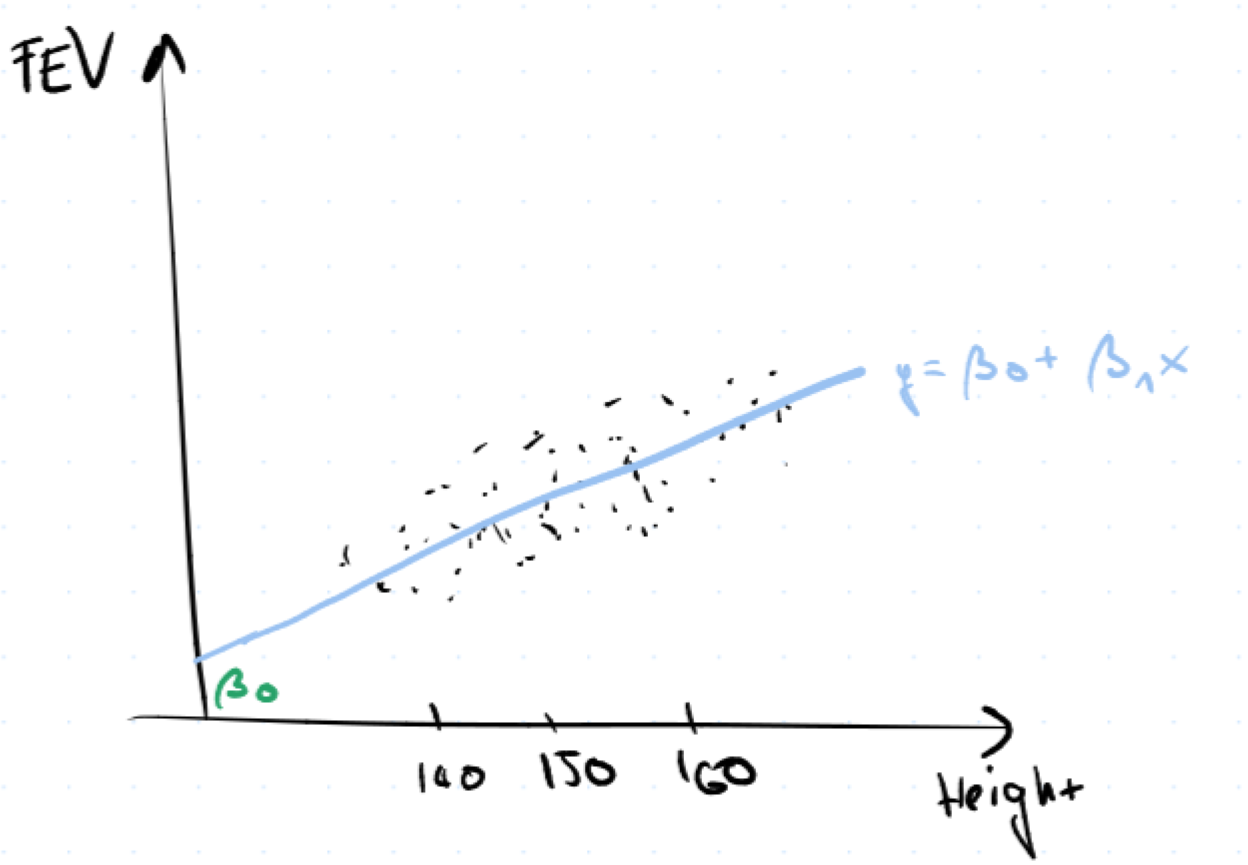


02 / c)

$$ FEV_i = \beta_0 + \beta_1 \cdot \text{Height}_i + \ve_i $$

Kapacitu plic modelujeme pomocí výšky. Tedy v tomto případě chceme funkci $y = \beta_0 + \beta_1 x$.

A graficky



A matice plánu tentokrát bude $$ \begin{pmatrix} 1 & \text{Height}_1 \ 1 & \text{Height}_2 \ \vdots \ 1 & \text{Height}_n \end{pmatrix}, $$ což dosazujeme do (LSM).

Tedy máme model hledáme $y = \beta_0 + \beta_1 x$ a zde

- β_0 ... střední hodnota predikce při nulových hodnotách ostatních prediktorů
- β_1 ... nárůst střední hodnoty predikce při nárůstu výšky (prediktor Height) o 1 cm

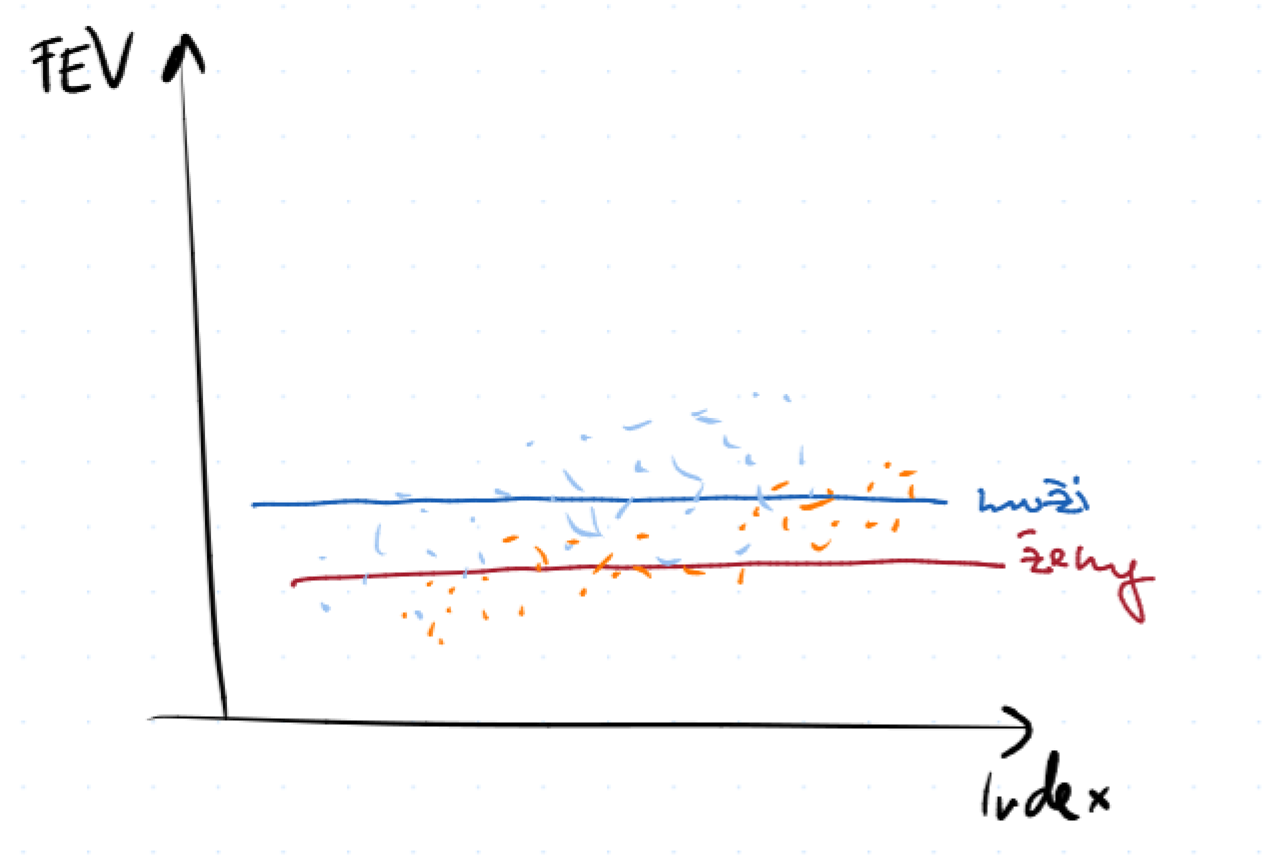
02 / b)

$$ FEV_i = \beta_0 + \beta_1 \cdot \text{Sex}_i + \ve_i $$

a β_1 zde reprezentuje rozdíl predikce mezi muži a ženami s maticí plánu $$ \begin{pmatrix} 1 & 1 \ 1 & 0 \ \vdots \ 1 & 1 \end{pmatrix}, $$

kde 1 reprezentuje muže.

A graficky



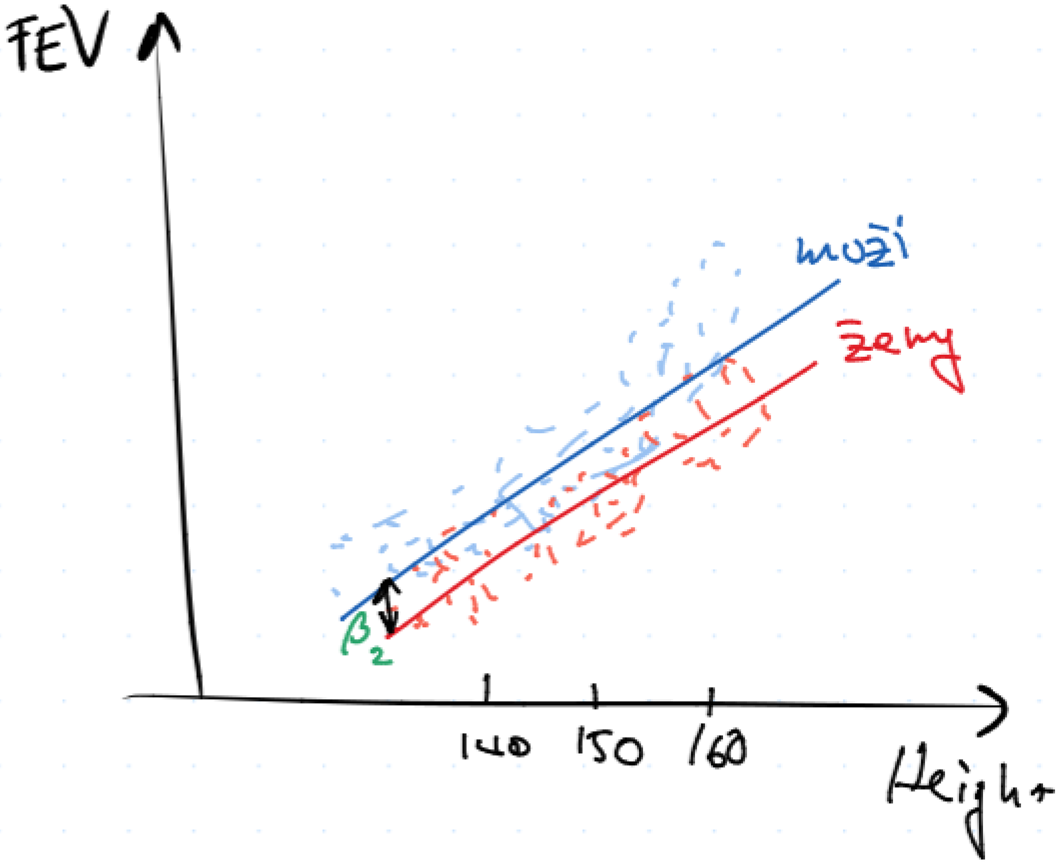
02/ d)

$$ FEV_i = \beta_0 + \beta_1 \cdot \text{Height}_i + \beta_2 \text{Height}_i^2 + \ve_i $$ V tomto případě je β_1, β_2 jsou složité na interpretaci

A matice plánu by v tomto případě byla $$ \begin{pmatrix} 1 & \text{Height}_1 & \text{Height}^2_1 \ 1 & \text{Height}_2 & \text{Height}^2_2 \ \vdots \ 1 & \text{Height}_n & \text{Height}^2_n \end{pmatrix}, $$

02 / e)

$$ FEV_i = \beta_0 + \beta_1 \cdot \text{Height}_i + \beta_2 \text{Sex}_i + \ve_i $$ s maticí plánu $$ \begin{pmatrix} 1 & \text{Height}_1 & 0 \ 1 & \text{Height}_2 & 1 \ \vdots \ 1 & \text{Height}_n & 1 \end{pmatrix}, $$ přičemž ve 3. sloupci jsou 1 značí muže.



A hledáme přímku $$ y = \beta_0 + \beta_1 x + \beta_2 \vv I \set{\text{Sex} = \text{"male"}} $$

s významem koeficientů

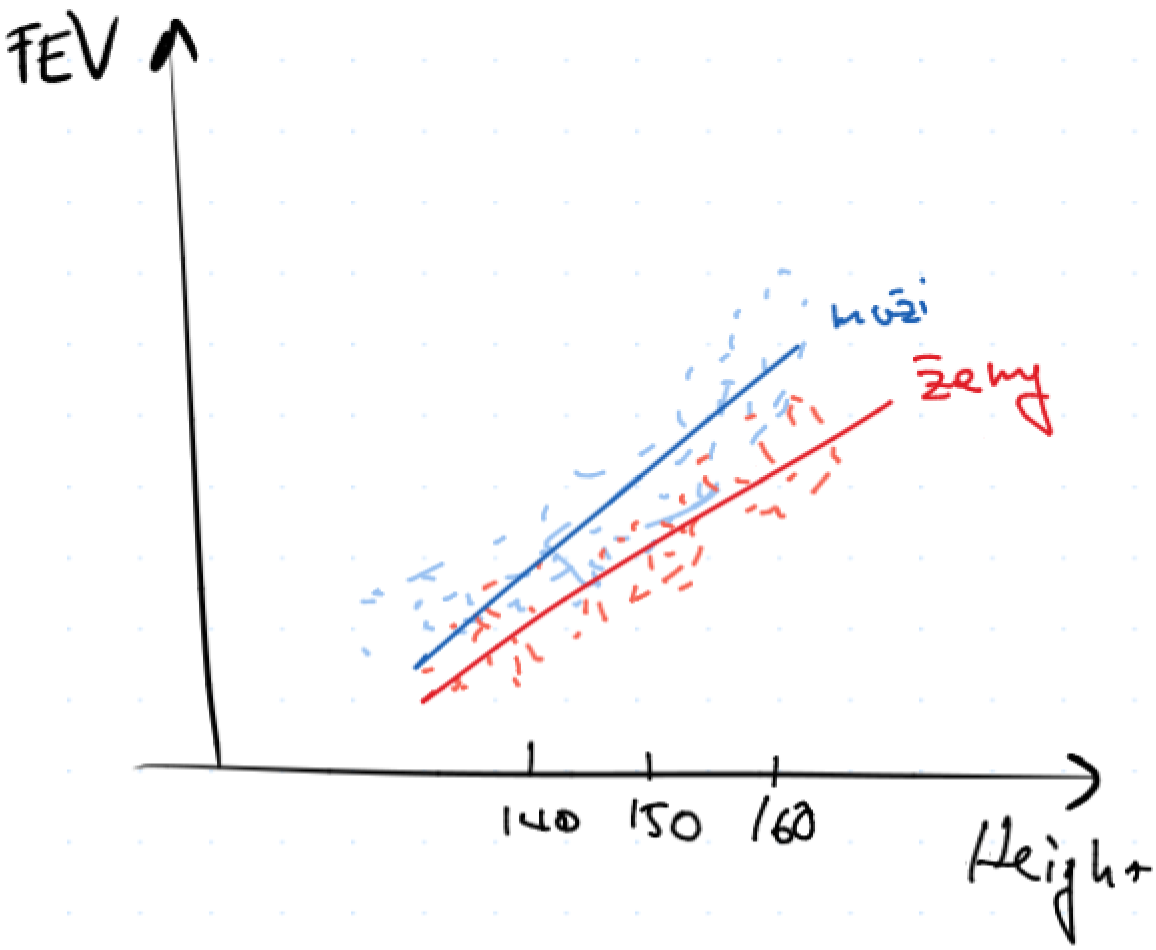
- β_0 ... střední hodnota predikce při nulových hodnotách ostatních prediktorů (nulová výška a ženské pohlaví)
- β_1 ... změna střední hodnoty predikce při nárůstu výšky (prediktor Height) o 1 cm pro ženy
- β_2 ... rozdíl střední hodnoty predikce mezi muži a ženami

02 / g)

$$ FEV_i = \beta_0 + \beta_1 \cdot \text{Height}_i + \beta_2 \text{Sex}_i + \beta_3 (\text{Sex}_i \times \text{Height}_i) + \ve_i, $$ kde členu $\text{Sex}_i \times \text{Height}_i$ **interakce** a matice plánu bude $$ \begin{pmatrix} 1 & \text{Height}_1 & 0 & 0 \ 1 & \text{Height}_2 & 1 & \text{Height}_2 \ 1 & \text{Height}_3 & 0 & 0 \ \vdots & \vdots & \vdots & \vdots \ 1 & \text{Height}_n & 1 & \text{Height}_n \end{pmatrix}, $$ přičemž ve 3. sloupci jsou 1 značí muže.

A hledáme přímku $$ y = \beta_0 + \beta_1 x + \beta_2 \vv I \set{\text{Sex} = \text{"male"}} + \beta_3 x \vv I \set{\text{Sex} = \text{"male"}}, $$ tj.

- žena ... $y = \beta_0 + \beta_1 x$
- muž ... $y = (\beta_0 + \beta_2) + (\beta_1 + \beta_3)x$



Zde β_3 značí rozdíl střední hodnoty predikce mezi muži a ženami při nárůstu výšky o 1 cm

> rozdíl rychlosti růstu FEV mezi muži a ženami

Interpretation

- "keeping the values of all the other covariates fixed, a unit increase in x_i is associated with a $\hat{\beta}_i$ increase in $\operatorname{E} Y$"
  - suitably adapted for categorical predictors and potentially interactions, and depends on the choice of the identifiability conditions
  - polynomials need a more complex interpretation
- is it meaningful to imagine that a covariate changes while all the other remain fixed?

3. cvičení

$$ \xdef\mcal#1{\mathcal{#1}} \xdef\scal#1#2{\langle #1, #2 \rangle} \xdef\N{\mathbb N} \xdef\R{\mathbb R} \xdef\Q{\mathbb{Q}} \xdef\Z{\mathbb{Z}} \xdef\D{\mathbb{D}} \xdef\bm#1{\boldsymbol{#1}} \xdef\vv#1{\mathbf{#1}} \xdef\vvp#1{\pmb{#1}} \xdef\floor#1{\lfloor #1 \rfloor} \xdef\ceil#1{\lceil #1 \rceil} \xdef\grad#1{\mathrm{grad} , #1} \xdef\ve{\varepsilon} $$

Příklad 1

Máme symetrickou, pozitivně definitní matici $m \times m$ značenou Σ.

a)

Poz. def $\iff$ všechna vlastní čísla jsou kladná $\implies$ $\det(\lambda_1 \dots \lambda_m) > 0$ $\implies$ inverze bude existovat $\implies h(\Sigma) = m$

b)

Matice Σ je *samoadjungovaný operátor* Inverzi sestrojíme pomocí spektrálního rozkladu. $$ \Sigma = U \Lambda U^T, $$ kde U je tvořená vlastními vektory a je ortogonální ($U \cdot U^T = I$) a Λ je diagonální matice vlastních čísel.

Pak inverze je $$ \Sigma^{-1} = U \Lambda^{-1} U^T $$

A jako ověření $$ \Sigma \Sigma^{-1} = U \Lambda U^T U \Lambda^{-1} U^T = I $$

c)

Ze spektrálního rozkladu je jistě matice Σ^{-1} pozitivně definitní. Nechť $\vv x \neq 0$ libovolné, pak $$ \vv x^T \Sigma^{-1} \vv x = \vv x^T U \Lambda^{-1} U^T \vv x = \vv x^T U \Lambda^{-\frac 1 2} \Lambda^{-\frac 1 2} U^T \vv x, $$ kde $\Lambda^{-\frac 1 2}$ je matice s převrácenými hodnotami odmocnin vlastních čísel matice Σ na diagonále a tedy $$ \underbrace{\vv x^T U \Lambda^{-\frac 1 2}}*{\vv y} \underbrace{\Lambda^{-\frac 1 2} U^T \vv x}*{\vv y^T} = \vert \vert \vv y \vert \vert^2 > 0 $$

d)

Mějme množinu S_c takovou, že $$ S_c = {\vv x; ; (\vv x - \vv \mu)^T \Sigma^{-1} (\vv x- \vv \mu) = c }, $$ kde $\vv \mu \in \R^m$ a $c \in \R$.

Pro $c < 0$ je $S_c \equiv \emptyset$.

Dále pro $c = 0$ je řešením pouze $S_c = {\vv \mu}$.

Nakonce pro $c > 0$ je $$ \underbrace{(\vv x - \vv \mu)^T U}_{\vv y} \Lambda^{-1} U^T (\vv x - \vv \mu) = \vv y^T \Lambda^{-1} \vv y $$ A pro $m = 2$ tedy $$ (y_1 ; y_2) \begin{pmatrix} \frac 1 {\lambda_1} & 0 \ 0 & \frac 1 {\lambda_2} \end{pmatrix} \begin{pmatrix} y_1 \ y_2 \end{pmatrix} = c $$ $$ \frac {y_1^2} {\lambda_1} + \frac{y_2^2} {\lambda_2} = c, $$ což je **rovnice elipsy** se středem (μ_1, μ_2) a směry os jsou právě vlastní vektory Σ. Nakonec délky poloos budou $\sqrt{c \lambda_i}$
77921418b8dee14fd3706b04914d2aff.png
Image not found or type unknown

Analogicky pro $m > 2$ dostaneme **elipsoid**.

Příklad 4

Matice Σ je poz. **semi**def matice symetrická matice $m \times m$.

a)

$$ h(\Sigma) = r, $$ kde $0 < r \leq m$.

Nahradíme U za matici $$ U_1 = \begin{pmatrix} \vv u_1 & \vert & \vv u_2 & \vert & \dots & \vert & \vv u_r \end{pmatrix} $$ a také $$ \Lambda_1 = \mathrm{diag} (\lambda_1, \dots, \lambda_r) $$

b)

Sestrojme matici $\tilde{\Sigma}$ takovou, že $\tilde{\Sigma} \tilde{\Sigma}^T = \Sigma$ jako $$ \tilde{\Sigma} = U_1 \Lambda_1^{\frac 1 2}, $$ když je $h(\Sigma) = r$ a pro $h(\Sigma) = m$ jako $\tilde{\Sigma} = U \Lambda^{\frac 1 2}$.

d)

> Zde značím pseudoinverzi matice A jako $A^\dagger$

$\tilde{\Sigma}^\dagger = \Lambda_1^{- \frac 1 2} U_1^T$ a pro tuto matici bychom ukázali všechny 4 vlastnosti pseudoinverze, tj.

- $A A^\dagger A = A$
- $A^\dagger A A^\dagger = A^\dagger$
- $(A^\dagger A)^T = A^\dagger A$
- $(A A^\dagger)^T = A A^\dagger$

4. cvičení

$$ \xdef\mcal#1{\mathcal{#1}} \xdef\scal#1#2{\langle #1, #2 \rangle} \xdef\N{\mathbb N} \xdef\R{\mathbb R} \xdef\Q{\mathbb{Q}} \xdef\Z{\mathbb{Z}} \xdef\D{\mathbb{D}} \xdef\bm#1{\boldsymbol{#1}} \xdef\vv#1{\mathbf{#1}} \xdef\vvp#1{\pmb{#1}} \xdef\floor#1{\lfloor #1 \rfloor} \xdef\ceil#1{\lceil #1 \rceil} \xdef\grad#1{\mathrm{grad} , #1} \xdef\ve{\varepsilon} \xdef\im#1{\mathrm{im}(#1)} \xdef\tr#1{\mathrm{tr}(#1)} \xdef\norm#1{\left\vert \left\vert #1 \right\vert\right\vert} \xdef\scal#1#2{\langle #1, #2 \rangle} $$

1. příklad

Máme symetrickou idempotentní matice P velikosti $n\times n$, což je matice **ortogonální projekce**.

a)

Jelikož je P symetrická, tak existuje spektrální rozklad tvaru $$ P = U \Lambda U^T $$

> matice U je ortogonální a plné hodnosti

a jelikož je idempotentní $$ P = PP $$ Potom $$ U\Lambda U^T = P = PP = U \Lambda \underbrace{U^T U}_{I} \Lambda U^T = U \Lambda^2 U^T $$ A celkem dostáváme $\Lambda^2 = \Lambda$. Navíc jelikož je Λ diagonální, tak pro všechna vlastní čísla λ matice Λ platí $$ \lambda^2 = \lambda\ (\lambda - 1)\lambda = 0 \implies \lambda = 0 \lor \lambda = 1 $$

> Idempotentní matice P je invertibilní právě tehdy, když $P = I$

b)

Jelikož $P = U\Lambda U^T$, pak U je plné hodnosti (je ortogonální), pak $$ h(P) = h(U\Lambda U^T) = h(\Lambda) = | \set{\lambda_i, \; \lambda_i = 1} | $$

c)

Z části **b)** máme $$ \implies h(P) = \dots = \sum \lambda_i = \tr \Lambda = \tr {\Lambda U U^T} $$ A jelikož pro stopu součinu matic platí $\tr {A B C} = \tr {C A B}$ (invariantnost vůči cyklickým operacím), pak $$ \implies h(P) = \dots = \tr {U \Lambda U^T} = \tr P $$

d)

Je-li vektor $\vv y \in \im P$, pak $P \vv y = \vv y$ Pokud $\vv y \in \im P$, pak $\exists \vv z$, že $\vv y = P \vv z$

> Definice obrazu $\im P$: $$\im P = \set{\vv y \in \R^n \mid \exists \vv z \in \R^n : \; P\vv z = \vv y}$$

Pak $$ P \vv y = P (P \vv z) = (P P) \vv z = P \vv z = \vv y $$

e)

> Chceme ukázat, že projekce do menšího podprostoru je zároveň projekce do onoho většího prostoru

Nechť $\vv z \in \R^n$ je libovolné, projekce z do menšího prostoru je prvek většího prostoru. $$ \tilde P \vv z \in \im {\tilde P} \leq \im P $$ Pak podle **d)** platí $$ P(\tilde P \vv z) = \tilde P \vv z\ P(\tilde P \vv z) - \tilde P \vv z = \vv 0 \ (P\tilde P - \tilde P) \vv z = \vv 0, $$ nicméně vektor $\vv z$ byl libovolný. To tedy znamená, že zobrazení $P \tilde P - \tilde P$ pošle všechny $\vv z \in \R^n$ na nulový vektor, tj. $$ \ker (P \tilde P - \tilde P) = \R^n $$ Z lineární algebry víme, že $\dim \ker(A) + \dim \im A = n$ pro A tvaru $n \times n$. A tedy $$ \dim \im {P \tilde P - \tilde P} = 0 \implies P \tilde P - \tilde P = \vv 0 $$ Neboť $P, \tilde P$ jsou ortogonální projekce, tak jsou symetrické. Celkem $$ \tilde P P = \tilde P^T P^T = (\tilde P P)^T = \tilde P^T = \tilde P = P\tilde P $$

f)

Nechť $\vv x \in \R^n$ pevné. Mějme $P \vv z \in \im P$ Vezměme libovolné $\vv y \in \im P$ a spočítáme $\norm {\vv y - \vv x}^2$ Pak $$ \norm {\vv y - \vv x}^2 = \norm{P \vv x - x + y - P \vv x}^2 = \scal {(P \vv x - \vv x) + (\vv y - P \vv x)} {(P \vv x - \vv x) + (\vv y - P \vv x)} $$

> Pro skalární součin platí $$\scal {\vv u} {\vv v} = \scal {\vv v} {\vv u} \ \scal {\vv u + \vv v} {\vv w} = \scal {\vv u} {\vv w} + \scal {\vv v} {\vv w} \ \scal {\vv u} {a\vv v} = a \scal {\vv u} {\vv v}, ; a \in \R $$

Pak dostáváme $$ = \norm{P \vv x - \vv x}^2 + 2 \scal {P \vv x - \vv x} {\vv y - P \vv x} + \norm{\vv y - P \vv x}^2 $$

A zajímá nás hlavně $\scal {P \vv x - \vv x} {\vv y - P \vv x}$, tedy

$$ \scal {P \vv x - \vv x} {\vv y - P \vv x} = (\vv y - P \vv x)^T (P \vv x - \vv x) = \vv y^T (P \vv x - \vv x) - (P \vv x)^T P \vv x - \vv x = $$ $$ = \vv y^T P \vv x - \vv y^T \vv x - \vv x^T \underbrace{P^T P}_ {P} \vv x + \vv x^T P^T \vv x = \vv y^T P \vv x - \vv y^T \vv x $$

Jelikož $\vv y \in \im P$, tak jistě $\exists \vv z \in \R^n$ takové, že $\vv y = P \vv z$. Z toho plyne

$$ \vv y^T P \vv x - \vv y^T \vv x = (P \vv z)^T P \vv x - (P \vv z)^T \vv x = \vv z^T \underbrace{P^T P}_ {P} \vv x - \vv z^T \underbrace{P^T}_ {P} \vv x = 0 $$

Celkem $\norm {\vv y - \vv x}^2$ závisí pouze na $\norm {\vv y - P\vv x}^2$ a $$ \norm{P \vv x - \vv x}^2 + \underbrace{2 \scal {P \vv x - \vv x} {\vv y - P \vv x}}_{0} + \norm{\vv y - P \vv x}^2 \geq \norm{P \vv x - \vv x}^2 $$ a rovnost nastane pouze v případě $P \vv x = \vv y$.

Zadání v `R`

Vykreslit si grafy hustot a distribučních funkcí pro

- $N(0,1)$ a spočítat $P(X \leq 2)$
- Studentovo $t(df = 5)$ a spočítat $P(X \leq 2)$
- $\chi^2(df = 10)$ a spočítat $P(X \leq 20)$
- Fisherovo $F(5, 10)$ a spočítat $P(X \leq 2)$

a vypočtěte 95% kvantil a zaznačte do grafu

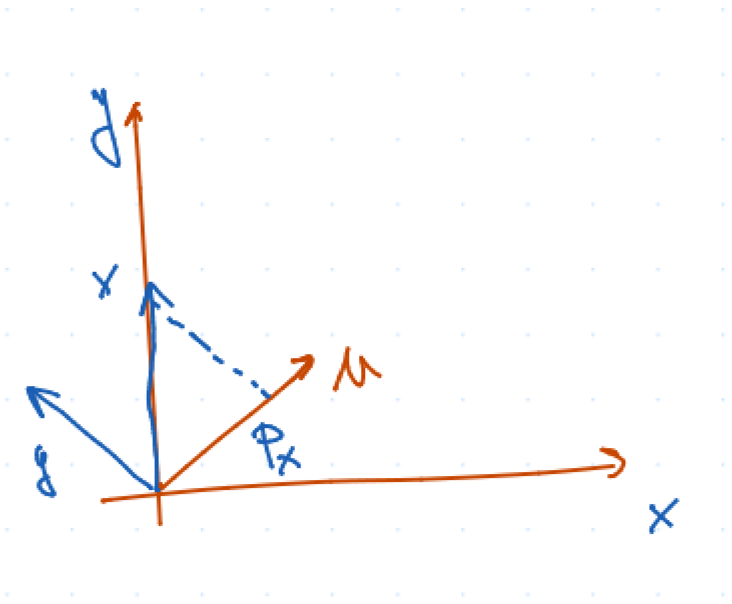
5. cvičení

$$ \xdef\mcal#1{\mathcal{#1}} \xdef\scal#1#2{\langle #1, #2 \rangle} \xdef\N{\mathbb N} \xdef\R{\mathbb R} \xdef\Q{\mathbb{Q}} \xdef\Z{\mathbb{Z}} \xdef\D{\mathbb{D}} \xdef\bm#1{\boldsymbol{#1}} \xdef\vv#1{\mathbf{#1}} \xdef\vvp#1{\pmb{#1}} \xdef\floor#1{\lfloor #1 \rfloor} \xdef\ceil#1{\lceil #1 \rceil} \xdef\grad#1{\mathrm{grad} , #1} \xdef\ve{\varepsilon} \xdef\im#1{\mathrm{im}(#1)} \xdef\tr#1{\mathrm{tr}(#1)} \xdef\norm#1{\left\vert \left\vert #1 \right\vert\right\vert} \xdef\scal#1#2{\langle #1, #2 \rangle} \xdef\ex#1{\mathrm{E} ,\left(#1\right)} \xdef\exv#1{\mathrm{E}, \vv{#1}} $$

4. cvičení / 2. příklad

a) i b)

Mějme vektor $\vv u \in \R^n$ (takový, že $\norm{\vv u} = 1$), pak matice $$ P = \frac {\vv u \vv u^T} {\norm {\vv u}^2} $$ je ortogonální projekce na $\im {\vv u}$.



Pro ortogonální projekci platí

1. $P = P \cdot P$ - idempotence projekce
2. P je symetrická (z ortogonality)

Pak $$ P = \frac {\vv u \vv u^T} {\norm {\vv u}^2} $$ a tedy $$ PP = \frac {\vv u \vv u^T} {\norm {\vv u}^2} \frac {\vv u \vv u^T} {\norm {\vv u}^2} $$

Z definice skalárního součiny $\scal {\vv u} {\vv u} = \norm {\vv u}^2$ a proto $$ PP = \frac {\vv u \vv u^T} {\norm{\vv u}^2} = P, $$ což zvláště platí pro $\norm{\vv u} = 1$. Pro nějaké $\vv x \in \R^n$ máme $$ P \vv x = \frac {\vv u \vv u^T \vv x} {\norm {\vv u}^2} = \frac {\vv u \scal {\vv u} {\vv x}} {\norm{\vv u}^2} = \underbrace{\frac {\scal {\vv u} {\vv x}} {\norm{\vv u}^2}}_ {\in \R} \vv u \in \im {\vv u} $$

c)

Mějme ${ \vv u_ 1, \dots, \vv u_ p }$ ortonormální vektory, pak matice $$ P = U U^T, $$ kde $U = (\vv u_ 1 \; \vv u_ 2 \; \dots \; \vv u_ p)$, je ortogonální projekce na $\im U$.

Ukažme $$ P \cdot P = U \underbrace{U^T U}_ {I} U^T = U U^T $$ a $$ P\vv x = U U^T \vv x = U \cdot \begin{pmatrix} \scal {\vv u_ 1} {\vv x} \ \scal {\vv u_ 2} {\vv x} \ \vdots \ \scal {\vv u_ p} {\vv x} \ \end{pmatrix} = \underbrace{\vv u_ 1 \scal {\vv u_ 1} {\vv x}}_ {\in \R} + \dots + \underbrace{\vv u_ p \scal {\vv u_ p} {\vv x}}_ {\in \R} \in \im U, $$ což je lineární kombinace vektorů $\vv u_1, \dots, \vv u_p$ a jistě tedy $P \vv x \in \im U$.

d) i e)

Máme lineárně nezávislé vektory ${\vv a_1, \dots, \vv a_n}$, pak $$ P = A(A^T A)^{-1}A^T $$ je ortogonální projekce.

Jednou možností by bylo použít spektrální rozklad $A = U \Sigma V$, čehož bychom dostali $P = U U^T$, což jsme ukázali v bodě **c)**.

Druhá možnost je $$ PP = A(A^T A)^{-1}\underbrace{A^T A(A^T A)^{-1}}_ {I}A^T = P, $$ což stejně fungovalo i pro pseudoinverzi. Dále $$ A \vv x = \vv y \in \im A, $$ pak $$ P \vv y = (A(A^T A)^{-1}A^T) \vv y = A \underbrace{(A^T A)^{-1}A^T A}_ {I} \vv x = A \vv x = \vv y $$

> Zde jsme jen ukázali něco o $\vv y \in \im A$, nikoliv o obecném $\vv x \in \R^p$

5. cvičení / 1. příklad

Nechť $$ \exv X = \begin{pmatrix} \ex X_1 \ \ex X_2 \ \vdots \ \ex X_n \end{pmatrix} $$ a $DX = Var X = Cov(\vv X, \vv X)$ platí $$ Cov(\vv X, \vv X) = \begin{pmatrix} Cov(X_1, X_1) & Cov(X_1, X_2) & \dots & Cov(X_1, X_n) \ Cov(X_2, X_1) & Cov(X_2, X_2) & \dots & Cov(X_2, X_n) \ \vdots & \vdots & \ddots & \vdots \ Cov(X_n, X_1) & Cov(X_n, X_2) & \dots & Cov(X_n, X_n) \end{pmatrix} $$

a)

Ukažme $$ \ex {A \vv X + \vv b} = A\cdot \exv X + \vv b, $$ což je analogické k jednorozměrnému případu.

Pro i-tý prvek platí $$ \ex {\sum_{k = 1}^n a_{i,k} X_k + b_i} = \sum_{k = 1}^n \ex {X_k} + b_i, $$ což je, co jsme potřebovali.

b)

Ukažme $$ Var(A\vv X + \vv b) = A \cdot Var \vv X \cdot A^T $$ což je opět analogie k $D(a X + b) = a^2 D X$.

Využijeme vlastnost kovariance. Tedy pro (i,j)-tý prvek matice $A \vv X + \vv b$ platí $$ Cov \left(\sum_{k = 1}^n a_{i,k} X_k + b_i, \sum_{k = 1}^n a_{j,k} X_k + b_j \right) = \sum_{k = 1}^n \sum_{l = 1}^n a_{i,k} a_{j,l} Cov (X_k, X_l) = $$ $$ = \sum_{l = 1}^n \left(\sum_{k = 1}^n a_{i,k} Cov(X_k, X_l) \right) a_{j,l} $$

c)

Máme ukázat $$ \ex {\vv X^T \vv X} = \exv X^T \exv X + tr(Var(\vv X)), $$ což je ekvivalentní s $$ \ex{X_1^2 + X_2^2 + \dots + X_n^2} = \ex{X_1}^2 + \dots + \ex{X_n}^2 + DX_1 + \dots + DX_n $$

Obecně platí $$ Var(\vv X) = \ex{\vv X \vv X^T} - \exv X \cdot \exv X^T $$ a tedy $$ \ex {\vv X^T \vv X} = \ex {tr(\vv X^T \vv X)} = $$ pak dle vlastnosti stopy matice $$ = \ex {tr(\vv X \vv X^T)} = tr(\ex{\vv X \vv X^T}) = tr(Var(\vv X) + \exv X \cdot \exv X^T) = $$ a opět dle vlastnosti stopy matice $$ = tr(Var(\vv X)) + tr(\exv X \cdot \exv X^T) = \exv X^T \cdot \exv X + tr(Var(\vv X)) $$

7. cvičení

$$ \xdef\mcal#1{\mathcal{#1}} \xdef\scal#1#2{\langle #1, #2 \rangle} \xdef\N{\mathbb N} \xdef\R{\mathbb R} \xdef\Q{\mathbb{Q}} \xdef\Z{\mathbb{Z}} \xdef\D{\mathbb{D}} \xdef\bm#1{\boldsymbol{#1}} \xdef\vv#1{\mathbf{#1}} \xdef\vvp#1{\pmb{#1}} \xdef\floor#1{\lfloor #1 \rfloor} \xdef\ceil#1{\lceil #1 \rceil} \xdef\grad#1{\mathrm{grad} , #1} \xdef\ve{\varepsilon} \xdef\im#1{\mathrm{im}(#1)} \xdef\tr#1{\mathrm{tr}(#1)} \xdef\norm#1{\left\vert \left\vert #1 \right\vert\right\vert} \xdef\scal#1#2{\langle #1, #2 \rangle} \xdef\ex#1{\mathrm{E} ,\left(#1\right)} \xdef\exv#1{\mathrm{E}, \vv{#1}} $$

Nechť $$ \vv Y $$ jsou data, $$ \hat {\vv Y} = \vv X \hat {\vvp \beta}, \qquad E \hat{\vv Y} = E \vv Y $$ je odhad $\vv Y$ a $$ \vv e $$ je odhad $\vvp \ve$.

A máme **celkovou sumu čtverců** $$ TSS = \sum_{i = 1}^n (Y_i - \overline Y_i)^2 $$ také **vysvětlovanou sumu čtverců** $$ ESS = \sum_{i = 1}^n(\hat Y_i - \overline Y_i)^2 $$ a neposlední řadě **reziduální sumu čtverců** $$ RSS = \sum_{i = 1}^n (Y_i- \hat Y_i)^2 $$

A platí $$ TSS = RSS + ESS $$

a nechť R^2 je **koeficient determinace** $$ R^2 = \frac {ESS} {TSS} \in (0, 1] $$ a **adjustovaný koeficient determinace** $$ R^2_{adj} = 1 - \frac {\cfrac {RSS} {n-p}} {\cfrac {TSS} {n-1}} $$

Dále $$ \hat \sigma^2 = \frac {RSS} {n - p} $$ a $$ var(\hat{\vvp \beta}) = \hat \sigma^2 (\vv X^T \vv X)^{-1} $$ Přičemž $var(\hat{\vvp \beta})$ dostaneme pomocí `vcov(<model>)`

9. cvičení

a) IS pro β_i: $$ T_i = \frac {\hat{\beta_i}} {\sqrt{\hat{\sigma} (\pmb X^T \pmb X)^{-1}_{i,i}}} \sim t(n-p) $$

Pak $$ P\left(T_ i \in \left[t_{\frac \alpha 2}(n-p), t_{1 - \frac \alpha 2}(n-p)\right]\right) = 1 - \alpha $$ $$ t_{\frac \alpha 2}(n-p) \leq T_ i \leq t_ {1 - \frac \alpha 2}(n-p) $$ $$ t_{\frac \alpha 2}(n-p) \leq \frac {\beta_i - \hat{\beta_ i}} {\sqrt{\hat{\sigma}^2 (\pmb X^T \pmb X)^{-1}_ {i,i}}} \leq t_{1 - \frac \alpha 2}(n-p) $$ Z čehož dostaneme $$ \beta_ i \in \left(\hat{\beta_ i} \pm t_{1-\frac \alpha 2}(n-p) \sqrt{\hat{\sigma}^2 (\pmb X^T \pmb X)^{-1}_ {i,i}}\right) $$

b) $$ T = \frac {\pmb a^T \pmb \beta^T} {\sqrt{\hat{\sigma}^2 \pmb a^T (\pmb X^T \pmb X)^{-1} \pmb a}} \sim t(n-p) $$

$$ \left(\pmb a^T \pmb \beta \pm t_{1-\frac \alpha 2}(n-p) \sqrt{\hat{\sigma}^2 \pmb a^T (\pmb X^T \pmb X)^{-1} \pmb a}\right) $$

A regresní přímka pro dívky bude mít tvar $$ y = \hat \beta_0 + \hat \beta_1 x $$ Pro chlapce: $$ y = \hat \beta_0 + \hat \beta_2 + (\hat \beta_1 + \hat \beta_3) x $$ IS pro $\beta_1 + \beta_3$, tedy $\pmb a = (0, 1, 0, 1)$

d) Při počítání predikčního intervalu zohledňujeme chybu u "nového pozorování". Tedy odhad rozptylu je $$ \hat{\sigma}^2 \pmb x^T (\pmb X^T \pmb X)^{-1} \pmb x + \hat \sigma^2 \implies T = \frac {\pmb x^T \pmb \beta^T} {\sqrt{\hat{\sigma}^2(1 + \pmb x^T (\pmb X^T \pmb X)^{-1} \pmb x)}} \sim t(n-p) $$

```
# Confidence interval
predict(..., interval = "confidence")
# Or
predict(..., interval = "prediction")
```

10. cvičení

$$ \xdef\mcal#1{\mathcal{#1}} \xdef\scal#1#2{\langle #1, #2 \rangle} \xdef\N{\mathbb N} \xdef\R{\mathbb R} \xdef\Q{\mathbb{Q}} \xdef\Z{\mathbb{Z}} \xdef\D{\mathbb{D}} \xdef\bm#1{\boldsymbol{#1}} \xdef\vv#1{\mathbf{#1}} \xdef\vvp#1{\pmb{#1}} \xdef\floor#1{\lfloor #1 \rfloor} \xdef\ceil#1{\lceil #1 \rceil} \xdef\grad#1{\mathrm{grad} , #1} \xdef\ve{\varepsilon} \xdef\im#1{\mathrm{im}(#1)} \xdef\tr#1{\mathrm{tr}(#1)} \xdef\norm#1{\left\vert \left\vert #1 \right\vert\right\vert} \xdef\scal#1#2{\langle #1, #2 \rangle} \xdef\ex#1{\mathrm{E} ,\left(#1\right)} \xdef\exv#1{\mathrm{E}, \vv{#1}} \xdef\mtrx#1{\begin{pmatrix}#1\end{pmatrix}} $$

Scheffeho věta $$ P\left([\vv b^T (A\hat\beta - A \beta)]^2 \leq m F_{1 - \alpha}(m, n - p) \hat \sigma^2 \vv b^T A (\vv X^T \vv X)^{-1} A^T \vv b\right) = 1 - \alpha $$ $\forall b \in \R^m$, je-li matice A typu $m \times p$ plné hodnosti.

Příklad $$ Y_i = \beta_0 + \beta_1 \cdot \text{Height}_I + \beta_2 \cdot \text{Sex}_i + \beta_3 \cdot (\text{Height} + \text{Sex})_i + \ve_i, \quad \ve \sim N(0, \sigma^2) $$ a chceme zkonstruovat 95% PS pro chlapce a dívky

1) Napíšeme tvar reg. křivky

- d: $y = \hat \beta_0 + \hat\beta_1 x$
- ch: $y = \hat\beta_0 + \hat\beta_2 + (\hat \beta_1 + \hat \beta_3)x$

2) Zvolíme vhodný tvar $\vv b$ a A:

- d: $\vv b = \mtrx{1 \ x} \in \R^2$, pak $$ \mtrx{1 & x} \overbrace{\mtrx{1 & 0 & 0 & 0 \ 0 & 1 & 0 & 0}}^A \mtrx{\hat \beta_1 \ \hat \beta_2 \ \hat \beta_3 \ \hat \beta_4} $$
- ch: $\vv b = \mtrx{1 \ x}$, pak $$ \mtrx{1 & x} \overbrace{\mtrx{1 & 0 & 1 & 0 \ 0 & 1 & 0 & 1}}^A \mtrx{\hat \beta_1 \ \hat \beta_2 \ \hat \beta_3 \ \hat \beta_4} $$

Nejprve počítejme pro dívky, Označme $$ \vv b^T A = \vv x^T = (1, x, 0, 0) $$ **3)** Odvodíme tvar pásu spolehlivosti (PS) $$ P\left([\vv x^T \hat \beta - \underbrace{\vv x^T \beta}_ {y = \beta_0 + \beta_1 x}]^2 \leq 2 F_ {1 - \alpha}(2, n - 4) \sigma^2 \vv x^T (\vv X^T \vv X)^{-1} \vv x \right) = 1 - \alpha $$ kde y je náhodná proměnná. Upravujme

$$ P\left(|\vv x^T \hat \beta - y| \leq \sqrt{2 F_{1 - \alpha}(2, n - 4) \sigma^2 \vv x^T (\vv X^T \vv X)^{-1} \vv x} \right) = 1 - \alpha $$

- pro $\vv x^T \hat \beta - y > 0$ dostáváme **dolní hranici** $$ P\left(y \geq \vv x^T \hat \beta - \sqrt{2 F_{1 - \alpha}(2, n - 4) \sigma^2 \vv x^T (\vv X^T \vv X)^{-1} \vv x} \right) = 1 - \alpha $$
- nebo pro $\vv x^T \hat \beta - y < 0$ dostáváme **horní hranici** $$ P\left(y \leq \vv x^T \hat \beta + \sqrt{2 F_{1 - \alpha}(2, n - 4) \sigma^2 \vv x^T (\vv X^T \vv X)^{-1} \vv x}

\right) = 1 - \alpha $$