ČASOVÉ ŘADY

VíTĚZSLAV VESELÝ

PODPORA VÝUKY NA POČÍTAČOVÉ SÍTI

`http://www.math.muni.cz/~vesely`

Algoritmy pro časové řady v MATLABu:

`d = vesely('cas_rady')` ... připojení a výpis obsahu knihovny

`d` ... výstup=úplná cesta do adresáře knihovny.

1. Úvod

1.1. **Ukázky časových řad.**

Časová řada = soubor pozorování nějaké veličiny $\{x_t, t \in T\}$, kde t je zpravidla čas a $T \subseteq \mathbb{R}$.

Obr. 1.1:
Měření proudu procházejícího odporem r v obvodu se střídavým napětím,
$v(t) = a\cos(vt + \Theta)$, tj. $x_t = \frac{a}{r}\cos(vt + \Theta)$.

Obr. 1.2 (MATLAB: `getdata('brockwell.dat/uspop.dat')`):
Růst populace v USA v létech 1790–1980 sledovaný v desetiletých intervalech.

Obr. 1.3 (MATLAB: `getdata('brockwell.dat/strikes.dat')`):
Počet stávek v USA v létech 1951–1980.

Obr. 1.4: Výsledky Národní a Americké ligy v baseballu v létech 1933–1980.

Obr. 1.5 (MATLAB: `getdata('brockwell.dat/sunspots.dat')`):
Počty slunečních skvrn v létech 1770–1869.

Obr. 1.6 (MATLAB: `getdata('brockwell.dat/deaths.dat')`):
Počet úmrtí při nehodách v USA v létech 1973–1978.

1.2. **Oblasti uplatnění metod analýzy časových řad.**

fyzika, technika:
- seismický záznam v geofyzice.
- řada nejvyšších denních teplot v meteorologii.
- průběh výstupního signálu určitého elektrického přístroje.
- tenzometrické měření povrchového napětí v provozu namáhané strojní součástky.

biologie, ekologie: sledování různých parametrů znečištění ovzduší.
medicína: záznam EKG nebo EEG.
společenské vědy: změny v počtu a složení obyvatelstva.
sociologie: vývoj rozvodovosti.
EKONOMIE: Teorie časových řad = jedna z nejdůležitějších kvantitativních metod pro analýzu ekonomických dat, např.:
- analýza poptávky po určitém výrobku
- analýza objemu zemědělské produkce
- analýza počtu cestujících v letecké dopravě
- analýza vývoje kurzu akcií na burze

Date: 25. října 2001.

1.3. **Cíl analýzy.**

Porozumění mechanismu, jímž se generují sledované údaje.
Znalost modelu tohoto mechanismu $\Rightarrow$ znalost algoritmu, jímž můžeme chování tohoto mechanismu simulovat na počítači $\Rightarrow$ schopnost popsat s jistou přesností jeho chování:

- mezi časovými okamžiky měření (**interpolace**)
- v budoucnosti (**extrapolace, prognóza**)
- s cílem řídit a optimalizovat činnost určitého systému vhodnou volbou vstupních a počátečních podmínek (**regulace**), např. regulace složitých technologických procesů.

1.4. **Některé specifické problémy analýzy časových řad.**

a) Volba okamžiků pozorování:
- jsou přímo **diskrétní svou povahou**, např. úroda obilí za jednotlivé roky.
- vznikají **diskretizací spojité časové řady**, např. teplota v danou denní dobu na daném místě.
- vznikají **akumulací (agregací)** hodnot za určité období, např. denní množství srážek, roční výroba závodu. Místo akumulace se někdy provádí průměrování.

Je-li dána možnost volby, je třeba jí věnovat pozornost:
- **málo bodů** $\Rightarrow$ unikne charakteristický rys řady.
- **mnoho bodů** $\Rightarrow$ zvýší se výpočetní náročnost.
- **ekvidistantní diskretizace** zpravidla usnadní numerické zpracování, ale neumožňuje adaptivně měnit hustotu diskretizace v závislosti na lokálním charakteru řady.
- **při agregaci** se mohou porušit vlastnosti původní řady.

b) Problémy s kalendářem:
- různá délka kalendářních měsíců.
- 4 nebo 5 víkendů v měsíci.
- různý počet pracovních dnů v měsíci.
- pohyblivé svátky: např. svátek na začátku měsíce sníží prodej potravin za tento měsíc, ale zvýší jej za předchozí v důsledku efektu předzásobení.
  Příklad 'očištění' časové řady např. od proměnlivé délky měsíce:
  zavedeme 'standardní' měsíc o délce 30 dnů a pak údaj třebas o produkci za leden přenásobíme korekčním faktorem $\frac{30}{31}$.

c) Problémy s nekompatibilitou jednotlivých měření:

Příklad: hodnota nějakého ukazatele se jeden rok týká např. 85 podniků, další rok jen 82 apod.

d) Problémy s délkou časových řad:
- zvětšení počtu měření (např. půlením časových intervalů mezi body) nemusí vždy znamenat zvětšení množství informace.
- někdy se mohou objevit protichůdné tendence: metoda zpracování vyžaduje delší řadu, ale na druhé straně řada vzniklá dlouhodobým sledováním může měnit charakteristiky svého modelu v čase $\Rightarrow$ obtíže s konstrukcí modelu.

1.5. **Označení a základy matematické statistiky.**

$s := v$, resp. $v =: s$ … označení výrazu v symbolem s.

Číselné obory:
$\mathbb{N}$ … množina všech přirozených čísel
$\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ … množina všech nezáporných celých čísel
$\mathbb{Z}$ … množina všech celých čísel
$\mathbb{R}$ … množina všech reálných čísel
$\mathbb{R}^+$ … množina všech nezáporných reálných čísel
$\mathbb{C}$ … množina všech komplexních čísel
$(\cdot)^+ : \mathbb{R} \to \mathbb{R}^+$ … zobrazení definované předpisem $(x)^+ = \max(0, x)$
(a, b) … otevřený interval

$[a, b]$ ... uzavřený interval
$\mathcal{I}(a, b) \;=\; \{x \mid \min(a, b) < x < \max(a, b)\}$
$\mathcal{I}[a, b] \;=\; \{x \mid \min(a, b) \leq x \leq \max(a, b)\}$.

Vektory:
$\boldsymbol{x} = (x_1, \ldots, x_n)^T \in \mathbb{C}^n$ ... zpravidla sloupcové
$\boldsymbol{x} + h = (x_1 + h, \ldots, x_n + h)^T,\ h \in \mathbb{C}$
$\boldsymbol{t} = (t_1, \ldots, t_k)^T \in \mathbb{N}^k,\ t_i \in \{1, \ldots, n\}$ pro $i = 1, \ldots, k,\ k \leq n \quad \Rightarrow \quad \boldsymbol{x_t} := (x_{t_1}, \ldots, x_{t_k})^T \in \mathbb{C}^k$
$1 \leq i \leq n \quad \Rightarrow \quad \boldsymbol{x}(i) = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)^T$
$f(\boldsymbol{x}) = f(x_1, \ldots, x_n),\ d\boldsymbol{x} = dx_1 \ldots dx_n$
$\mathbf{0}, \mathbf{0}_{n\times 1}$ ... nulový sloupcový vektor délky n.

Matice:
$A, A_{m\times n} = [a_{ij}] = [A(i, j)]$ ... matice rozměru $m \times n$
$\mathcal{R}(A) = \{\boldsymbol{y} \mid \boldsymbol{y} = A\boldsymbol{x}\}$ ... obor hodnot operátoru A
$\mathcal{N}(A) = \{\boldsymbol{x} \mid A\boldsymbol{x} = 0\}$ ... jádro operátoru A
$A^T = [a_{ji}]$ ... transponovaná matice
$A^* = [\bar{a}_{ji}]$ ... hermitovsky sdružená matice
$I, I_n = I_{n\times n} = [\delta_{ij}]$ ... jednotková matice řádu n
$|A|$ ... determinant čtvercové matice A
$0, 0_{m\times n} = [0]$ ... nulová matice rozměru $m \times n$

$$\mathrm{diag}(\boldsymbol{x}) = \begin{bmatrix} x_1 & 0 & \ldots & 0 \\ 0 & x_2 & \ldots & 0 \\ & & \vdots & \\ 0 & 0 & \ldots & x_n \end{bmatrix} \ldots \text{ diagonální matice}$$

$A(i, :) = (a_{i1}, \ldots, a_{in}) =: r_i$ ... i-tý řádek matice A ve stylu MATLAB
$A(:, j) = (a_{1j}, \ldots, a_{mj})^T =: s_j$ ... j-tý sloupec matice A ve stylu MATLAB
$A = [r_1; \ldots; r_m] = [s_1, \ldots, s_n]$ ... blokový zápis matice A ve stylu MATLAB
$A > 0$ (resp. $A \geq 0$) ... pozitivně (semi)definitní matice.

Norma a skalární součin vektorů:
$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^n x_i\, \bar{y}_i = \boldsymbol{y}^* \boldsymbol{x}$, speciálně:
$\boldsymbol{y} = A^* \boldsymbol{x} \Leftrightarrow y_i = \langle \boldsymbol{x}, A(:, i) \rangle$ pro $i = 1, 2, \ldots, n$
$\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} = \sqrt{\sum_{i=1}^n |x_i|^2}$ ... Euklidovská norma.

Schwarzova nerovnost:
$|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\|$, kde rovnost nastane právě když vektory $\boldsymbol{x}$ a $\boldsymbol{y}$ jsou lineárně závislé.

Pravděpodobnostní prostor $(\Omega, \mathcal{A}, P)$:
Ω ... základní prostor elementárních jevů
$\mathcal{A} \subseteq 2^\Omega$ ... σ-algebra náhodných jevů
P ... pravděpodobnostní míra na $\mathcal{A}$
Všechny náhodné veličiny budeme vždy uvažovat nad týmž pravděpodobnostním prostorem. Komplexní náhodnou veličinou budeme rozumět veličinu $X = X_r + iX_i$, kde X_r a X_i jsou reálné náhodné veličiny představující po řadě reálnou a imaginární část X
$\mathbb{X} = (X_1, \ldots, X_n)^T$ ... (komplexní) náhodný vektor tvořený (komplexními) náhod. veličinami X_i.

Střední hodnota:
$\mu = \mu_X = \mathrm{E}X$ ... střední hodnota náhodné veličiny X
$\boldsymbol{\mu} = \boldsymbol{\mu}_{\mathbb{X}} = \mathrm{E}\mathbb{X} = (\mathrm{E}X_1, \ldots, \mathrm{E}X_n)^T$ ... střední hodnota náhodného vektoru $\mathbb{X}$.

Rozptyl a kovariance náhodných veličin:
$\sigma^2 = \sigma_X^2 = \mathrm{var}X = \mathrm{E}|X - \mathrm{E}X|^2 = \mathrm{E}|X|^2 - |\mathrm{E}X|^2 \geq 0$ ... rozptyl X
$\sigma_{XY} = \mathrm{cov}(X, Y) = \mathrm{E}(X - \mathrm{E}X)\overline{(Y - \mathrm{E}Y)} = \mathrm{E}X\overline{Y} - (\mathrm{E}X)(\overline{\mathrm{E}Y})$ ... kovariance X a Y
$\mathrm{cov}(X, X) = \mathrm{var}X,\ \mathrm{cov}(Y, X) = \overline{\mathrm{cov}(X, Y)}$
$\mathrm{cov}(\sum_r X_r, \sum_s Y_s) = \sum_r \sum_s \mathrm{cov}(X_r, Y_s)$ a odtud speciálně:
$\mathrm{var}(X + Y) = \mathrm{var}X + \mathrm{cov}(X, Y) + \mathrm{cov}(Y, X) + \mathrm{var}Y = \mathrm{var}X + 2Re\,\mathrm{cov}(X, Y) + \mathrm{var}Y$.

Varianční a kovarianční matice náhodných vektorů:
$\Sigma_{\mathbb{X}} = \mathrm{var}\mathbb{X} = [\mathrm{cov}(X_i, X_j)] = \mathrm{E}(\mathbb{X} - \mathrm{E}\mathbb{X})(\mathbb{X} - \mathrm{E}\mathbb{X})^* = \mathrm{E}\mathbb{X}\mathbb{X}^* - (\mathrm{E}\mathbb{X})(\mathrm{E}\mathbb{X})^*$ ... varianční matice $\mathbb{X}$
$\Sigma_{\mathbb{X}\mathbb{Y}} = \mathrm{cov}(\mathbb{X}, \mathbb{Y}) = [\mathrm{cov}(X_i, Y_j)] = \mathrm{E}(\mathbb{X} - \mathrm{E}\mathbb{X})(\mathbb{Y} - \mathrm{E}\mathbb{Y})^* = \mathrm{E}\mathbb{X}\mathbb{Y}^* - (\mathrm{E}\mathbb{X})(\mathrm{E}\mathbb{Y})^*$ ...
... kovarianční matice $\mathbb{X}$ a $\mathbb{Y}$

$\operatorname{cov}(\mathbb{X},\mathbb{X}) = \operatorname{var}\mathbb{X},\ \operatorname{cov}(\mathbb{Y},\mathbb{X}) = \operatorname{cov}(\mathbb{X},\mathbb{Y})^* \Rightarrow \operatorname{var}\mathbb{X} = (\operatorname{var}\mathbb{X})^* \ldots$
… **varianční matice $\mathbb{X}$ je hermitovská**.

Pro konstantní komplexní vektory $\boldsymbol{a}$ a $\boldsymbol{c}$ a matice B a D odpovídajících rozměrů platí:
$\operatorname{cov}(\boldsymbol{a} + B\mathbb{X}, \boldsymbol{c} + D\mathbb{Y}) = \operatorname{cov}(B\mathbb{X}, D\mathbb{Y}) = B\operatorname{cov}(\mathbb{X},\mathbb{Y})\,D^*$.

$$\Downarrow \mathbb{X} = \mathbb{Y}$$

$\operatorname{var}(\boldsymbol{a} + B\mathbb{X}) = \operatorname{cov}(\boldsymbol{a} + B\mathbb{X}, \boldsymbol{a} + B\mathbb{X}) = \operatorname{cov}(B\mathbb{X}, B\mathbb{X}) = B\operatorname{var}(\mathbb{X})\,B^*$.

$$\Downarrow \boldsymbol{b}^* = B$$

$0 \le \operatorname{var}(\boldsymbol{b}^*\mathbb{X}) = \boldsymbol{b}^*\operatorname{var}\mathbb{X}\,\boldsymbol{b} \Rightarrow \operatorname{var}\mathbb{X} \ge 0 \ldots$ **varianční matice je pozitivně semidefinitní**.
$\Sigma_{\mathbb{X}}$ je tedy celkem hermitovská a pozitivně semidefinitní a má proto reálná nezáporná vlastní čísla λ_i.
Zřejmě existuje matice $\Sigma_{\mathbb{X}}^{\frac{1}{2}}$, jejíž vlastní čísla jsou $\lambda_i^{\frac{1}{2}}$ a přitom platí: $\Sigma_{\mathbb{X}} = \Sigma_{\mathbb{X}}^{\frac{1}{2}}\,\Sigma_{\mathbb{X}}^{\frac{1}{2}}$.

Dále platí
$\operatorname{cov}(\sum_r \mathbb{X}_r, \sum_s \mathbb{Y}_s) = \sum_r\sum_s \operatorname{cov}(\mathbb{X}_r, \mathbb{Y}_s)$ a odtud speciálně:
$\operatorname{var}(\mathbb{X}+\mathbb{Y}) = \operatorname{var}\mathbb{X} + \operatorname{cov}(\mathbb{X},\mathbb{Y}) + \operatorname{cov}(\mathbb{Y},\mathbb{X}) + \operatorname{var}\mathbb{Y} = \operatorname{var}\mathbb{X} + 2\mathit{Re}\operatorname{cov}(\mathbb{X},\mathbb{Y}) + \operatorname{var}\mathbb{Y}$.

1.6. Normální rozdělení a rozdělení z něj odvozená.

1. <u>*Normální rozdělení:*</u>
   $X \sim N(\mu, \sigma^2),\ \mu = \mathrm{E}X,\ \sigma^2 = \operatorname{var}X \ldots$
   … reálná náhodná veličina s **normálním (gaussovským) rozdělením**;
   $f(x) = (\sqrt{2\pi}\sigma)^{-1}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \ldots$ **hustota** náhodné veličiny X;
   $\Phi(t) := \mathrm{E}e^{itX} = e^{it\mu - \frac{1}{2}\sigma^2 t^2} \ldots$ **charakteristická funkce** náhodné veličiny X;
   $U \sim N(0,1) \ldots$ **standardizované normální rozdělení** náhodné veličiny U;
   $u_\alpha = F^{-1}(\alpha) \ldots$ α-kvantil pro U, kde $F(x) = \int_{-\infty}^{x} f(t)\,dt$ značí distribuční funkci U;
   $\mathbb{X} \sim N_n(\boldsymbol{\mu}, V),\ \boldsymbol{\mu} = \mathrm{E}\mathbb{X},\ V = \operatorname{var}\mathbb{X} \ldots$
   … reálný náhodný vektor s **n-rozměrným normálním rozdělením**;
   $f(\boldsymbol{x}) = (\sqrt{(2\pi)^n|V|})^{-1}e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T V^{-1}(\boldsymbol{x}-\boldsymbol{\mu})} \ldots$ **n-rozměrná hustota** náhodného vektoru $\mathbb{X}$;
   $\Phi_n(\boldsymbol{t}) := \mathrm{E}e^{i\boldsymbol{t}^T\mathbb{X}} = e^{i\boldsymbol{t}^T\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{t}^T V \boldsymbol{t}} \ldots$ **charakteristická funkce** náhodného vektoru $\mathbb{X}$;
   $\mathbb{U} \sim N_n(\boldsymbol{0}, I_n) \ldots$ **standardizované normální rozdělení** náhodného vektoru $\mathbb{U}$.

   Platí

   $X \sim N(\mu,\sigma^2) \quad\Rightarrow\quad a + bX \sim N(a + b\mu, b^2\sigma^2)$ pro $a, b \in \mathbb{R}$;
   $\mathbb{X} \sim N_n(\boldsymbol{\mu}, V) \quad\Rightarrow\quad \boldsymbol{a} + B\mathbb{X} \sim N_m(\boldsymbol{a} + B\boldsymbol{\mu}, BVB^T)$ pro $\boldsymbol{a} \in \mathbb{R}^m$ a matici $B = B_{m\times n}$ nad $\mathbb{R}$.

2. <u>*Rozdělení $\chi^2(n)$:*</u>
   Nechť $U_i \sim N(0,1)$ pro $i = 1, \ldots, n$ jsou stochasticky nezávislé, pak náhodná veličina
   $C = \sum_{i=1}^n U_i^2 \sim \chi^2(n)$ má **Pearsonovo "chí kvadrát" rozdělení o n stupních volnosti**;
   $\chi_\alpha(n) \ldots$ α-kvantil pro C.

   Platí

   $C_i \sim \chi^2(n_i)$ pro $i = 1, \ldots, m$ stochasticky nezávislé $\quad\Rightarrow\quad C = \sum_{i=1}^m C_i \sim \chi^2(n_1+n_2+\cdots+n_m)$.

3. <u>*Studentovo t rozdělení:*</u>
   Nechť $U \sim N(0,1)$ a $C \sim \chi^2(k)$ jsou stochasticky nezávislé, pak náhodná veličina
   $T = \frac{U}{\sqrt{C/k}} \sim t(k)$ má **Studentovo t rozdělení o k stupních volnosti**;
   $t_\alpha(k) \ldots$ α-kvantil pro T.

4. <u>*Fisher-Snedecorovo F rozdělení:*</u>
   Nechť $C_1 \sim \chi^2(n_1)$ a $C_2 \sim \chi^2(n_2)$ jsou stochasticky nezávislé, pak náhodná veličina
   $F = \frac{C_1/n_1}{C_2/n_2} \sim F(n_1, n_2)$ má **Fisher-Snedecorovo F rozdělení s n_1 a n_2 stupni volnosti**;
   $F_\alpha(n_1, n_2) \ldots$ α-kvantil pro F.

1.7. **Prostor $L_2(\Omega, \mathcal{A}, P)$.**

$L_2(\Omega, \mathcal{A}, P)$ definujeme jako množinu všech (komplexních) náhodných veličin nad týmž pravděpodobnostním prostorem $(\Omega, \mathcal{A}, P)$, které mají konečné druhé momenty (resp. rozptyly - viz dále 1.11), tj. $L_2(\Omega, \mathcal{A}, P) := \{X \mid X \text{ náhodná veličina nad } (\Omega, \mathcal{A}, P),\ \mathrm{E}|X|^2 < \infty\}$.

Poznamenejme, že do tohoto prostoru zahrnujeme také všechny konstanty z $\mathbb{C}$, které považujeme za náhodné veličiny s nulovým rozptylem.

Věta 1.8. *$L_2(\Omega, \mathcal{A}, P)$ je* **Hilbertův prostor** *se skalárním součinem $\langle X, Y\rangle = \mathrm{E}X\overline{Y}$ a normou $\|X\|_2 = \sqrt{\langle X, X\rangle} = \sqrt{\mathrm{E}|X|^2}$.*

Důkaz. $L_2(\Omega, \mathcal{A}, P)$ je obdobou funkcionálního prostoru $L_2(\mathcal{J})$ tvořeného funkcemi absolutně integrovatelnými v kvadrátu na intervalu $\mathcal{J} \subseteq \mathbb{R}$. Totiž $\mathrm{E}|X|^2 = \int_\Omega |X(\omega)|^2\, dP(\omega)$, takže namísto s Lebesgueovým integrálem pracujeme s obecnějším pojetím integrálu, kde Lebesgueova míra je nahrazena pravděpodobnostní mírou P:

- **Skalární součin $\langle X, Y\rangle$ existuje a je konečný** pro každé $X, Y \in L_2(\Omega, \mathcal{A}, P)$, jak snadno nahlédneme z nerovnosti
$$4|X\overline{Y}| = (|X| + |\overline{Y}|)^2 - (|X| - |\overline{Y}|)^2 \le (|X| + |\overline{Y}|)^2 + (|X| - |\overline{Y}|)^2 = 2(|X|^2 + |\overline{Y}|^2),$$
odkud užitím $|\overline{Y}| = |Y|$ dostáváme
$$|X\overline{Y}| \le \frac{1}{2}(|X|^2 + |Y|^2),$$
takže
$$|\mathrm{E}X\overline{Y}| \le \int_\Omega |X(\omega)\overline{Y(\omega)}|\, dP(\omega) \le \frac{1}{2}\left[\int_\Omega |X(\omega)|^2\, dP(\omega) + \int_\Omega |Y(\omega)|^2\, dP(\omega)\right] < \infty.$$
- $L_2(\Omega, \mathcal{A}, P)$ **je vektorovým prostorem**. Je uzavřený na násobení skaláry $c \in \mathbb{C}$, neboť $\mathrm{E}|cX|^2 = |c|^2\mathrm{E}|X|^2 < \infty$. Uzavřenost vzhledem ke sčítání plyne z:
$$|X + Y|^2 \le (|X| + |Y|)^2 = |X|^2 + 2|XY| + |Y|^2 \Rightarrow \mathrm{E}|X + Y|^2 \le \mathrm{E}|X|^2 + 2\mathrm{E}|XY| + \mathrm{E}|Y|^2 < \infty.$$
- **Ověření, že $L_2(\Omega, \mathcal{A}, P)$ je úplný, neboli Hilbertův prostor**, je složitější, ale provádí se opět zcela analogicky jako v případě funkcionálního prostoru $L_2(\mathcal{J})$. Podrobnosti lze nalézt například v monografii [1, §2.10].

□

Důsledek 1.9 (Schwarzova nerovnost)**.**
$$|\mathrm{E}XY|,\ |\mathrm{E}X\overline{Y}| \le \|X\|_2\|Y\|_2 = \sqrt{\mathrm{E}|X|^2}\sqrt{\mathrm{E}|Y|^2}, \quad X, Y \in L_2(\Omega, \mathcal{A}, P).$$

Důsledek 1.10. *$X \in L_2(\Omega, \mathcal{A}, P) \ \Rightarrow\ \mathrm{E}X$ existuje.*

Důkaz.
$$|\mathrm{E}X| = |E(1.X)| \le \underbrace{\sqrt{\mathrm{E}|1|^2}}_{1}\sqrt{\mathrm{E}|X|^2} = \sqrt{\mathrm{E}|X|^2} < \infty.$$

□

Důsledek 1.11.
$$X, Y \in L_2(\Omega, \mathcal{A}, P) \ \Rightarrow\ X - \mathrm{E}X,\ Y - \mathrm{E}Y \in L_2(\Omega, \mathcal{A}, P)$$
a
$$\langle X - \mathrm{E}X,\ Y - \mathrm{E}Y\rangle = \mathrm{E}(X - \mathrm{E}X)\overline{(Y - \mathrm{E}Y)} = \mathrm{cov}(X, Y)$$
existuje a splňuje Schwarzovu nerovnost
$$|\mathrm{cov}(X, Y)| \le \sqrt{\mathrm{E}|X - \mathrm{E}X|^2}\sqrt{\mathrm{E}|Y - \mathrm{E}Y|^2} = \sigma_X\sigma_Y.$$

Důsledek 1.12.
$$\rho(X, Y) = \begin{cases} \frac{\mathrm{cov}(X,Y)}{\sigma_X\sigma_Y} & \text{pro} \quad \sigma_x\sigma_y \ne 0 \\ 0 & \text{pro} \quad \sigma_x\sigma_y = 0 \end{cases}$$
je tzv. **korelační koeficient náhodných veličin** X **a** Y, *pro nějž platí $|\rho(X, Y)| \le 1$ a speciálně $-1 \le \rho(X, Y) \le 1$ v případě reálných náhodných veličin X a Y.*

Poznámka . Náhodné veličiny $X, Y \in L_2(\Omega, \mathcal{A}, P)$ se nazývají **nekorelované**, jestliže $\rho(X,Y) = 0$. Vzhledem k 1.11 je nekorelovanost ekvivalentní s $\operatorname{cov}(X,Y) = 0$, tj. s **ortogonalitou** centrovaných veličin $X - \mathrm{E}X$ a $Y - \mathrm{E}Y$ v $L_2(\Omega, \mathcal{A}, P)$.
$\rho(\mathbb{X}, \mathbb{Y}) = [\rho(X_i, Y_j)]_{i,j}$ je tzv. **vzájemná korelační matice náhodných vektorů** $\mathbb{X}$ **a** $\mathbb{Y}$.
$\rho(\mathbb{X}, \mathbb{X}) = [\rho(X_i, X_j)]_{i,j}$ je tzv. **korelační matice náhodného vektoru** $\mathbb{X}$.

POZOR! Nekorelovanost indikuje neexistenci **stochastické závislosti pouze lineárního typu**. Tedy platí
X, Y stochasticky nezávislé $\Rightarrow$ X, Y nekorelované, avšak nikoliv naopak:
X, Y nekorelované $\not\Rightarrow$ X, Y stochasticky nezávislé.