

M5120 Lineární statistické modely

- [2. cvičení](#)
- [3. cvičení](#)
- [4. cvičení](#)
- [5. cvičení](#)
- [7. cvičení](#)
- [9. cvičení](#)
- [10. cvičení](#)

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 β_0, \dots, β_k a \ve_i je náhodná proměnná chyby.

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

Celkem máme
$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{1,1} & \dots & x_{1,k} \\ 1 & x_{2,1} & \dots & x_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n,1} & \dots & x_{n,k} \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \ve_1 \\ \ve_2 \\ \vdots \\ \ve_n \end{pmatrix}$$
$$\tag{LSM} \quad A \text{ vektorově } \vv{Y} = \underbrace{\vv{X}}_{\text{matice plánu}} \cdot \vvp{\beta} + \vvp{\ve}$$

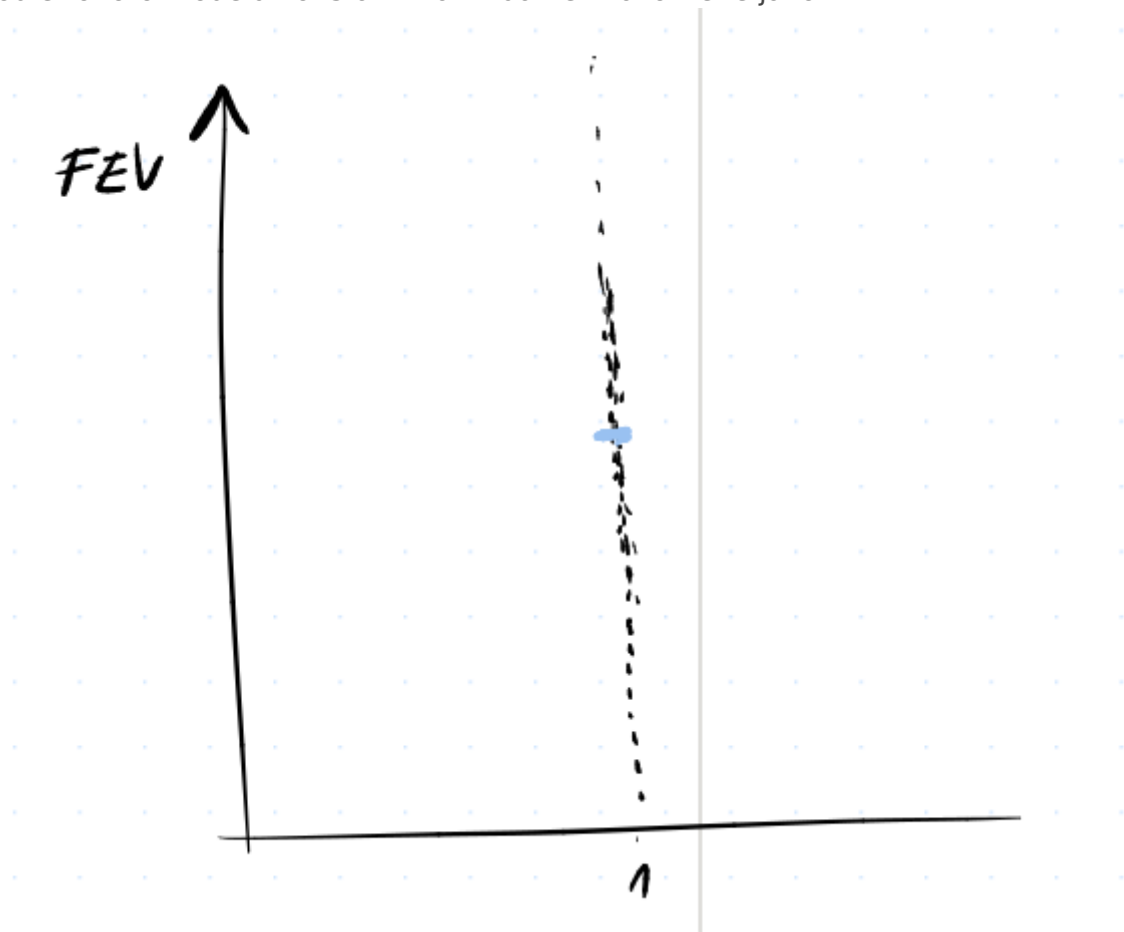
Tedy pro 2. cvičení

02 / a)

$$FEV_i = \beta_0 + \ve_i$$

Můžeme si představit jako funkci $y = \beta_0$

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

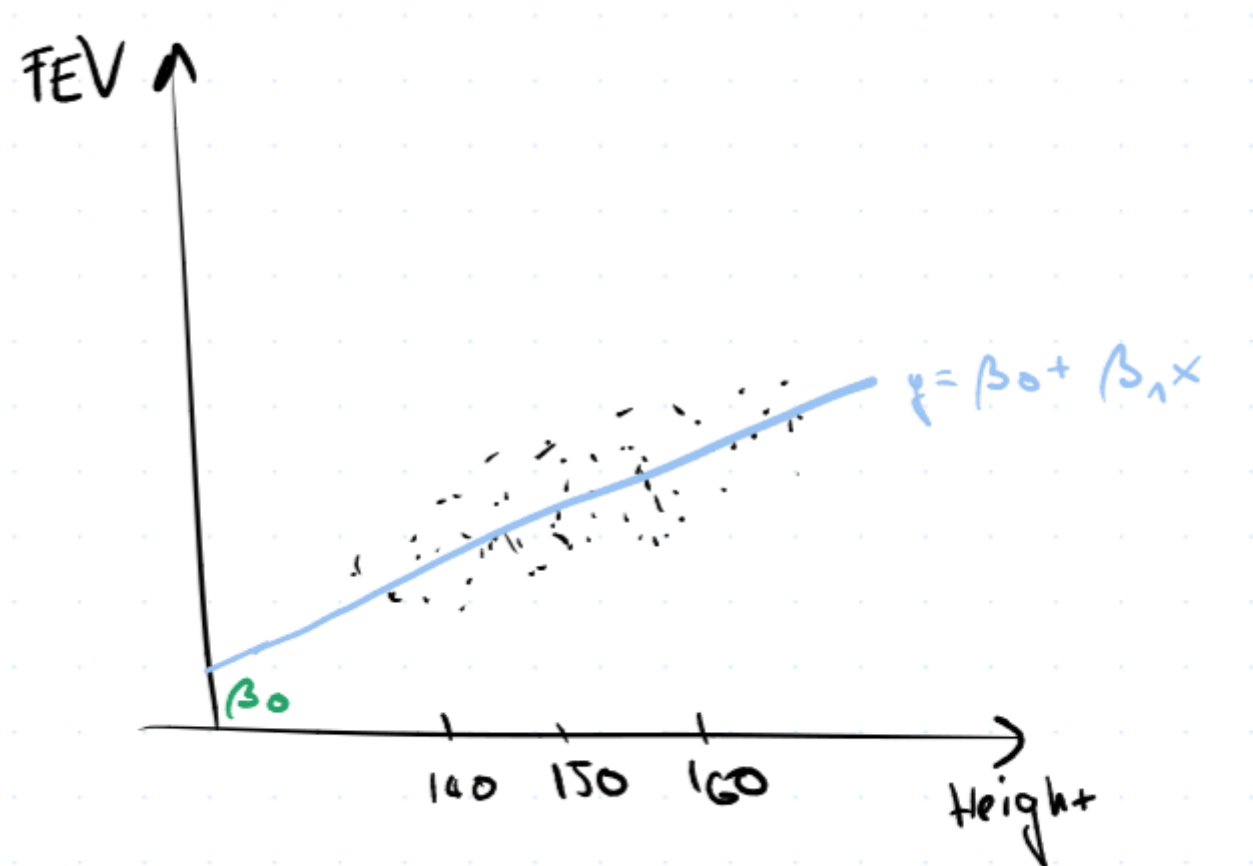


02 / c)

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

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

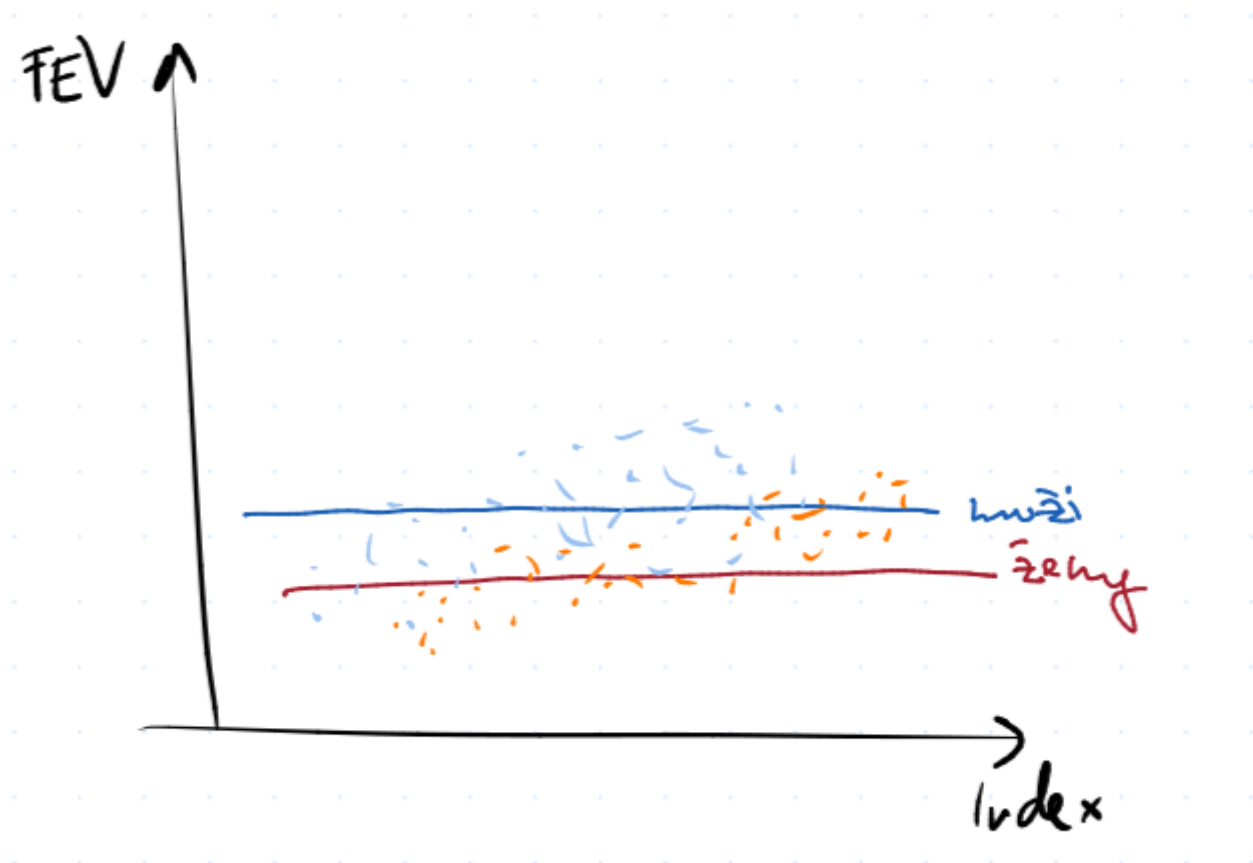
02 / b)

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

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



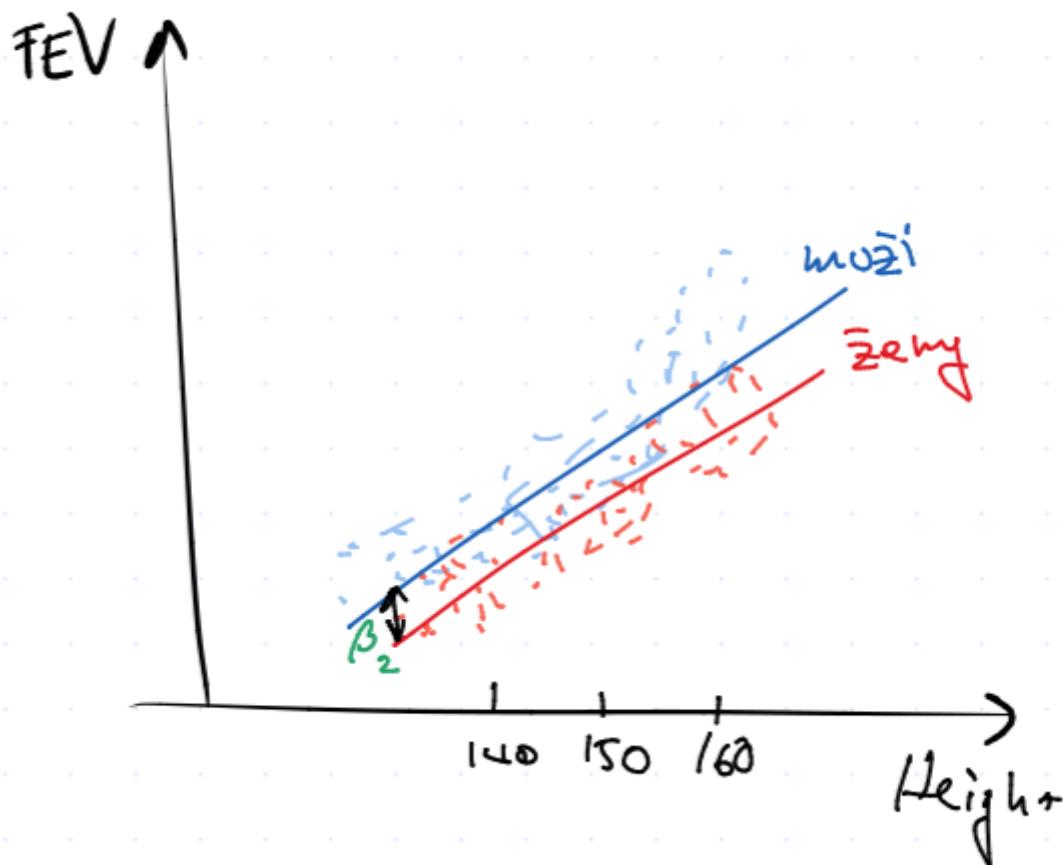
02/ d)

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

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

02 / e)

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



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

s významem koeficientů

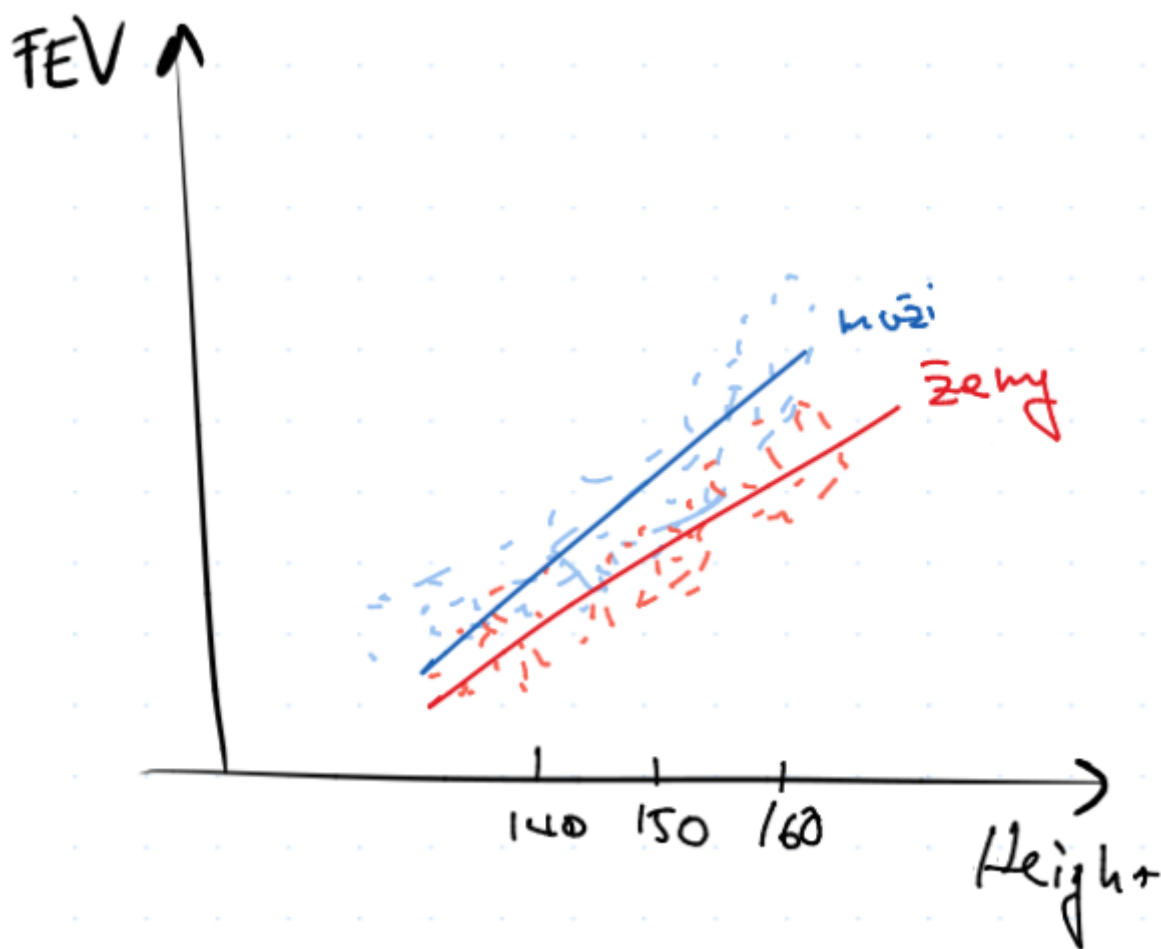
- β_0 ... střední hodnota predikce při nulových hodnotách ostatních prediktorů (nulová výška a ženské pohlaví)
- β_1 ... změna střední hodnoty predikce při nárůstu výšky (prediktor Height) o 1 cm pro ženy
- β_2 ... rozdíl střední hodnoty predikce mezi muži a ženami

02 / g)

$FEV_i = \beta_0 + \beta_1 \cdot \text{Height}_i + \beta_2 \text{Sex}_i + \beta_3 (\text{Sex}_i \times \text{Height}_i) + \epsilon_i$, 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 $y = \beta_0 + \beta_1 x + \beta_2 \text{I} \set{\text{Sex} = \text{"male"}} + \beta_3 x \text{I} \set{\text{Sex} = \text{"male"}}$, 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 $E Y$ ”
 - ▶ 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

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

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 sestojíme 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 jisté matice Σ^{-1} pozitivně definitní. Nechť $\forall x \neq 0$ libovolné, pak $\forall x^T \Sigma^{-1} \forall x = \forall x^T U \Lambda^{-1} U^T \forall x = \forall x^T U \Lambda^{-\frac{1}{2}} \Lambda^{-\frac{1}{2}} U^T \forall x$, kde $\Lambda^{-\frac{1}{2}}$ je matice s

převrácenými hodnotami odmocnin vlastních čísel matice Σ na diagonále a tedy
$$\underbrace{(\mathbf{x}^T U \Lambda^{-\frac{1}{2}})}_{\mathbf{y}} \underbrace{\Lambda^{-\frac{1}{2}} U^T \mathbf{x}}_{\mathbf{y}^T} = \sum_{i=1}^m y_i^2 > 0$$

d)

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

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

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

Nakonec pro $c > 0$ je
$$(\mathbf{x} - \boldsymbol{\mu})^T U \Lambda^{-1} U^T (\mathbf{x} - \boldsymbol{\mu}) = \mathbf{y}^T \Lambda^{-1} \mathbf{y}$$
 A pro $m = 2$ tedy
$$\begin{pmatrix} \frac{1}{\lambda_1} & 0 \\ 0 & \frac{1}{\lambda_2} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \mathbf{c}$$

$$\frac{y_1^2}{\lambda_1} + \frac{y_2^2}{\lambda_2} = c,$$
 což je **rovnice elipsy** se středem (μ_1, μ_2) a směry os jsou právě vlastní vektory Σ . Nakonec délky poloos budou $\sqrt{c \lambda_i}$

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

Příklad 4

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

a)

$$h(\Sigma) = r, \quad \text{kde } 0 < r \leq m.$$

Nahradíme U za matici
$$U_1 = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_r \\ \mathbf{u}_{r+1} & \dots & \mathbf{u}_m \end{pmatrix}$$
 a také
$$\Lambda_1 = \text{diag}(\lambda_1, \dots, \lambda_r)$$

b)

Sestrojíme 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

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

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 $\lambda^2 = \lambda \implies \lambda(\lambda - 1) = 0 \implies \lambda = 0 \text{ or } \lambda = 1$

“ 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 $h(P) = h(U\Lambda U^T) = h(\Lambda) = |\{\lambda_i, \lambda_i = 1\}|$

c)

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

d)

Je-li vektor $y \in \text{im } P$, pak $P y = y$. Pokud $y \in \text{im } P$, pak $\exists z$, že $y = P z$

Definice obrazu $\text{im } P$: $\text{im } P = \{y \in \mathbb{R}^n \mid \exists z \in \mathbb{R}^n : y = P z\}$

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

e)

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

Nechť $z \in \mathbb{R}^n$ je libovolné, projekce z do menšího prostoru je prvek většího prostoru. $\tilde{P} z \in \text{im } \{\tilde{P}\} \leq \text{im } P$ Pak podle **d)** platí $P(\tilde{P} z) = \tilde{P} z$
 $P(\tilde{P} z) - \tilde{P} z = 0 \implies (P - \tilde{P}) z = 0$, nicméně vektor z byl libovolný. To tedy znamená, že zobrazení $P - \tilde{P}$ pošle všechny $z \in \mathbb{R}^n$ na nulový vektor, tj. $\ker(P - \tilde{P}) = \mathbb{R}^n$. Z lineární algebry víme, že $\dim \ker(A) + \dim \text{im } A = n$ pro A tvaru $n \times n$. A tedy $\dim \text{im } \{P - \tilde{P}\} = 0 \implies P - \tilde{P} = 0$.
 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 $\mathbf{x} \in \mathbb{R}^n$ pevné. Mějme $\mathbf{P} \mathbf{z} \in \lim P$ Vezměme libovolné $\mathbf{y} \in \lim P$ a spočítáme $\|\mathbf{y} - \mathbf{x}\|^2$ Pak $\|\mathbf{y} - \mathbf{x}\|^2 = \|\mathbf{P} \mathbf{x} - \mathbf{x} + \mathbf{y} - \mathbf{P} \mathbf{x}\|^2 = \text{scal} \{(\mathbf{P} \mathbf{x} - \mathbf{x}) + (\mathbf{y} - \mathbf{P} \mathbf{x})\} \{(\mathbf{P} \mathbf{x} - \mathbf{x}) + (\mathbf{y} - \mathbf{P} \mathbf{x})\}$

Pro skalární součin platí $\text{scal} \{\mathbf{u}\} \{\mathbf{v}\} = \text{scal} \{\mathbf{v}\} \{\mathbf{u}\}$
 $\text{scal} \{\mathbf{u} + \mathbf{v}\} \{\mathbf{w}\} = \text{scal} \{\mathbf{u}\} \{\mathbf{w}\} + \text{scal} \{\mathbf{v}\} \{\mathbf{w}\}$
 $\text{scal} \{\mathbf{u}\} \{a\mathbf{v}\} = a \text{scal} \{\mathbf{u}\} \{\mathbf{v}\}, ; a \in \mathbb{R}$

Pak dostáváme $\|\mathbf{y} - \mathbf{x}\|^2 = \|\mathbf{P} \mathbf{x} - \mathbf{x}\|^2 + 2 \text{scal} \{\mathbf{P} \mathbf{x} - \mathbf{x}\} \{\mathbf{y} - \mathbf{P} \mathbf{x}\} + \|\mathbf{y} - \mathbf{P} \mathbf{x}\|^2$

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

$$\text{scal} \{\mathbf{P} \mathbf{x} - \mathbf{x}\} \{\mathbf{y} - \mathbf{P} \mathbf{x}\} = (\mathbf{y} - \mathbf{P} \mathbf{x})^T (\mathbf{P} \mathbf{x} - \mathbf{x}) = \mathbf{y}^T (\mathbf{P} \mathbf{x} - \mathbf{x}) - (\mathbf{P} \mathbf{x})^T \mathbf{P} \mathbf{x} - \mathbf{x}^T \mathbf{x} = \mathbf{y}^T \mathbf{P} \mathbf{x} - \mathbf{y}^T \mathbf{x} - \mathbf{x}^T \mathbf{P} \mathbf{x} + \mathbf{x}^T \mathbf{x} = \mathbf{y}^T \mathbf{P} \mathbf{x} - \mathbf{y}^T \mathbf{x}$$

Jelikož $\mathbf{y} \in \lim P$, tak jistě $\exists \mathbf{z} \in \mathbb{R}^n$ takové, že $\mathbf{y} = \mathbf{P} \mathbf{z}$. Z toho plyne

$$\mathbf{y}^T \mathbf{P} \mathbf{x} - \mathbf{y}^T \mathbf{x} = (\mathbf{P} \mathbf{z})^T \mathbf{P} \mathbf{x} - (\mathbf{P} \mathbf{z})^T \mathbf{x} = \mathbf{z}^T \mathbf{P}^T \mathbf{P} \mathbf{x} - \mathbf{z}^T \mathbf{P}^T \mathbf{x} = \mathbf{z}^T (\mathbf{P}^T \mathbf{P} \mathbf{x} - \mathbf{P}^T \mathbf{x}) = 0$$

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

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 vypočtete 95% kvantil a zaznačte do grafu

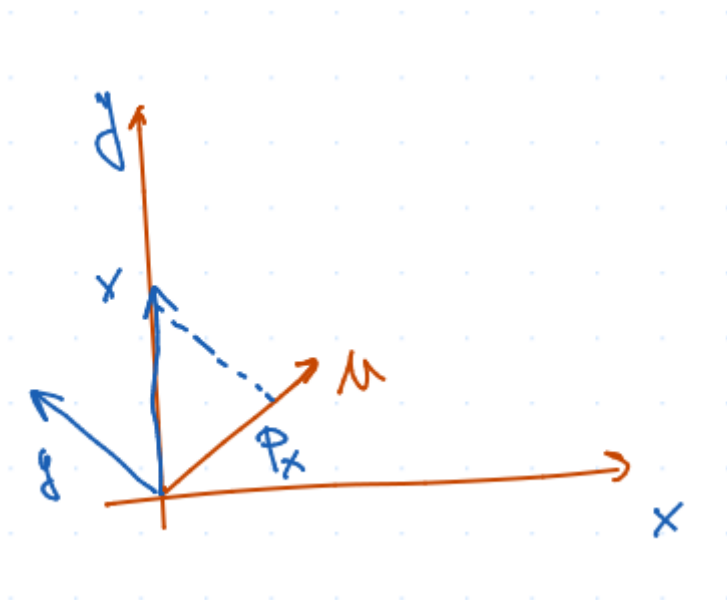
5. cvičení

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

4. cvičení / 2. příklad

a) i b)

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



Pro ortogonální projekci platí

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

Pak $P = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\|\mathbf{u}\|^2}$ a tedy $PP = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\|\mathbf{u}\|^2} \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\|\mathbf{u}\|^2}$

Z definice skalárního součinu $\langle \mathbf{v}, \mathbf{u} \rangle = \|\mathbf{u}\|^2$ a proto $PP = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\|\mathbf{u}\|^2} = P$, což zvláště platí pro $\|\mathbf{u}\| = 1$. Pro nějaké $\mathbf{x} \in \mathbb{R}^n$ máme $P \mathbf{x} = \frac{\langle \mathbf{v}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{x} \rangle}{\|\mathbf{u}\|^2} = \frac{\langle \mathbf{v}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{x} \rangle}{\|\mathbf{u}\|^2} = \underbrace{\frac{\langle \mathbf{v}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{x} \rangle}{\|\mathbf{u}\|^2}}_{\in \mathbb{R}} \mathbf{u} \in \text{im } U$

c)

Mějme $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ ortonormální vektory, pak matice $P = U U^T$, kde $U = (\mathbf{u}_1 \mid \mathbf{u}_2 \mid \dots \mid \mathbf{u}_p)$, je ortogonální projekce na $\text{im } U$.

Ukažme $P \cdot P = U \underbrace{U^T U}_I U^T = U U^T$ a $P \mathbf{x} = U U^T \mathbf{x} = U \begin{pmatrix} \langle \mathbf{u}_1, \mathbf{x} \rangle \\ \langle \mathbf{u}_2, \mathbf{x} \rangle \\ \vdots \\ \langle \mathbf{u}_p, \mathbf{x} \rangle \end{pmatrix} = \underbrace{\mathbf{u}_1 \langle \mathbf{u}_1, \mathbf{x} \rangle}_{\in \mathbb{R}} + \dots + \underbrace{\mathbf{u}_p \langle \mathbf{u}_p, \mathbf{x} \rangle}_{\in \mathbb{R}} \in \text{im } U$, což je lineární kombinace vektorů $\mathbf{u}_1, \dots, \mathbf{u}_p$ a jistě tedy $P \mathbf{x} \in \text{im } U$.

d) i e)

Máme lineárně nezávislé vektory $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$, 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 $A \mathbf{x} = \mathbf{y} \in \text{im } A$, pak $P \mathbf{y} = (A(A^T A)^{-1} A^T) \mathbf{y} = A \underbrace{(A^T A)^{-1} A^T A}_I \mathbf{x} = A \mathbf{x} = \mathbf{y}$

„ Zde jsme jen ukázali něco o $\mathbf{y} \in \text{im } A$, nikoliv o obecném $\mathbf{x} \in \mathbb{R}^p$

5. cvičení / 1. příklad

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

a)

Ukažme $\mathbf{A} \mathbf{X} + \mathbf{b} = \mathbf{A} \mathbf{X} + \mathbf{b}$, což je analogické k jednorozměrnému případu.

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

b)

Ukažme $\text{Var}(\mathbf{A} \mathbf{X} + \mathbf{b}) = \mathbf{A} \text{Var } \mathbf{X} \mathbf{A}^T$ což je opět analogie k $D(aX + b) = a^2 DX$.

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

c)

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

Obecně platí $\text{Var}(\mathbf{X}) = \mathbf{X} \mathbf{X}^T - \mathbf{X} \mathbf{X}^T$ a tedy $\mathbf{X}^T \mathbf{X} = \mathbf{X}^T \mathbf{X} + \text{tr}(\mathbf{X} \mathbf{X}^T) = \text{tr}(\mathbf{X} \mathbf{X}^T) = \text{tr}(\text{Var}(\mathbf{X}) + \mathbf{X} \mathbf{X}^T) = \text{tr}(\text{Var}(\mathbf{X})) + \text{tr}(\mathbf{X} \mathbf{X}^T) = \mathbf{X}^T \mathbf{X} + \text{tr}(\text{Var}(\mathbf{X}))$

7. cvičení

$\$ \def\mc{#1}{\mathcal{#1}} \def\sc{#1#2}{\langle #1, #2 \rangle} \def\N{\mathbb{N}} \def\R{\mathbb{R}} \def\Q{\mathbb{Q}} \def\Z{\mathbb{Z}} \def\D{\mathbb{D}} \def\bm{#1}{\boldsymbol{#1}} \def\vv{#1}{\mathbf{#1}} \def\vp{#1}{\mathrm{#1}} \def\fl{#1}{\lfloor #1 \rfloor} \def\ce{#1}{\lceil #1 \rceil} \def\grad{#1}{\mathrm{grad} , #1} \def\ve{\mathrm{varepsilon}} \def\im{#1}{\mathrm{im}(#1)} \def\tr{#1}{\mathrm{tr}(#1)} \def\norm{#1}{\left\| \right\|_{#1}} \def\sc{#1#2}{\langle #1, #2 \rangle} \def\ex{#1}{\mathrm{E} , \left(#1 \right)} \def\exv{#1}{\mathrm{E} , \vv{#1}} \$$

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

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

A platí $TSS = RSS + ESS$

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

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

9. cvičení

a) IS pro β_i :
$$T_i = \frac{\hat{\beta}_i}{\sqrt{\hat{\sigma}^2 (\mathbf{X}^T \mathbf{X})^{-1}_{i,i}}} \sim t(n-p)$$

Pak
$$P\left(t_{\frac{\alpha}{2}(n-p)} \leq T_i \leq t_{1-\frac{\alpha}{2}(n-p)}\right) = 1 - \alpha$$
$$t_{\frac{\alpha}{2}(n-p)} \leq T_i \leq t_{1-\frac{\alpha}{2}(n-p)} \iff t_{\frac{\alpha}{2}(n-p)} \leq \frac{\hat{\beta}_i - \hat{\beta}_i}{\sqrt{\hat{\sigma}^2 (\mathbf{X}^T \mathbf{X})^{-1}_{i,i}}} \leq t_{1-\frac{\alpha}{2}(n-p)}$$
$$Z \text{ čehož dostaneme } \hat{\beta}_i \in \left(\hat{\beta}_i \pm t_{1-\frac{\alpha}{2}(n-p)} \sqrt{\hat{\sigma}^2 (\mathbf{X}^T \mathbf{X})^{-1}_{i,i}}\right)$$

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

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

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 $\mathbf{a} = (0, 1, 0, 1)$

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

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

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

10. cvičení

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

Scheffeho věta $\sum_{j=1}^m \left(\frac{(\hat{\beta}_j - \beta_j)^T A (\hat{\beta}_j - \beta_j)}{\sum_{j=1}^m (\hat{\beta}_j - \beta_j)^T A (\hat{\beta}_j - \beta_j)} \right) \leq m F_{1-\alpha}(m, n-p)$ pro všechny $\beta \in \mathbb{R}^m$, je-li matice A typu $m \times p$ plně hodnosti.

Příklad
$$Y_i = \beta_0 + \beta_1 \cdot \text{Height}_i + \beta_2 \cdot \text{Sex}_i + \beta_3 \cdot (\text{Height} + \text{Sex})_i + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2)$$
 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 \mathbf{v} a \mathbf{A} :

- d: $\forall b = \begin{pmatrix} 1 \\ x \end{pmatrix} \in \mathbb{R}^2$, pak $\overbrace{\begin{pmatrix} 1 & x \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}}^A \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \hat{\beta}_4 \end{pmatrix}$
- ch: $\forall b = \begin{pmatrix} 1 \\ x \end{pmatrix}$, pak $\overbrace{\begin{pmatrix} 1 & x \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}}^A \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \hat{\beta}_4 \end{pmatrix}$

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

$$P(\|\hat{\beta} - y\| \leq \sqrt{2 F_{1-\alpha}(2, n-4)} \sigma \mid X) = 1 - \alpha$$

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

$$\right) = 1 - \alpha \quad \square \square$$