

# Obsah

<b>1</b>	<b>Neurčitost a citlivost</b>	<b>2</b>
1.1	Analýza neurčitosti . . . . .	2
1.1.1	Pravděpodobnostní analýza neurčitosti . . . . .	4
1.1.2	Bayesova metoda . . . . .	5
1.2	Zobrazovací metody a lokální analýza citlivosti . . . . .	6
1.2.1	Metoda konečných diferencí . . . . .	7
1.2.2	Přímá metoda . . . . .	7
1.2.3	Lokální citlivosti a analýza neurčitosti . . . . .	8
1.3	Globální analýza citlivosti . . . . .	9
1.3.1	Standardizované regresní koeficienty . . . . .	9
1.3.2	Spearmanův koeficient pořadové korelace . . . . .	10
1.3.3	Metody založené na rozptylu . . . . .	11
1.4	Ilustrační příklady . . . . .	18
1.4.1	Příklad 1 . . . . .	18
1.4.2	Příklad 2 . . . . .	20
1.4.3	Příklad 3 . . . . .	23
1.4.4	Příklad 4 . . . . .	25
1.5	Příklady k procvičení . . . . .	30

# 1 Neurčitost a citlivost

## 1.1 Analýza neurčitosti

S výpočetní vědou je neodmyslitelně spjatá vlastnost zvaná neurčitost. Jakékoliv řešení reálného problému z praxe v sobě zahrnuje neurčitosti a není možné je odstranit<sup>1</sup>. To je na jednu stranu trochu frustrující, na druhou stranu nám nezbývá, než se s tím smířit a „minimalizovat škody“. Každé řešení problému, které vede na nějaké „důležité“ rozhodnutí musí sestávat z analýzy neurčitosti.

V knize [5] jsme neurčitosti rozdělili do tří hlavních skupin na neurčitosti v matematickém popisu, neurčitosti v datech a neurčitosti v aplikaci modelu. V literatuře je možné najít různá dělení neurčitostí, vždy se však týkají tří zmíněných oblastí. Každý popis reality je nepřesný, podobně nepřesná jsou data (závislá jednak na přesnosti měření, ale také na jejich „reprezentativnosti“ – viz dále) a zejména v případě vědeckých výpočtů se neobejdeme bez použití ICT (a případných dalších nepřesných metod hledání řešení problému).

Pro lepší pochopení a uvědomění si důležitosti analýzy neurčitosti uvedme tři jednoduché příklady.

**Příklad 1.1:** Na vymezeném území žije populace nějakých živých organismů. Určete, jak se bude vyvíjet její velikost v následujícím období (dnech, měsících, ...), jestliže počáteční velikost populace v čase  $t = 0$  je  $N(0) = 100$  a vnitřní koeficient růstu populace je  $r = 0,1$ .

*Řešení:* V knize [5] jsme uvedli několik matematických modelů růstu populace živých organismů. Předpokládejme, že naše populace splňuje kvantovací a vzorkovací podmínku (viz [5]). Jedním z možných popisů jejího růstu je rovnice

$$\frac{d}{dt}N(t) = r \cdot N(t), \quad (1.1)$$

kde  $N(t)$  je velikost populace v čase  $t$ . Řešení rovnice (1.1) má tvar:

$$N(t) = N(0) \cdot e^{r \cdot t}.$$

Pokud nemáme k dispozici žádnou další informaci o tom, na čem závisí růst populace (a jak), odpovědí na zadaný problém je, že se velikost populace bude v čase zvětšovat podle vztahu

$$N(t) = 100 \cdot e^{0,1 \cdot t}.$$

Pokud ovšem zjistíme, že na vymezeném území je omezená kapacita prostředí hodnoty  $K$ , je třeba pro predikci růstu populace použít jiný model, např.

$$\frac{d}{dt}N(t) = r \cdot N(t) \cdot \left(1 - \frac{N(t)}{K}\right). \quad (1.2)$$

---

<sup>1</sup>Je však většinou možné je snížit.

V takovém případě by už však naše odpověď na zadaný problém byla, že velikost populace se „po určitém čase“ ustálí na hodnotě  $K$ . K této hodnotě se bude velikost populace přibližovat buď neustálým růstem, či neustálým poklesem v závislosti na hodnotě  $K$  v porovnání s hodnotou  $N(0)$  (viz [5]).

Další změnu přístupu k modelování růstu populace může přinést doplňující informace například o výskytu nějakého predátora na vymezeném území či jiných vlivů ovlivňujících počet jedinců sledované populace. To vše zahrnujeme do neurčitosti v matematickém popisu modelu.

**Příklad 1.2:** Uvažujme opět, že na vymezeném území žije populace nějakých živých organismů. V určitý den byla změřena její velikost, která činila  $N(0) = 100$  jedinců. Nechť dále víme, že prostředí, v němž populace žije, užíví (dlouhodobě) maximálně 500 jedinců (tj. kapacita prostředí  $K = 500$ ) a že pro vnitřní koeficient růstu populace  $r$  platí  $r \in [-0, 1; 0, 1]$ . Určete vývoj velikosti populace v následujícím období (dnech, měsících).

*Řešení:* Předpokládejme nyní, že růst sledované populace odpovídá rovnici (1.2). V tomto případě se nám však objevuje neurčitost v hodnotě vnitřního koeficientu růstu populace  $r$ , která způsobuje, že odpověď je opět velmi neurčitá. Velikost populace se bude buď postupně zvyšovat až k hodnotě  $K = 500$  (pro  $r > 0$ ) nebo naopak klesat až po úplné vymření populace (pro  $r < 0$ ), či může zůstat konstantní v závislosti na hodnotě (pro  $r = 0$ ).

**Příklad 1.3:** Nalezněte pevný bod funkce  $f(x) = 11 \cdot x - 2$ .

*Řešení:* Nejprve si připomeňme, že pevným bodem  $x$  funkce  $f : M \rightarrow M$  je prvek  $x \in M$ , pro nějž platí  $x = f(x)$ . Nejjednodušší způsob nalezení správné odpovědi je zřejmě vyřešení rovnice

$$x = 11 \cdot x - 2, \tag{1.3}$$

kdy okamžitě získáme  $x = 0, 2$ . Zkusme však hledat řešení metodou prosté iterace (viz [4]) pomocí ICT s počáteční iterací rovnou správnému řešení, tj.  $x_1 = 0, 2^2$ . V systémech Maple a Matlab zkusme vypočítat několik prvních iterací. Nehledě na jejich počet bychom očekávali, že každá z iterací by měla mít hodnotu 0, 2.

V systému Maple bychom pro výpočet a výpis prvních 15 iterací v prostředí *Standard Worksheet* mohli použít následující kód:

```
x[1]:=0.2
for i from 2 to 16 do x[i] := 11*x[i-1]-2 end do
```

V Matlabu bychom pro totéž použili kód:

```
x(1)=0.2;
for i=2:16 x(i) = 11*x(i-1)-2; end
x
```

Zatímco v systému Maple dostaneme všechny iterace rovné správné hodnotě 0, 2, v Matlabu se dvanáctá iterace (při standardním zobrazení na 4 desetinná místa) již liší od správné

---

<sup>2</sup>Počáteční iteraci značíme jako  $x_1$  kvůli Matlabu, v němž je pole indexováno od čísla 1.

hodnoty a tento rozdíl se s dalšími iteracemi stále zvyšuje. Například dvacátá iterace (pokud bychom si ji vypočetli) již má hodnotu řádově  $10^4!$  Když si nastavíme příkazem `format long` zobrazení čísel s vyšší přesností, zjistíme, že již první počítané iterace nebyly přesně rovné hodnotě 0, 2. Na vině je zobrazení čísel v počítači a fakt, že pevný bod funkce  $f$  je odpuzujícím pevným bodem (viz [4]).

Čísla jsou v počítači ukládána v binární soustavě, v níž má většina desetinných čísel nekonečný rozvoj. Již při uložení čísla tak ztrácíme přesnost. Relativně špatná podmíněnost úlohy a odpuzující vlastnost pevného bodu tuto nepřesnost při každé další iteraci zvětšují. V systému Maple se chyba neprojevuje, neboť ten má implementováno ukládání čísel v desítkové soustavě, takže číslo 0, 2 je uloženo přesně.

Předchozí tři příklady byly vybrány záměrně. V praxi nemusí (ale samozřejmě mohou) neurčitosti způsobovat takové rozdíly (neurčitosti, chyby) v řešení zadaného problému, navíc příklad 1.3 je trochu „umělý“. Ovšem i v praxi se může stát, že řešení problému povede na nějakou iterativní metodu, či že bude provedeno několik výpočtů, které zpočátku zanedbatelnou chybu mnohonásobí. Z uvedeného vyplývá, že analýza neurčitosti je velmi důležitá k ověření „správnosti“ (resp. „věrohodnosti“) nalezeného řešení.

Neurčitosti můžeme analyzovat postupně po skupinách, jak jsme si je dříve rozdělili. Neurčitost v matematickém popisu souvisí s nedostatkem informací (resp. provedenými předpoklady) při sestavování modelu. Jejich analýza pak sestává z porovnání různých struktur matematického modelu s měřenými daty. Jedná se o nejsložitější část celé analýzy.

Datová neurčitost zahrnuje nepřesnost měření parametrů vystupujících v modelu (způsobenou jak chybou měřicího přístroje, tak lidským faktorem) a nepřesnost vzniklou omezeným množstvím vzorků ze sledovaného systému. V závislosti na informacích o hodnotách parametrů reprezentujeme datovou neurčitost nejčastěji za pomoci intervalové aritmetiky, fuzzy teorie, či pravděpodobnostní analýzy [5].

Poslední typ neurčitosti je předmětem numerické analýzy. Zkoumá se, jaký vliv na hodnotu řešení má použití ICT (uložení čísel v počítači/softwareovém systému) a metod hledajících přibližné řešení. Více k této problematice je možné najít např. v [4].

### 1.1.1 Pravděpodobnostní analýza neurčitosti

Parametry matematického modelu obsahují neurčitost, kterou je sice většinou možné popsat intervalem, typicky má však parametr nějakou střední hodnotu, jež je v jistém smyslu „nejpravděpodobnější“. Snahou tak je chápat parametr jako náhodnou veličinu s příslušným rozdělením pravděpodobnosti. Tzv. propagace nečitostí modelem (viz [5]) je pak zpravidla<sup>3</sup> prováděna metodou Monte Carlo, při níž se vygeneruje „dostatečně mnoho“ realizací náhodných veličin reprezentujících parametry modelu, které poslouží k různým vyhodnocením modelu. Získáme tak pravděpodobnostní rozdělení hodnot řešení.

Použití metody Monte Carlo „přináší“ další neurčitosti, neboť se jedná o přibližnou metodu. Se vzrůstajícím počtem generovaných realizací náhodných veličin a vyhodnocení modelu konverguje získané rozdělení hodnot ke skutečnému rozdělení hodnot řešení modelu. Přesnost získané aproximace pravděpodobnostního rozdělení testujeme porovnáním základních statistických parametrů (střední hodnota, rozptyl) u dvou různých skupin vygenerovaných řešení nebo použitím odhadu chyby Monte Carlo metody.

Uvažujme model tvaru

$$Y = f(X_1, \dots, X_m),$$

---

<sup>3</sup>U (velmi) jednoduchých modelů, případně u výpočetně náročných, tomu tak být nemusí.

kde  $Y$  je hledané řešení,  $X_1, \dots, X_m$  ( $m \in \mathbb{N}$ ) jsou parametry modelu reprezentované náhodnými veličinami s pravděpodobnostními rozděleními  $p_1(x_1), \dots, p_m(x_m)$ . Označme  $\mathbf{X} = (X_1, \dots, X_m)$ ,  $\mathbf{x} = (x_1, \dots, x_m)$  a  $p(\mathbf{x})$  sdruženou hustotu pravděpodobnosti náhodného vektoru  $\mathbf{X}$ . Nechť dále  $M$  je počet vygenerovaných realizací náhodného vektoru  $\mathbf{X}$ ,  $\mathbf{X}_{(k)}$  jeho  $k$ -tá realizace. Pak pro odhad střední hodnoty funkce  $f(\mathbf{X})$  máme:

$$\widehat{E}(f(\mathbf{X})) = \frac{1}{M} \cdot \sum_{k=1}^M f(\mathbf{X}_{(k)}). \quad (1.4)$$

Tento odhad konverguje (podle silného zákona velkých čísel) ke skutečné střední hodnotě pro  $M \rightarrow \infty$ , jestliže

$$\int |f(\mathbf{x})| \cdot p(\mathbf{x}) \, d\mathbf{x} < \infty.$$

Rozptyl  $v_M$  takto určených odhadů středních hodnot pro dané  $M$  zapíšeme jako:

$$v_M = V \left( \frac{1}{M} \cdot \sum_{k=1}^M f(\mathbf{X}_{(k)}) \right) = \frac{1}{M} \cdot V(f(\mathbf{X})),$$

kde jsme využili vlastnosti rozptylu, že pro nezávislé náhodné veličiny  $Z_1, Z_2$  a konstanty  $a, b$  platí:

$$V(a \cdot Z_1 + b \cdot Z_2) = a^2 \cdot V(Z_1) + b^2 \cdot V(Z_2),$$

a současně předpokladu

$$\int (f(\mathbf{x}))^2 \cdot p(\mathbf{x}) \, d\mathbf{x} < \infty.$$

Hodnotu  $\sqrt{v_M}$  nazýváme chybou Monte Carlo metody (viz [8]). Aplikací nevychýleného odhadu pro rozptyl  $V(f(\mathbf{X}))$  můžeme  $v_M$  odhadnout vztahem

$$\widehat{v}_M = \frac{1}{M \cdot (M-1)} \cdot \sum_{k=1}^M \left( f(\mathbf{X}_{(k)}) - \widehat{E}(f(\mathbf{X})) \right)^2, \quad (1.5)$$

případně

$$\widehat{v}_M = \frac{1}{M \cdot (M-1)} \cdot \left( \sum_{k=1}^M f^2(\mathbf{X}_{(k)}) - M \cdot \left( \widehat{E}(f(\mathbf{X})) \right)^2 \right). \quad (1.6)$$

## 1.1.2 Bayesova metoda

Bayesova metoda nám umožňuje na základě apriorní informace o hodnotě parametru (resp. řešení modelu) a experimentálních výsledků vyvodit posteriorní informaci o parametru (resp. řešení modelu).

Mějme opět model tvaru

$$Y = f(X_1, \dots, X_m)$$

se stejným označením jako dříve. Dále označme  $O_j$  hodnotu  $j$ -tého ( $j = 1, \dots, n$ , kde  $n \in \mathbb{N}$ ) pozorování (experimentálního výsledku). Monte Carlo simulací jsme získali pravděpodobnostní rozdělení hodnot  $Y$  skládajícího se z celkem  $M$  vygenerovaných (a následně vyhodnocených) řešení, které označíme  $Y_{(k)}$  pro  $k = 1, \dots, M$ . Pravděpodobnost  $p(Y_{(k)})$ , že náhodně

zvolené řešení je rovno  $Y_{(k)}$  pro konkrétní  $k$ , je  $\frac{1}{M}$ . Dále definujme pravděpodobnost  $p(O|Y_{(k)})$  vyjadřující, že pozorování  $\mathbf{O}$  ( $\mathbf{O} = (O_1, \dots, O_n)$ ) odpovídá predikci modelu  $Y_{(k)}$ , opět pro konkrétní  $k$ . Za předpokladu, že chyby pozorování jsou nezávislé a normálně rozdělené s nulovou střední hodnotou a rozptylem  $\sigma^2$ , zmíněnou pravděpodobnost spočítáme podle vztahu<sup>4</sup>

$$p(\mathbf{O}|Y_{(k)}) = \prod_{j=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(O_j - Y_{(k)})^2}{2 \cdot \sigma^2}\right). \quad (1.7)$$

Užitím Bayesova vzorce můžeme definovat tzv. posteriorní pravděpodobnost  $p(Y_{(k)}|\mathbf{O})$  vygenerované hodnoty řešení  $Y_{(k)}$ :

$$p(Y_{(k)}|\mathbf{O}) = \frac{p(\mathbf{O}|Y_{(k)}) \cdot p(Y_{(k)})}{\sum_{i=1}^M p(\mathbf{O}|Y_{(i)}) \cdot p(Y_{(i)})}, \quad (1.8)$$

a získat tak nové rozložení hodnot řešení modelu  $Y'$  více odpovídající skutečnějším pozorováním se střední hodnotou:

$$E(Y') = \sum_{i=1}^M p(Y_{(i)}|\mathbf{O}) \cdot Y_{(i)}$$

a rozptylem

$$V(Y') = \sum_{i=1}^M p(Y_{(i)}|\mathbf{O}) \cdot (Y_{(i)} - E(Y'))^2.$$

Zcela analogicky získáváme též nová rozdělení pravděpodobnosti pro parametry  $X_1, \dots, X_m$ . Jejich střední hodnoty a rozptyly počítáme dle stejných vztahů (pouhou záměnou veličiny  $Y$  za zvolený parametr).

## 1.2 Zobrazovací metody a lokální analýza citlivosti

Analýzu neurčitosti obvykle doplňuje analýza citlivosti. Díky té se dozvíme, jak které parametry ovlivňují řešení modelu. Metody analýzy citlivosti můžeme rozdělit do tří skupin na metody *zobrazovací*, *lokální* a *globální*.

Účelem zobrazovacích metod je identifikace nejvýznamnějších parametrů, které nejvíce ovlivňují řešení modelu. Zobrazovací metody bývají tzv. „one-at-a-time“, což znamená, že v jednom kroku zkoumají vždy vliv jednoho parametru na řešení modelu, zatímco ostatní parametry nabývají svých středních hodnot (a jsou pro tento krok konstantní). Používají se především při analýze citlivosti modelů s velkým počtem parametrů (až stovky). Získaný výsledek je kvalitativní. Zjistíme, které parametry jsou citlivější (tj. ovlivňují více řešení modelu) než jiné, nedozvíme se však, o kolik. Výhodou je nižší výpočetní náročnost (oproti ostatním metodám analýzy citlivosti), nevýhodou pak zmíněná „nedostatečná“ informace o citlivosti parametrů. Více o zobrazovacích metodách analýzy citlivosti lze najít např. v [13].

Lokální analýza citlivosti byla již stručně představena v [5]. Deterministický matematický model bývá obvykle zadán systémem algebraických rovnic

$$0 = \mathbf{f}(\mathbf{y}(\mathbf{x}), \mathbf{x}), \quad (1.9)$$

---

<sup>4</sup> $\exp(z)=e^z$

kde  $\mathbf{y}(\mathbf{x}) = (y_1(\mathbf{x}), y_2(\mathbf{x}), \dots, y_n(\mathbf{x}))^T$  je  $n$ -rozměrný vektor hledaného řešení (závisle proměnných),  $\mathbf{x} = (x_1, x_2, \dots, x_m)^T$  je  $m$ -rozměrný vektor parametrů modelu a  $\mathbf{f} = (f_1, f_2, \dots, f_k)^T$  vektorová funkce; nebo systémem diferenciálních rovnic, jako je např. počáteční problém:

$$\frac{d\mathbf{y}(t, \mathbf{x})}{dt} = \mathbf{f}(\mathbf{y}(t, \mathbf{x}), \mathbf{x}), \quad \mathbf{y}(0) = \mathbf{y}_0, \quad (1.10)$$

kde  $\mathbf{y}(t, \mathbf{x})$  je hledané řešení,  $t$  čas,  $\mathbf{x}$  je  $m$ -rozměrný vektor parametrů modelu,  $\mathbf{f}$  je  $k$ -rozměrná vektorová funkce a  $\mathbf{y}_0$  vektor počátečních podmínek. Jelikož je rovnice (1.9) speciálním případem rovnice (1.10) pro konstantní  $t$ , budeme nadále uvažovat pouze rovnici (1.10). Analogicky k [5] definujeme *lokální citlivosti prvního řádu* výrazem

$$\frac{\partial y_i(t, \mathbf{x})}{\partial x_j} \quad \text{pro } i = 1, \dots, n, \quad j = 1, \dots, m$$

a *lokální citlivosti druhého řádu* výrazem

$$\frac{\partial^2 y_i(t, \mathbf{x})}{\partial x_j \partial x_l} \quad \text{pro } i = 1, \dots, n, \quad j, l = 1, \dots, m.$$

Nakonec definujeme *matici citlivostí*  $S$ , která je tvořena lokálními citlivostmi prvního řádu, tedy

$$S = \frac{\partial y_i(t, \mathbf{x})}{\partial x_j} \quad \text{pro } i = 1, \dots, n, \quad j = 1, \dots, m.$$

Pokud nejsme schopni určit popsané citlivosti analyticky pomocí uvedených vztahů, používáme některou z numerických metod. Uvedme pro ilustraci dvě nejjednodušší metody, další je možné nalézt v [13].

### 1.2.1 Metoda konečných diferencí

Nejjednodušší metoda analýzy lokální citlivosti spočívá v nahrazení derivace diferencí.

$$\frac{\partial y(t, \mathbf{x})}{\partial x_j} \approx \frac{y(t, x_1, \dots, x_j + \Delta x_j, \dots, x_m) - y(t, \mathbf{x})}{\Delta x_j}, \quad j = 1, \dots, m. \quad (1.11)$$

Přesnost této metody závisí na velikosti difference  $\Delta x_j$ . Čím větší hodnota, tím méně je výsledek přesný zejména u nelineárních modelů. Pokud je difference příliš malá, zvyšuje se vliv zaokrouhlovací chyby počítače. V praxi se často volí difference rovná 1 % velikosti příslušného parametru, nicméně nalezení optimální (resp. akceptovatelné) difference je zpravidla otázkou metody pokus-omyl [13].

### 1.2.2 Přímá metoda

Jestliže zderivujeme rovnici (1.10) podle parametru  $x_j$  dostaneme (za použití pravidla o derivaci složené funkce) systém

$$\frac{d}{dt} \frac{\partial \mathbf{y}}{\partial x_j} = \frac{\partial \mathbf{f}}{\partial \mathbf{y}} \cdot \frac{\partial \mathbf{y}}{\partial x_j} + \frac{\partial \mathbf{f}}{\partial x_j}, \quad \frac{\partial}{\partial x_j} \mathbf{y}(0) = S_0. \quad (1.12)$$

Systém (1.12) můžeme dále přepsat do maticové formy

$$\frac{d}{dt}S = J \cdot S + F, \quad S(0) = S_0, \quad (1.13)$$

kde  $S_0$  je nulová matice,  $J$  a  $F$  matice zadané vztahy

$$J = \frac{\partial f_l}{\partial y_i}, \quad F = \frac{\partial f_l}{\partial x_j}, \quad \text{pro } l = 1, \dots, k, \quad i = 1, \dots, n, \quad j = 1, \dots, m.$$

Přímé metody jsou založeny na řešení rovnice (1.12). Pro numerické řešení je třeba znát matice  $J$  a  $F$ , a proto se rovnice (1.12) musí řešit společně s rovnicí (1.10). Pro asymptoticky stabilní stacionární bod nebo v případě řešení systému (1.9), můžeme rovnici (1.13) transformovat na tvar (viz [13]):

$$S = -J^{-1} \cdot F. \quad (1.14)$$

### 1.2.3 Lokální citlivosti a analýza neurčitosti

Závěrem části o lokálních metodách analýzy citlivosti si ještě ukažme jejich možné využití při analýze neurčitosti. Lokální citlivosti lze totiž použít k lineárnímu odhadu rozptylu hledaného řešení a vlivu jednotlivých parametrů na tento rozptyl. Rozptyl řešení  $y_i(t, \mathbf{x})$  způsobený rozptylem parametru  $x_j$  můžeme odhadnout vztahem

$$\sigma_j^2(y_i) = \left( \frac{\partial y_i}{\partial x_j} \cdot \sigma(x_j) \right)^2. \quad (1.15)$$

Za předpokladu, že mezi parametry  $x_1, \dots, x_m$  nedochází k interakcím, pak můžeme celkový rozptyl řešení  $y_i$  vyjádřit jako součet právě vypočítaných rozptylů:

$$\sigma^2(y_i) = \sum_{j=1}^m \sigma_j^2(y_i). \quad (1.16)$$

Z rovnic (1.15) a (1.16) je následně možné určit příspěvek parametru  $x_j$  do celkového rozptylu řešení  $y_i$  způsobeného všemi parametry zlomkem

$$S_{ij} = \frac{\sigma_j^2(y_i)}{\sigma^2(y_i)}, \quad (1.17)$$

případně v procentech

$$S\%_{ij} = S_{ij} \cdot 100 \%. \quad (1.18)$$

V následující části se budeme věnovat metodám globální analýzy citlivosti, které určí vliv parametrů na rozptyl v hledaném řešení (mnohem) přesněji. Přesto může být výhodné v některých případech použít představené lokální metody k získu alespoň řádové přesnosti (a tedy prvotní představy).



## 1.3 Globální analýza citlivosti

Analýza lokální citlivosti zkoumá lokální efekt změny parametrů (kolem jednoho bodu v  $m$ -rozměrném prostoru parametrů) na řešení modelu. Pokud však parametry nabývají širšího spektra hodnot, využívají se spíše metody globální analýzy citlivosti.

Vraťme se pro jednoduchost k modelu popsaném rovnicí

$$Y = f(X_1, \dots, X_m), \quad (1.19)$$

kde  $Y$  je hledané řešení a  $X_1, \dots, X_m$  ( $m \in \mathbb{N}$ ) parametry modelu reprezentované náhodnými veličinami s pravděpodobnostními rozděleními  $p_1(x_1), \dots, p_m(x_m)$ . Metodou Monte Carlo můžeme získat odhad pravděpodobnostního rozdělení  $p_Y(y)$  pro řešení  $Y$  a s ním i odhady střední hodnoty a rozptylu. V této části nás bude zajímat, jak se na zmíněném rozptylu podílejí parametry  $\mathbf{X} = (X_1, \dots, X_m)$ .

Předpokládejme, že jsme provedli Monte Carlo simulaci, a máme tak k dispozici  $M$  realizací pro každou z náhodných veličin  $X_1, \dots, X_m$  a  $Y$  označené  $x_{11}, \dots, x_{1M}, \dots, x_{m1}, \dots, x_{mM}$  a  $y_1, \dots, y_M$ . Nejprve zmíníme dvě hojně používané metody v globální analýze citlivosti, které jsou však závislé na typu modelu.

### 1.3.1 Standardizované regresní koeficienty

Jestliže je uvažovaný model (skoro) lineární, můžeme provést regresní analýzu a popsat jej rovnicí

$$Y = b_0 + \sum_{i=1}^m b_i \cdot X_i + \varepsilon, \quad (1.20)$$

kde  $b_i$  jsou regresní koeficienty a  $\varepsilon$  chyba regresního modelu (rozdíl skutečné hodnoty původního modelu a regresního odhadu). Hodnoty regresních koeficientů získáme např. metodou nejmenších čtverců. Označme:

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_M \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{m1} \\ \vdots & \vdots & & \vdots \\ 1 & x_{1M} & \cdots & x_{mM} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_0 \\ \vdots \\ b_m \end{pmatrix}, \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_M \end{pmatrix}.$$

Minimalizací výrazu

$$S(\mathbf{b}) = (\mathbf{y} - \mathbf{x} \cdot \mathbf{b})^T \cdot (\mathbf{y} - \mathbf{x} \cdot \mathbf{b})$$

dostaneme (viz ODKAZ)

$$\mathbf{b} = (\mathbf{x}^T \cdot \mathbf{x})^{-1} \cdot \mathbf{x}^T \cdot \mathbf{y}. \quad (1.21)$$

Označme dále

$$SS_{tot} = \sum_{k=1}^M (y_k - \bar{y})^2, \quad SS_{reg} = \sum_{k=1}^M (\hat{y}_k - \bar{y})^2, \quad SS_{res} = \sum_{k=1}^M (\hat{y}_k - y_k)^2,$$

kde  $\hat{y}_k$  značí odhad hodnoty  $y_k$  získaný z regresního modelu a  $\bar{y}$  průměr hodnot  $y_k$  (pro všechna  $k = 1, \dots, M$ ). Adekvátnost regresního modelu pak můžeme posoudit z hodnoty koeficientu determinace definovaného vzáthem

$$R^2 = \frac{SS_{reg}}{SS_{tot}}.$$

Koeficient determinace nabývá hodnoty z intervalu  $[0, 1]$  a udává, jaký podíl neurčitosti (rozptylu) se podařilo regresi vysvětlit. Regresní model můžeme algebraicky přeformulovat na tvar

$$\frac{\hat{Y} - \bar{y}}{\hat{s}} = \sum_{i=1}^m \frac{b_i \cdot \hat{s}_i}{\hat{s}} \cdot \frac{X_i - \bar{x}_i}{\hat{s}_i}, \quad (1.22)$$

kde

$$\hat{Y} = b_0 + \sum_{i=1}^m b_i \cdot X_i, \quad \bar{y} = \sum_{k=1}^M \frac{y_k}{M}, \quad \hat{s} = \sqrt{\sum_{k=1}^M \frac{(y_k - \bar{y})^2}{M-1}},$$

$$\bar{x}_i = \sum_{k=1}^M \frac{x_{ik}}{M}, \quad \hat{s}_i = \sqrt{\sum_{k=1}^M \frac{(x_{ik} - \bar{x}_i)^2}{M-1}}.$$

Nyní koeficienty  $\frac{b_i \cdot \hat{s}_i}{\hat{s}}$  za předpokladu, že  $X_i$  (pro  $i = 1, \dots, m$ ) jsou nezávislé, vyjadřují vliv proměnné  $X_i$  na hledané řešení  $Y$ . Nazýváme je *standardizovanými regresními koeficienty* a značíme  $\beta_i$ . Platí, že  $\beta_0 = 0$  a (viz např. [14]):

$$\sum_{i=1}^m (\beta_i)^2 = R^2.$$

Čím větší je hodnota  $R^2$ , tím lépe popisuje regresní model závislost řešení původního modelu na jeho parametrech. K posouzení významnosti regresních koeficientů slouží několik statistických testů (viz [13]), které však vychází z předpokladů, jež v tomto případě (globální analýzy citlivosti deterministických modelů) neplatí. Pro dané nastavení parametrů  $x_1, \dots, x_m$  totiž dostaneme vždy stejné řešení  $y$ .

### 1.3.2 Spearmanův koeficient pořadové korelace

V případě, že je model (1.19) (výrazně) nelineární a koeficient determinace lineárního regresního modelu malý, zavádí se často transformace pořadí. Po vygenerování  $M$  realizací každé z náhodných veličin  $X_1, \dots, X_m, Y$  se přiřadí pořadí realizacím veličiny  $Y$  (od 1 do  $M$  podle pořadí velikosti původní hodnoty ze všech realizací této veličiny) a provede se lineární regrese. Podobně je možné transformovat i náhodné veličiny  $X_1, \dots, X_m$  a provádět lineární regresi až poté.

Další možný postup je přímý výpočet Spearmanových koeficientů pořadové korelace. Jednotlivým vygenerovaným realizacím (u příslušné náhodné veličiny) přiřadíme pořadí podle velikosti (od 1 do  $M$ ). Koeficient pořadové korelace  $r_i$  vyjadřující závislost mezi  $Y$  a  $X_i$  následně vypočítáme podle vztahu

$$r_i = 1 - \frac{6 \cdot \sum_{k=1}^M d_k^2}{M^3 - M}, \quad (1.23)$$

kde  $d_k$  je rozdíl v přiřazeném pořadí  $k$ -té realizace  $X_i$  a  $k$ -té realizace  $Y$ . Koeficient korelace je číslo z intervalu  $[-1, 1]$ . Hodnota blízká jedné značí rostoucí závislost mezi  $X_i$  a  $Y$ , hodnota blízká minus jedné klesající (tj. s rostoucí hodnotou  $X_i$  klesá hodnota  $Y$ ).

Za předpokladu, že je model monotónní a že mezi vstupními proměnnými modelu nedochází k žádným interakcím, lze získat příspěvek  $S_i$  parametru  $X_i$  do celkového rozptylu výstupu modelu  $Y$  jako (viz [9]):

$$S_i = \frac{r_i^2}{\sum_{j=1}^m r_j^2}. \quad (1.24)$$

V případě použití pořadové transformace můžeme k posouzení adekvátnosti modelu a významnosti získaných regresních koeficientů použít opět koeficient determinace a statistické testy významnosti. Při užití přímého výpočtu Spearmanova koeficientu pořadové korelace se též nabízí některé statistické testy významnosti (viz např. [9], [19]).

Nevýhodou transformace podle pořadí je zvýšení relativních vlivů parametrů prvního řádu, tj. dojde k potlačení vlivů způsobených interakcemi, což může vést k přehlédnutí vlivů parametrů, které působí především v interakcích. Tato nevýhoda je tím významnější, čím víc má model parametrů [13].

### 1.3.3 Metody založené na rozptylu

Předchozí dvě metody globální analýzy citlivosti jsou závislé na typu modelu. Následující dvě metody patří do skupiny metod založených na rozkladu rozptylu a na typu modelu nezávisí. Princip spočívá v rozložení rozptylu řešení  $V(Y)$  na rozptyly tohoto řešení způsobené jak jednotlivými parametry, tak jejichmi interakcemi. Vlivem prvního řádu parametru  $X_i$  na řešení  $Y$  rozumíme výraz

$$V_i = V_{X_i}(E_{X_{-i}}(Y|X_i)), \quad (1.25)$$

kde zápis  $X_{-i}$  vyjadřuje všechny parametry  $X_j, j = 1, \dots, m$  kromě  $X_i$ . Upozorníme, že nemůžeme vliv prvního řádu definovat výrazem  $V_{X_{-i}}(Y|X_i)$ , neboť jednak bychom určovali rozptyl řešení při zafixovaném parametru  $X_i$  na jedné hodnotě (ten však nabývá intervalu hodnot) a navíc by se nám mohlo stát, že  $V_{X_{-i}}(Y|X_i) > V(Y)$ . Naproti tomu pro dříve definovaný vliv parametru  $X_i$  výrazem (1.25) platí požadovaná nerovnost  $V_i \leq V(Y)$ . Relativní hodnotu vlivu prvního řádu (v porovnání s celkovým rozptylem řešení) pak nazveme *indexem citlivosti prvního řádu* a označíme  $S_i$ :

$$S_i = \frac{V_{X_i}(E_{X_{-i}}(Y|X_i))}{V(Y)}. \quad (1.26)$$

Za povšimnutí stojí, že v případě lineárního modelu dostáváme:  $S_i = \beta_i^2$ , kde  $\beta_i$  je standardizovaný regresní koeficient (viz část 1.3.1). Vliv druhého řádu parametrů  $X_i$  a  $X_j$  (pro  $i \neq j$ ) na řešení  $Y$  definujeme výrazem

$$V_{ij} = V_{X_{i,j}}(E_{X_{-(i,j)}}(Y|X_i, X_j)) - V_i - V_j, \quad (1.27)$$

kde podobně  $X_{-(i,j)}$  vyjadřuje všechny parametry  $X_k, k = 1, \dots, m$  kromě  $X_i$  a  $X_j$ . Od celkového rozptylu způsobeného parametry  $X_i$  a  $X_j$  odečítáme na pravé straně (1.27) vlivy prvního řádu, abychom dostali pouze vliv způsobený interakcí mezi  $X_i$  a  $X_j$ . Následně můžeme definovat *index citlivosti druhého řádu*  $S_{ij}$ :

$$S_{ij} = \frac{V_{ij}}{V(Y)}. \quad (1.28)$$

Tímto způsobem bychom mohli pokračovat až po vliv (index citlivosti)  $m$ -tého řádu. Celkově jsme tak popsali vlivy parametrů  $X_1, \dots, X_m$  na rozptyl řešení  $Y$ , přičemž platí (viz např. [15]):

$$V(Y) = \sum_{i=1}^m V_i + \sum_{i=1}^{m-1} \sum_{\substack{j=2 \\ i < j}}^m V_{ij} + \dots + V_{12\dots m} \quad (1.29)$$

a tedy i

$$1 = \sum_{i=1}^m S_i + \sum_{i=1}^{m-1} \sum_{\substack{j=2 \\ i < j}}^m S_{ij} + \dots + S_{12\dots m}. \quad (1.30)$$

V praxi většinou nepočítáme všechny indexy citlivosti (všech řádů), neboť je to výpočetně náročné (počet indexů citlivosti je  $2^m - 1$ ) a ani to není nutné. Zpravidla nás zajímá, jak ovlivňuje výstup modelu každý parametr samostatně (tj. index citlivosti prvního řádu) a které parametry jsou nevýznamné (a je případně možné je nahradit konstantou). Nízký index citlivosti prvního řádu ještě nemusí znamenat, že je příslušný parametr nevýznamný. Jeho vliv se totiž může projevit především v interakcích s jinými parametry. Z toho důvodu definujeme *celkový index citlivosti*  $S_{T_i}$  výrazem

$$S_{T_i} = 1 - \frac{V_{X_{-i}}(E_{X_i}(Y|X_{-i}))}{V(Y)} = \frac{E_{X_{-i}}(V_{X_i}(Y|X_{-i}))}{V(Y)}. \quad (1.31)$$

Pro konkrétní představu uvažujme, že  $m = 3$ . Máme tedy model tvaru  $Y = f(X_1, X_2, X_3)$ . Indexy citlivosti prvního řádu jsou  $S_1, S_2, S_3$ , indexy citlivosti druhého řádu  $S_{12}, S_{13}, S_{23}$ , index citlivosti třetího řádu  $S_{123}$  a celkové indexy citlivosti  $S_{T_1} = S_1 + S_{12} + S_{13} + S_{123}$ ,  $S_{T_2} = S_2 + S_{12} + S_{23} + S_{123}$ ,  $S_{T_3} = S_3 + S_{13} + S_{23} + S_{123}$ . Většinou určujeme pouze indexy citlivosti prvního řádu  $S_i$  a celkové indexy citlivosti  $S_{T_i}$ .

## Sobol'ova metoda

Hlavní myšlenkou Sobol'ovy metody je dekompozice (též známá jako FANOVA dekompozice, viz [10]) funkce  $f(X_1, \dots, X_m)$  z rovnice (1.19) na ortogonální sčítance:

$$f(X_1, \dots, X_m) = f_0 + \sum_{i=1}^m f_i(X_i) + \sum_{i=1}^{m-1} \sum_{j=i+1}^m f_{ij}(X_i, X_j) + \dots + f_{1,2,\dots,m}(X_1, \dots, X_m). \quad (1.32)$$

Pro členy rovnice (1.32) přitom platí:

$$\begin{aligned} f_0 &= E(f(\mathbf{X})), & f_i(X_i) &= E_{X_{-i}}(f(\mathbf{X})|X_i) - f_0, \\ f_{ij}(X_i, X_j) &= E_{X_{-(i,j)}}(f(\mathbf{X})|X_i, X_j) - f_i(X_i) - f_j(X_j) - f_0, \quad \dots \end{aligned}$$

Poslední člen  $f_{1,2,\dots,m}(X_1, \dots, X_m)$  rovnice (1.32) se určí přímo z této rovnice. Jelikož jsou členy pravé strany rovnice (1.32) ortogonální, dostáváme po aplikaci rozptylu na levou i pravou stranu rovnici (1.29), viz např. [12], [13].

Potřebné rozptyly  $V(Y), V_i, V_{ij}, \dots$  pak můžeme vypočítat následovně:

$$V(Y) = \int_{\mathbb{R}^m} f^2(\mathbf{x}) \cdot p(\mathbf{x}) \, d\mathbf{x} - f_0^2,$$

$$V_{i_1, \dots, i_s} = \int_{\mathbb{R}^s} f_{i_1, \dots, i_s}^2(x_{i_1}, \dots, x_{i_s}) \cdot p_{i_1, \dots, i_s}(x_{i_1}, \dots, x_{i_s}) dx_{i_1} \dots dx_{i_s},$$

kde  $i_1, \dots, i_s \in \{1, 2, \dots, m\}$ ,  $i_1 < \dots < i_s$  a  $p_{i_1, \dots, i_s}(x_{i_1}, \dots, x_{i_s})$  je sdružená hustota pravděpodobnosti náhodných veličin  $X_{i_1}, \dots, X_{i_s}$ .

Většinou nejsme schopni určit rozptyly (a tedy ani indexy citlivosti) analyticky, a tak je třeba využít následujících numerických odhadů<sup>5</sup>:

$$\hat{f}_0 = \frac{1}{M} \cdot \sum_{k=1}^M f(x_{1k}, \dots, x_{mk}), \quad (1.33)$$

$$\hat{V}(Y) = \frac{1}{M} \cdot \sum_{k=1}^M f^2(x_{1k}, \dots, x_{mk}) - \left(\hat{f}_0\right)^2, \quad (1.34)$$

$$\hat{V}_i = \frac{1}{M} \cdot \sum_{k=1}^M f(x_{-ik}^{(1)}, x_{ik}^{(1)}) \cdot f(x_{-ik}^{(2)}, x_{ik}^{(1)}) - \left(\hat{f}_0\right)^2, \quad (1.35)$$

$$\hat{V}_{-i} = \frac{1}{M} \cdot \sum_{k=1}^M f(x_{-ik}^{(1)}, x_{ik}^{(1)}) \cdot f(x_{-ik}^{(1)}, x_{ik}^{(2)}) - \left(\hat{f}_0\right)^2, \quad (1.36)$$

kde  $x_{-ik} = x_{1k}, \dots, x_{i-1,k}, x_{i+1,k}, \dots, x_{mk}$ . Horní indexy <sup>(1)</sup> a <sup>(2)</sup> v rovnicích (1.35) a (1.36) vyjadřují, že pro získání odhadů  $\hat{V}_i$  a  $\hat{V}_{-i}$  potřebujeme ve skutečnosti  $2 \cdot M$  realizací náhodných veličin  $X_1, \dots, X_m, Y$  rozdělených do dvou skupin po  $M$  hodnotách (pro každou z veličin). Index <sup>(1)</sup> značí, že příslušné realizace vybíráme z první skupiny, index <sup>(2)</sup> představuje druhou skupinu. Analogicky bychom mohli odvodit odhady pro zbylé rozptyly ( $V_{ij}, V_{ijk}, \dots$ ).

Přesnost odhadů (1.33) – (1.36) je možné určit podle Monte Carlo chyby, již jsme definovali dříve (viz (1.5), resp. (1.6)). Na členy v rovnicích (1.33) – (1.36) se totiž můžeme dívat jako na odhady příslušných středních hodnot, tj.:

$$\begin{aligned} \hat{f}_0 &= \hat{E}(f(\mathbf{X})), \\ \hat{V}(Y) + \left(\hat{f}_0\right)^2 &= \hat{E}(f^2(\mathbf{X})), \\ \hat{V}_i + \left(\hat{f}_0\right)^2 &= \hat{E}\left(f\left(X_{-i}^{(1)}, X_i^{(1)}\right) \cdot f\left(X_{-i}^{(2)}, X_i^{(1)}\right)\right), \\ \hat{V}_{-i} + \left(\hat{f}_0\right)^2 &= \hat{E}\left(f\left(X_{-i}^{(1)}, X_i^{(1)}\right) \cdot f\left(X_{-i}^{(1)}, X_i^{(2)}\right)\right), \end{aligned}$$

Odhad chyby celkového rozptylu bychom mohli například vyjádřit jako:

$$\hat{v}_M\left(\hat{V}(Y)\right) = \frac{1}{M \cdot (M-1)} \cdot \left(\sum_{k=1}^M f^4(\mathbf{X}_{(k)}) - M \cdot \left(\hat{E}(f^2(\mathbf{X}))\right)^2\right) + \hat{v}_M\left(\left(\hat{f}_0\right)^2\right). \quad (1.37)$$

Přestože se zvyšujícím se  $M$  jsou zmíněné odhady stále přesnější,  $M$  bude vždy konečné a odhad nepřesný. Uvažujme, že proměnná  $X_i$  vůbec neovlivňuje řešení  $Y = f(\mathbf{X})$ . Jí příslušný index citlivosti tak má být  $S_i = 0$ .

<sup>5</sup>Opět vycházíme z toho, že máme k dispozici  $M$  realizací náhodného vektoru  $\mathbf{X}$  a náhodné veličiny  $Y$  z Monte Carlo simulace.

Podle (1.35):

$$\widehat{V}_i = \frac{1}{M} \cdot \sum_{k=1}^M f(x_{-ik}^{(1)}, x_{ik}^{(1)}) \cdot f(x_{-ik}^{(2)}, x_{ik}^{(1)}) - \left(\widehat{f}_0\right)^2.$$

Tato hodnota však nikdy nebude nižší než

$$\left| \frac{1}{M} \cdot \sum_{k=1}^M f(\mathbf{x}_{(k)}^{(1)}) \cdot f(\mathbf{x}_{(k)}^{(2)}) - \left(\widehat{f}_0\right)^2 \right|. \quad (1.38)$$

Homma a Saltelli (viz [3]) proto doporučují korekční člen (1.38) aplikovat bez absolutní hodnoty<sup>6</sup> na jednotlivé rozptyly  $V_{i_1, \dots, i_s}$  (pro  $i_1, \dots, i_s \in \{1, 2, \dots, m\}, i_1 < \dots < i_s$ ) tak, že od rozptylů prvního řádu ( $V_i$ ) se korekční člen odečte, k rozptylům druhého řádu ( $V_{ij}$ ) se přičte, od rozptylů třetího řádu ( $V_{ijk}$ ) se odečte atd. Vlivy všech řádů jsou tímto nedostatkem přesnosti ovlivněny stejně, takže pokud odečteme korekční člen od  $V_i$  i od  $V_j$ , odečítáme jej od  $V_{ij}$  již dvakrát (viz definice  $V_{ij}$ ), a proto v tomto případě jednou korekční člen přičteme. Z tohoto důvodu se znaménka střídají (při aplikaci korekčního členu na již vypočtené rozptyly  $V_{i_1, \dots, i_s}$ ).

## FAST metoda

Fourierův test citlivosti amplitudy (angl. Fourier Amplitude Sensitivity Test) je další metodou založenou na rozkladu rozptylu. Hlavní myšlenkou metody je transformace  $m$ -rozměrného integrálu v proměnné  $\mathbf{x}$  na jednorozměrný integrál v proměnné  $s$  použitím transformačních funkcí  $G_i$  ( $i = 1, \dots, m$ ), konkrétně

$$X_i = G_i(\sin(\omega_i \cdot S)) \quad \text{nebo} \quad X_i = G_i((\omega_i \cdot S) \bmod 2\pi), \quad (1.39)$$

kde  $S$  je rovnoměrně rozdělená náhodná veličina na intervalu  $(-\pi, \pi)$  a  $\{\omega_i\}$  je množina celočíselných úhlových frekvencí. Rovnici modelu (1.19) pak můžeme přepsat následovně:

$$Y = f(X_1, \dots, X_m) = f(G_1(\omega_1 \cdot S), \dots, G_m(\omega_m \cdot S)) = f(S).$$

Pro vhodně zvolené  $\omega_i$  a  $G_i$  je pak možné aproximovat střední hodnotu řešení  $Y$  jako

$$E(Y) \approx \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) ds. \quad (1.40)$$

Pro rozptyl pak dostáváme:

$$V(Y) \approx \frac{1}{2\pi} \int_{-\pi}^{\pi} f^2(s) ds - (E(Y))^2. \quad (1.41)$$

Předpokládejme nyní, že  $f(s) \in L^2([-\pi, \pi])$ , tj. že platí:

$$\int_{-\pi}^{\pi} |f(s)|^2 ds < \infty.$$

<sup>6</sup>Absolutní hodnota byla přidána, aby příslušná věta měla správný matematický význam. Znaménko členu v absolutní hodnotě je však podstatné!

Pak Fourierova řada

$$\sum_{k=-\infty}^{\infty} c_k \cdot e^{i \cdot k \cdot s} \quad (1.42)$$

konverguje k funkci  $f(s)$  právě tehdy, když platí Parsevalova rovnost (viz ODKAZ)

$$\sum_{k=-\infty}^{\infty} |c_k|^2 = \frac{1}{2\pi} \cdot \int_{-\pi}^{\pi} |f(s)|^2 ds. \quad (1.43)$$

Pro Fourierovy koeficienty přitom platí:

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) \cdot e^{-i \cdot k \cdot s} ds.$$

Dosazením (1.43) do (1.41) a následnou úpravou získáváme:

$$\begin{aligned} V(Y) &\approx \sum_{k=-\infty}^{\infty} |c_k|^2 - \left( \frac{1}{2\pi} \cdot \int_{-\pi}^{\pi} f(s) ds \right)^2 \\ &\approx \sum_{k=-\infty}^{\infty} (a_k^2 + b_k^2) - (a_0^2 + b_0^2) \\ &\approx 2 \cdot \sum_{k=1}^{\infty} (a_k^2 + b_k^2), \end{aligned} \quad (1.44)$$

kde

$$a_k = \frac{1}{2} \cdot (c_k + c_{-k}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) \cdot \cos(k \cdot s) ds, \quad (1.45)$$

$$b_k = \frac{1}{2} \cdot i \cdot (c_k - c_{-k}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) \cdot \sin(k \cdot s) ds. \quad (1.46)$$

Vliv  $V_i$  prvního řádu parametru  $X_i$  na řešení  $Y$  definovaný vztahem (1.25) je složen z komponent o základní frekvenci  $\omega_i$  (přiřazené parametru  $X_i$ ) a jejích vyšších harmonických složek ( $2 \cdot \omega_i, 3 \cdot \omega_i, \dots$ ), tj.

$$V_i \approx 2 \cdot \sum_{p=1}^{\infty} (a_{p \cdot \omega_i}^2 + b_{p \cdot \omega_i}^2),$$

za předpokladu, že  $\omega_i$  jsou celá čísla. Amplitudy jednotlivých složek s rostoucím  $p$  klesají, takže pro odhad vlivu  $V_i$  můžeme použít jen prvních několik členů:

$$\widehat{V}_i = 2 \cdot \sum_{p=1}^N (a_{p \cdot \omega_i}^2 + b_{p \cdot \omega_i}^2), \quad (1.47)$$

kde  $N$  je nejvyšší harmonická složka, již uvažujeme. Obvykle se volí  $N = 4$ , resp.  $N = 6$  (viz [13]). Abychom mohli použít vztah (1.47), měly by být frekvence  $\omega_i$  pro  $i = 1, \dots, m$  tzv. nepoměrné, tj. mělo by platit:

$$\sum_{i=1}^m \gamma_i \cdot \omega_i = 0 \iff \forall i = 1, \dots, m : \gamma_i = 0,$$

přičemž  $\gamma_i \in \mathbb{Z}$  pro všechna  $i = 1, \dots, m$ . Jelikož chceme, aby také  $\omega_i$  byla celá čísla, požadovanou vlastnost nemůžeme splnit. Z toho důvodu volíme  $\omega_i$  alespoň taková, aby tuto vlastnost splňovala pro  $|\gamma_i| \in 1, \dots, N$ .

Volba frekvencí  $\omega_i$  je důležitá také z pohledu „vyplněnosti“ prostoru parametrů  $X_1, \dots, X_m$  pro generované hodnoty veličiny  $S$  z intervalu  $(-\pi, \pi)$ . Ilustrujme to na následujícím příkladu.

**Příklad 1.4:** Uvažujme model tvaru  $Y = f(X_1, X_2)$ , pro nějž  $X_1 \sim U(0, 1)$ ,  $X_2 \sim U(0, 1)$ .<sup>7</sup> Nechtě  $G_1 = \frac{1}{2} + \frac{1}{\pi} \cdot \arcsin(\sin(\omega_1 \cdot s))$ ,  $G_2 = \frac{1}{2} + \frac{1}{\pi} \cdot \arcsin(\sin(\omega_2 \cdot s))$ . Vykreslete graf získaných bodů  $[x_1, x_2]$  po vygenerování  $M = 5000$  realizací veličiny  $S \sim U(-\pi, \pi)$  a aplikaci funkcí  $G_1, G_2$  pro

(a)  $\omega_1 = 3, \omega_2 = 7$ ,

(b)  $\omega_1 = 11, \omega_2 = 21$ ,

(c)  $\omega_1 = \pi, \omega_2 = 21$ .

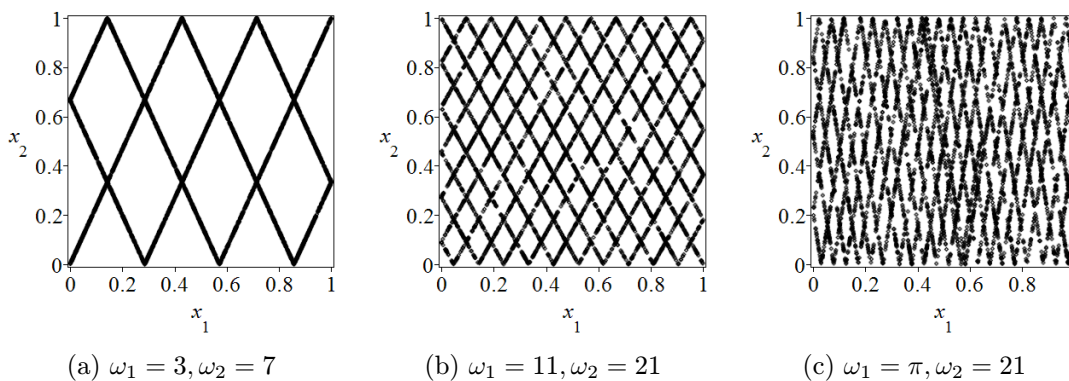
*Řešení:* V systému Maple můžeme k vykreslení grafu v případě (a) použít následující kód (pro zbylé případy se pouze změní proměnné  $\omega_1, \omega_2$ ):

```
M:=5000: omega[1]:=3: omega[2]:=7:
G[1]:= s -> 1/2 + (1/Pi)*arcsin(sin(omega[1]*s)):
G[2]:= s -> 1/2 + (1/Pi)*arcsin(sin(omega[2]*s)):

with(Statistics):
S := RandomVariable(Uniform(-Pi, Pi)):
Ss := Sample(S, M):
for i from 1 to M do
  A[i] := evalf(G[1](Ss[i])); B[i] := evalf(G[2](Ss[i]));
end do:
body := seq([A[i], B[i]], i = 1 .. M):

plot([body], style = point, axes = boxed,
      labels = [typeset(x[1]), typeset(x[2])]);
```

Z obrázku 1.1 vidíme, že v případě (a) je vyplněnost prostoru parametrů velmi „špatná“, což je způsobeno nevhodnou volbou frekvencí  $\omega_1, \omega_2$ . Čím „více“ jsou  $\omega_1, \omega_2$  nepoměrné, tím lepší dostáváme výsledek.



Obrázek 1.1: Vykreslení grafů z příkladu 1.4.

<sup>7</sup> $U(0, 1)$  značí uniformní (tj. rovnoměrné) rozdělení na intervalu  $(0, 1)$ .



Doplňme ještě zdrojový kód pro vykreslení případu (a) v Matlabu:

```
M=5000;
Ss = unifrnd(-pi, pi, [M 1]);
omega1=3; omega2=7;
x1 = .5 + asin(sin(omega1*Ss))/pi;
x2 = .5 + asin(sin(omega2*Ss))/pi;

plot(x1,x2,'.')
xlabel('x[1]')
ylabel('x[2]')
```

Vyplněnost prostoru parametrů (resp. zachování sdruženého pravděpodobnostního rozdělení  $p(\mathbf{x})$  pro parametry  $X_1, \dots, X_m$ ) je též ovlivněna volbou transformačních funkcí  $G_i$ . V literatuře jsou uváděné zejména transformace pro případ, že parametry  $X_1, \dots, X_m$  jsou (rovnoměrně) rozdělené na intervalu  $(0, 1)$ :

$$G_i(\sin(\omega_i \cdot s)) = x_i^* \cdot (1 + \nu_i^* \cdot \sin(\omega_i \cdot s)), \quad (1.48)$$

$$G_i(\sin(\omega_i \cdot s)) = \frac{1}{2} + \frac{1}{\pi} \cdot \arcsin(\sin(\omega_i \cdot s)), \quad (1.49)$$

kde  $x_i^*$  je nominální (střední) hodnota parametru  $X_i$  a  $\nu_i^*$  definuje koncové body intervalu neurčitosti. Pokud potřebujeme vytvořit transformaci  $G_i$  pro rozdělení pravděpodobnosti  $p_i(x)$  parametru  $X_i$ , můžeme využít navrhované rovnice, kterou by taková transformace měla splňovat (viz [13]):

$$\pi \cdot \sqrt{1 - x_i^2} \cdot p_i(G_i) \cdot \frac{dG_i(x_i)}{dx_i} = 1. \quad (1.50)$$

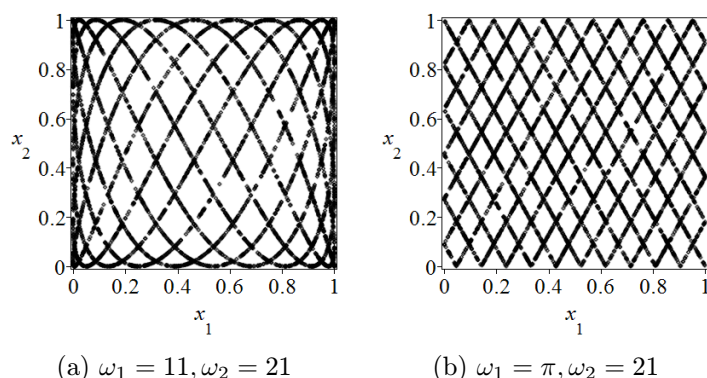
Všimněme si, že transformační funkce (1.49) splňuje rovnici (1.50).

**Příklad 1.5:** Uvažujme opět model tvaru  $Y = f(X_1, X_2)$ , pro nějž  $X_1 \sim U(0, 1)$ ,  $X_2 \sim U(0, 1)$ . Nechtě  $\omega_1 = 11, \omega_2 = 21$ . Vykreslete graf získaných bodů  $[x_1, x_2]$  po vygenerování  $M = 5000$  realizací veličiny  $S \sim U(-\pi, \pi)$  a aplikaci funkcí  $G_1, G_2$  pro

(a)  $G_1(\sin(\omega_1 \cdot s)) = \frac{1}{2} \cdot (1 + \sin(\omega_1 \cdot s)), G_2(\sin(\omega_2 \cdot s)) = \frac{1}{2} \cdot (1 + \sin(\omega_2 \cdot s)),$

(b)  $G_1 = \frac{1}{2} + \frac{1}{\pi} \cdot \arcsin(\sin(\omega_1 \cdot s)), G_2 = \frac{1}{2} + \frac{1}{\pi} \cdot \arcsin(\sin(\omega_2 \cdot s)).$

*Řešení:* Zdrojový kód k vykreslení grafů je analogický příkladu 1.4.



Obrázek 1.2: Vykreslení grafů z příkladu 1.5.

Počet vygenerovaných hodnot parametru  $S$  by měl splňovat Shanonův<sup>8</sup> vzorkovací teorém (viz [18]): „Přesná rekonstrukce spojitého, frekvenčně omezeného, signálu z jeho vzorků je možná tehdy, pokud byl vzorkován frekvencí alespoň dvakrát vyšší, než je maximální frekvence rekonstruovaného signálu.“ V našem případě proto nastavujeme hodnotu  $M$  minimálně na (viz [13]):

$$M = 2 \cdot N \cdot \max_{i=1, \dots, m} \{\omega_i\} + 1.$$

Ukázali jsme, jak s pomocí rozkladu modelové funkce  $f$  do fourierovy řady odhadnout celkový rozptyl funkce  $f$  v závislosti na parametrech  $X_1, \dots, X_m$  a rozptyly způsobené pouze jednotlivými parametry (tj. vlivy prvního řádu). Ještě bychom potřebovali odhadnout celkový vliv parametru  $X_i$  včetně jeho interakcí s ostatními parametry. K tomu využijeme stejné myšlenky jako dříve. Parametru  $X_i$  přiřadíme frekvenci  $\omega_i$  „velmi odlišnou“ od všech ostatních frekvencí příslušných zbylým parametrům, které budou navíc „téměř shodné“. Například bychom tedy volili:  $\omega_i = 20$  a  $\omega_j = 1$  pro  $j \neq i$ .

Jak jsme však viděli dříve, volit stejná (resp. „málo nepoměrná“)  $\omega_i$  pro  $i = 1, \dots, m$  není vhodné kvůli zachování pravděpodobnostního rozdělení parametrů  $X_1, \dots, X_m$ . Zpravidla proto volíme vyšší hodnotu pro  $\omega_i$  a ostatní  $\omega_j$  ( $j \neq i$ ) rozdělíme rovnoměrně v intervalu  $[1, \frac{\omega_i}{2 \cdot N}]$ , viz [13]. Vliv  $V_{-i}$  všech ostatních parametrů pak odhadneme jako:

$$\widehat{V}_{-i} = 2 \cdot \sum_{p=1}^N \sum_{\substack{j=1 \\ j \neq i}}^m \left( a_{p \cdot \omega_j}^2 + b_{p \cdot \omega_j}^2 \right). \quad (1.51)$$

Výhodou FAST metody je, že pro výpočet koeficientů citlivosti prvního řádu stačí jediný soubor vyhodnocení modelu (díky nastavení  $\omega_1, \dots, \omega_m$ ). Tato výhoda se smazává ve chvíli, kdy chceme počítat i celkové indexy citlivosti. Pro ně je totiž třeba vždy přenastavit frekvence  $\omega_1, \dots, \omega_m$  (pro každý parametr  $X_i$ ). Pokud však vytváříme frekvence  $\omega_1, \dots, \omega_m$  podle výše popsaného postupu, lze pro jedno jejich nastavení počítat současně  $S_i$  a  $S_{T_i}$ .

## 1.4 Ilustrační příklady

V této části si ukážeme čtyři jednoduché příklady, na nichž demonstrujeme dříve uváděné metody analýzy neurčitosti a citlivosti.

### 1.4.1 Příklad 1

Uvažujme model tvaru:

$$Y = X_1 + X_2 + X_3, \quad (1.52)$$

kde  $X_1 \sim U(\frac{1}{2}, \frac{3}{2})$ ,  $X_2 \sim U(\frac{3}{2}, \frac{9}{2})$ ,  $X_3 \sim U(\frac{9}{2}, \frac{27}{2})$ .

#### Analýza neurčitosti

V tomto jednoduchém příkladu bychom mohli střední hodnotu, rozptyl i pravděpodobnostní rozdělení veličiny  $Y$  vypočítat přímo na základě znalostí ze základního kurzu teorie

<sup>8</sup>Též známý jako Nyquistův či Kotělnikovův.

pravděpodobnosti. Pokud využijeme nějaký systém počítačové algebry, tj. např. Maple, získáme požadované následovně:

```
with(Statistics):
X[1] := RandomVariable(Uniform(1/2, 3/2)):
X[2] := RandomVariable(Uniform(3/2, 9/2)):
X[3] := RandomVariable(Uniform(9/2, 27/2)):

Y := X[1] + X[2] + X[3]:

Mean(Y);          # stredni hodnota Y
Variance(Y);      # rozptyl Y
PDF(Y,x);         # pravdepodobnostni distribucni funkce Y
```

Dostaneme:

$$E(Y) = 13, \quad V(Y) = \frac{91}{12}.$$

### Lokální analýza citlivosti

Přímo z definice získáváme:

$$S_1 = \frac{\partial Y}{\partial X_1} = 1, \quad S_2 = \frac{\partial Y}{\partial X_2} = 1, \quad S_3 = \frac{\partial Y}{\partial X_3} = 1.$$

Při použití nepřímé metody (viz 1.2.1) je nutné zvolit bod  $[x_1, x_2, x_3]$ , v němž lokální citlivosti numericky aproximujeme. Necht'  $[x_1, x_2, x_3] = [E(X_1), E(X_2), E(X_3)] = [1, 3, 9]$ . Velikost difference zvolíme na obvyklé 1 % velikosti parametru (tj.  $\Delta x_1 = 0.01, \Delta x_2 = 0.03, \Delta x_3 = 0.09$ ) a pro výpočet citlivosti prvního parametru použijeme vztah

$$S_1 = \frac{E(|x_1 \pm \Delta x_1 + x_2 + x_3 - x_1 - x_2 - x_3|)}{x_1 + x_2 + x_3} \cdot 100, \quad (1.53)$$

který jednak vztahuje citlivost na jednotku velikosti parametru (tj. počítá relativní citlivost) a současně uvažuje jak vliv přičtení konečné difference, tak jejího odečtení<sup>9</sup>. Hodnoty  $S_2$  a  $S_3$  se získají analogicky. Celkem dostaneme:

$$S_1 = \frac{E(0,01;0,01)}{13} \cdot 100 = \frac{1}{13}, \quad S_2 = \frac{E(0,03;0,03)}{13} \cdot 100 = \frac{3}{13}, \quad S_3 = \dots = \frac{9}{13}.$$

Vidíme, že získáváme jiné výsledky. Metodou konečných diferencí jsme de facto určili ne citlivost, ale elasticitu. Ta bývá definována vztahem:

$$\frac{\partial Y}{\partial X_i} \cdot \frac{X_i}{Y} \quad \left( \text{resp. } \frac{\partial \ln(Y)}{\partial \ln(X_i)} \right). \quad (1.54)$$

Stejných hodnot jak z definice citlivosti bychom dosáhli, kdybychom uvažovali diferenci rovnou jedné pro všechny tři parametry (tj.  $\Delta x_1 = \Delta x_2 = \Delta x_3 = 1$ ) a nevztahovali výsledek k hodnotě řešení. Je ke zvážení, který přístup (citlivost/elasticita) je vhodnější k posouzení (lokálního) vlivu parametru na řešení.

<sup>9</sup>a tyto dva vlivy zprůměruje (aplikací střední hodnoty)

## Globální analýza citlivosti

Model (1.52) je lineární, takže  $R^2 = 1$ . Standardizované regresní koeficienty můžeme určit z rovnice

$$\frac{Y - E(Y)}{\sqrt{V(Y)}} = \beta_1 \cdot \frac{X_1 - E(X_1)}{\sqrt{V(X_1)}} + \beta_2 \cdot \frac{X_2 - E(X_2)}{\sqrt{V(X_2)}} + \beta_3 \cdot \frac{X_3 - E(X_3)}{\sqrt{V(X_3)}}.$$

Dostaneme tak

$$\beta_1 = \frac{1}{\sqrt{91}}, \quad \beta_2 = \frac{3}{\sqrt{91}}, \quad \beta_3 = \frac{9}{\sqrt{91}}.$$

Použitím rozkladu rozptylu dostáváme:

$$\begin{aligned} V_1 &= V_{X_1}(E(Y|X_1)) = V(X_1) = \frac{1}{12}, \\ V_2 &= V_{X_2}(E(Y|X_2)) = V(X_2) = \frac{9}{12}, \\ V_3 &= V_{X_3}(E(Y|X_3)) = V(X_3) = \frac{81}{12}. \end{aligned}$$

a následně

$$S_1 = \frac{1}{91}, \quad S_2 = \frac{9}{91}, \quad S_3 = \frac{81}{91}.$$

### 1.4.2 Příklad 2

Uvažujme model tvaru:

$$Y = X_1 + X_2^4, \tag{1.55}$$

kde  $X_1 \sim U(0, 5)$ ,  $X_2 \sim U(0, 5)$ .

### Analýza neurčitosti

Analýzu neurčitosti provedeme stejným způsobem jako v předešlém příkladu. Tentokrát získáme:

$$E(Y) = \frac{255}{2}, \quad V(Y) = \frac{1000075}{36}.$$

### Lokální analýza citlivosti

Přímo z definice získáváme:

$$S_1 = \frac{\partial Y}{\partial X_1} = 1, \quad S_2 = \frac{\partial Y}{\partial X_2} = 4 \cdot X_2^3.$$

Tentokrát vychází index citlivosti  $S_2$  závislý na parametru  $X_2$ , je tedy podstatné, v jakém bodě lokální citlivost vyhodnotíme.

## Globální analýza citlivosti

Tentokrát již náš model lineární není (v proměnný  $X_1, X_2$ ), v případě počítání regresních koeficientů jej musíme vytvořit. Vygenerujeme proto  $M = 10^5$  realizací veličiny  $X_1$  a stejný počet realizací veličiny  $X_2$ . Následně z nich vytvoříme  $M$  řešení modelu (1.55). V systému Maple získáme lineární model příkazem `LinearFit`, v Matlabu můžeme použít příkaz `regress`.

### Maple

```
M := 10^5;
with(Statistics):
X[1] := RandomVariable(Uniform(0, 5)):
X[2] := RandomVariable(Uniform(0, 5)):

X1s := Sample(X[1], M): # generovani M realizaci veliciny X[1]
X2s := Sample(X[2], M): # generovani M realizaci veliciny X[2]
Ys := Vector(M):
for i from 1 to M do Ys[i] := X1s[i]+X2s[i]^4 end do:

# uprava formatu hodnot X1s a X2s pro potreby prikazu LinearFit
sloupec1 := LinearAlgebra[Transpose](X1s):
sloupec2 := LinearAlgebra[Transpose](X2s):
Matice := Matrix(M, 2, [sloupec1, sloupec2]):

# linearni regrese
fit := LinearFit([1, x[1], x[2]], Matice, Ys, [x[1], x[2]], output = solutionmodule):
funkce := fit:-Results([leastsquaresfunction]): # linearni model
```

### Matlab

```
M=100000;
X1s = unifrnd(0, 5,[M 1]); % generovani M realizaci veliciny X[1]
X2s = unifrnd(0, 5,[M 1]); % generovani M realizaci veliciny X[2]

Ys = X1s + X2s.^4; % model

b=regress(Ys,[ones(M,1), X1s, X2s]) % regresni koeficienty
```

Získáme lineární model tvaru:

$$\hat{Y} \doteq 0,59 \cdot X_1 + 99,96 \cdot X_2 - 123,97. \quad (1.56)$$

Dalšími výpočty dostáváme:

$$R^2 \doteq 0,75, \quad \beta_1 \doteq 0,01, \quad \beta_2 \doteq 0,87.$$

Jestliže přiřadíme všem realizacím (u příslušných veličin) pořadí<sup>10</sup> (podle velikosti) a provedeme lineární regresi znovu, dostaneme koeficient determinace  $R^2 \doteq 0,98$  a model tvaru:

$$\hat{Y} \doteq 0,08 \cdot X_1 + 0,99 \cdot X_2 - 3390,25. \quad (1.57)$$

Další možností je výpočet Spearmanových koeficientů podle vztahu (1.23), po němž obdržíme:

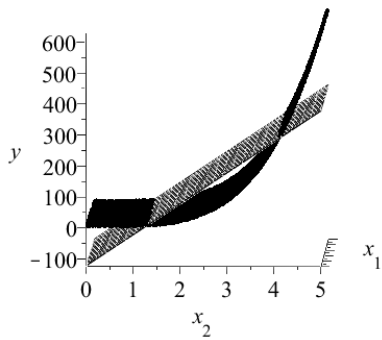
$$r_1 \doteq 0,07, \quad r_2 \doteq 0,99, \quad S_1 \doteq 0,01, \quad S_2 \doteq 0,99.$$

Pro posouzení kvalitativního vlivu proměnných  $X_1$  a  $X_2$  na řešení  $Y$  můžeme zobrazit grafy závislosti  $Y$  na  $X_1$  a  $Y$  na  $X_2$ , které nám v tomto případě velmi pomohou. Současně je možné zobrazit generované realizace veličin  $Y$ ,  $X_1$  a  $X_2$  s lineární funkcí (1.56).

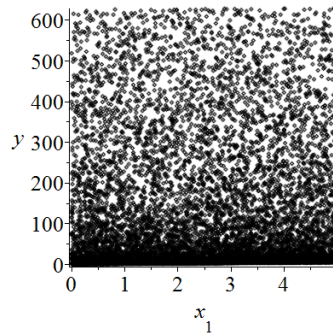
K výpočtu Spearmanových koeficientů a vykreslení grafů z obrázku 1.3 můžeme použít následující kód:

---

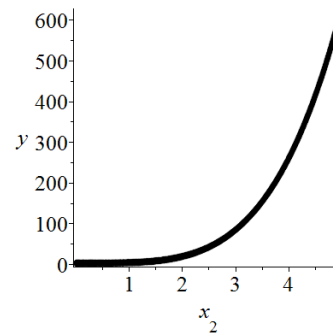
<sup>10</sup>V systému Maple k tomu lze využít příkaz `Rank`, v Matlabu příkaz `sort`.



(a) Generované body a lineární model (1.56)



(b) Závislost  $Y$  na  $X_1$



(c) Závislost  $Y$  na  $X_2$

Obrázek 1.3: Vykreslení grafů z ilustračního příkladu 2.

## Maple

```

R1 := Rank(X1s): # poradi prvku X1s podle velikosti
R2 := Rank(X2s): # poradi prvku X2s podle velikosti
RY := Rank(Ys): # poradi prvku Ys podle velikosti

# Spearmanovy koeficienty
r[1] := evalf(1-6*add((RY[i]-R1[i])^2, i = 1 .. M)/(M^3-M));
r[2] := evalf(1-6*add((RY[i]-R2[i])^2, i = 1 .. M)/(M^3-M));

# Vygenerovane body a linearni model
body := seq([X1s[i], X2s[i], Ys[i]], i = 1 .. M):
plot1 := plot3d(funkce, x[1] = 0 .. 5, x[2] = 0 .. 5, axes = frame,
labels = [typeset(x[1]), typeset(x[2]), typeset(y)]):
plot2 := plots[pointplot3d]([body], axes = frame):
plots[display]({plot1, plot2}):

# Graf zavislosti Y na X[1]
body2 := seq([X1s[i], Ys[i]], i = 1 .. M):
plots[pointplot]([body2], axes = frame, labels = [typeset(x[1]), typeset(y)]):

# Graf zavislosti Y na X[2]
body3 := seq([X2s[i], Ys[i]], i = 1 .. 10000):
plots[pointplot]([body3], axes = frame, labels = [typeset(x[2]), typeset(y)]):

```

## Matlab

```

% Spearmanovy koeficienty
r_1 = corr(X1s, Ys, 'type', 'Spearman')
r_2 = corr(X2s, Ys, 'type', 'Spearman')

% Vygenerovane body a linearni model
[x,y] = meshgrid(0:0.1:5, 0:0.1:5);
surf(x,y,b(1) + b(2)*x + b(3)*y) % lin. model
title('Graf_vygenerovanych_bodu_a_linearni_model')
xlabel('X[1]')
ylabel('X[2]')
zlabel('Y')
set(get(gca, 'ZLabel'), 'Rotation', 0.0)
hold on
plot3(X1s,X2s,Ys, 'r') % vygenerovane body

figure % novy obrazek
subplot(1,2,1); plot(X1s,Ys, 'r'); % Graf zavislosti Y na X[1]
title('Graf_zavislosti_Y_na_X[1]')
xlabel('X[1]')
ylabel('Y')
set(get(gca, 'YLabel'), 'Rotation', 0.0)
subplot(1,2,2); plot(X2s,Ys, 'r') % Graf zavislosti Y na X[2]
title('Graf_zavislosti_Y_na_X[2]')
xlabel('X[2]')
ylabel('Y')
set(get(gca, 'YLabel'), 'Rotation', 0.0)

```

Použitím rozkladu rozptylu dostáváme:

$$V_1 = V_{X_1}(E(Y|X_1)) = V(X_1) = \frac{25}{12},$$

$$V_2 = V_{X_2}(E(Y|X_2)) = V(X_2^4) = \frac{250000}{9}.$$

a následně

$$S_1 = \frac{V_1}{V(Y)} = \frac{3}{40003} \doteq 0,0001; \quad S_2 = \frac{V_2}{V(Y)} = \frac{40000}{40003} \doteq 0,9999.$$

### 1.4.3 Příklad 3

Uvažujme model tvaru:

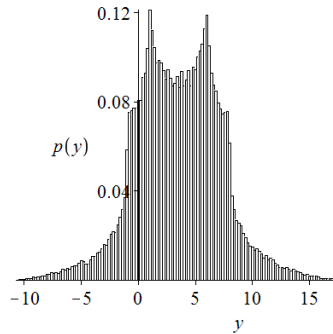
$$Y = \sin(X_1) + 7 \cdot \sin^2(X_2) + \frac{1}{10} \cdot X_3 \cdot \sin(X_1), \quad (1.58)$$

kde  $X_1 \sim U(-\pi, \pi)$ ,  $X_2 \sim U(-\pi, \pi)$ ,  $X_3 \sim U(-\pi, \pi)$ .

#### Analýza neurčitosti

$$E(Y) = \frac{7}{2}, \quad V(Y) = \frac{1}{1800} \cdot \pi^8 + \frac{1}{50} \cdot \pi^4 + \frac{53}{8}.$$

Pravděpodobnostní rozdělení veličiny  $Y$  můžeme odhadnout vygenerováním realizací pro  $X_1, X_2, X_3$ , následným vyhodnocením a zobrazením histogramu (v Maple příkaz `Histogram`, v Matlabu příkaz `hist`). Pro  $10^5$  vyhodnocení získáme v Maple histogram z obrázku 1.4



Obrázek 1.4: Histogram z ilustračního příkladu 3.

#### Lokální analýza citlivosti

Přímo z definice získáváme:

$$S_1 = \frac{\partial Y}{\partial X_1} = \cos(X_1) + \frac{1}{10} \cdot X_3^4 \cdot \cos(X_1), \quad S_2 = \frac{\partial Y}{\partial X_2} = 14 \cdot \sin(X_2) \cdot \cos(X_2),$$

$$S_3 = \frac{\partial Y}{\partial X_3} = \frac{2}{5} \cdot X_3^3 \cdot \sin(X_1).$$

## Globální analýza citlivosti

Pokusíme-li se znovu vytvořit lineární model, dostaneme  $R^2 \doteq 0,19$ . Přiřazení pořadí tentokrát nepomůže,  $R^2$  „zůstane“ na stejné hodnotě.

Použitím rozkladu rozptylu dostáváme:

$$V_1 = V_{X_1}(E(Y|X_1)) = V\left(\sin(X_1) + \frac{\pi^4}{50} \cdot \sin(X_1)\right) = \frac{1}{5000} \cdot \pi^8 + \frac{1}{50} \cdot \pi^4 + \frac{1}{2},$$

$$V_2 = V_{X_2}(E(Y|X_2)) = V(-7 \cdot \cos^2(X_2)) = \frac{49}{8},$$

$$V_3 = V_{X_3}(E(Y|X_3)) = V\left(\frac{7}{2}\right) = 0.$$

a následně

$$S_1 = \frac{V_1}{V(Y)} \doteq 0,3139; \quad S_2 = \frac{V_2}{V(Y)} \doteq 0,4424; \quad S_3 = \frac{V_3}{V(Y)} = 0.$$

Přestože jsme právě vypočítali globální indexy citlivosti 1. řádu, vypočítejme je nyní pomocí Sobol'ovy metody. Postupně dostáváme:

$$\widehat{E}(Y) \doteq 3,4958; \quad \widehat{V}(Y) \doteq 13,9476; \quad (V(Y) \doteq 13,8446);$$

$$\widehat{V}_1 \doteq 4,3680; \quad \widehat{V}_2 \doteq 6,1619; \quad \widehat{V}_3 \doteq -0,0561;$$

$$\widehat{S}_1 \doteq 0,3131; \quad \widehat{S}_2 \doteq 0,4418; \quad \widehat{S}_3 \doteq -0,0040.$$

Jestliže aplikujeme korekční člen (1.38), jehož hodnota je přibližně  $-0,0571$ , získáme:

$$\widehat{V}_1 \doteq 4,4250; \quad \widehat{V}_2 \doteq 6,2190; \quad \widehat{V}_3 \doteq 0,0010;$$

$$\widehat{S}_1 \doteq 0,3173; \quad \widehat{S}_2 \doteq 0,4459; \quad \widehat{S}_3 \doteq 0,0001;$$

Předchozí výsledky lze získat následujícími zdrojovými kódy:



## Maple

```
M := 10^5;
with(Statistics):
X[1] := RandomVariable(Uniform(-Pi, Pi)):
X[2] := RandomVariable(Uniform(-Pi, Pi)):
X[3] := RandomVariable(Uniform(-Pi, Pi)):

X1s := Sample(X[1], M): # M realizaci veliciny X[1]
X2s := Sample(X[2], M): # M realizaci veliciny X[2]
X3s := Sample(X[3], M): # M realizaci veliciny X[3]
Ys := Vector(M):

for i to M do Ys[i] := sin(X1s[i])+7*sin(X2s[i])^2+(1/10)*X3s[i]^4*sin(X1s[i]) end do:

f[0] := Mean(Ys): # odhad stredni hodnoty
V[0] := Variance(Ys): # odhad rozptylu

X1s2 := Sample(X[1], M): # "druha skupina" M realizaci veliciny X[1]
X2s2 := Sample(X[2], M): # "druha skupina" M realizaci veliciny X[2]
X3s2 := Sample(X[3], M): # "druha skupina" M realizaci veliciny X[3]
Ys2 := Vector(M):

# citlivost parametru X[1]
for i to M do Ys2[i] := sin(X1s2[i])+7*sin(X2s2[i])^2+(1/10)*X3s2[i]^4*sin(X1s2[i]) end do:
V[1] := add(Ys[i]*Ys2[i], i = 1 .. M)/M-f[0]^2;
S[1] = V[1]/V[0];

# citlivost parametru X[2]
for i to M do Ys3[i] := sin(X1s2[i])+7*sin(X2s2[i])^2+(1/10)*X3s2[i]^4*sin(X1s2[i]) end do:
V[2] := add(Ys[i]*Ys3[i], i = 1 .. M)/M-f[0]^2;
S[2] = V[2]/V[0];

# citlivost parametru X[3]
for i to M do Ys4[i] := sin(X1s2[i])+7*sin(X2s2[i])^2+(1/10)*X3s2[i]^4*sin(X1s2[i]) end do:
V[3] := add(Ys[i]*Ys4[i], i = 1 .. M)/M-f[0]^2;
S[3] = V[3]/V[0];

# korekčni clen
for i to M do Ysk[i] := sin(X1s2[i])+7*sin(X2s2[i])^2+(1/10)*X3s2[i]^4*sin(X1s2[i]) end do:
V[k] := add(Ys[i]*Ysk[i], i = 1 .. M)/M-f[0]^2;
```

## Matlab

```
M=100000;

% M realizaci jednotlivych velicin
X1s = unifrnd(-pi, pi, [M 1]);
X2s = unifrnd(-pi, pi, [M 1]);
X3s = unifrnd(-pi, pi, [M 1]);

Ys = sin(X1s) + 7*sin(X2s).^2 + (1/10)*X3s.^4.*sin(X1s);

mean_Ys = mean(Ys); % odhad stredni hodnoty
var_Ys = var(Ys); % odhad rozptylu

% "druha skupina" M realizaci jednotlivych velicin
X1s2 = unifrnd(-pi, pi, [M 1]);
X2s2 = unifrnd(-pi, pi, [M 1]);
X3s2 = unifrnd(-pi, pi, [M 1]);

% "druha reseni" pro jednotlivy vlivy
Ys2 = sin(X1s) + 7*sin(X2s2).^2 + (1/10)*X3s2.^4.*sin(X1s);
Ys3 = sin(X1s2) + 7*sin(X2s).^2 + (1/10)*X3s2.^4.*sin(X1s2);
Ys4 = sin(X1s2) + 7*sin(X2s2).^2 + (1/10)*X3s.^4.*sin(X1s2);
Ysk = sin(X1s2) + 7*sin(X2s2).^2 + (1/10)*X3s2.^4.*sin(X1s2);

% citlivost parametru X[1]
V1 = sum(Ys.*Ys2)/M - mean_Ys^2;
S1 = V1/var_Ys

% citlivost parametru X[2]
V2 = sum(Ys.*Ys3)/M - mean_Ys^2;
S2 = V2/var_Ys

% citlivost parametru X[3]
V3 = sum(Ys.*Ys4)/M - mean_Ys^2;
S3 = V3/var_Ys

% korekčni clen
Vk = sum(Ys.*Ysk)/M - mean_Ys^2;
S3 = Vk/var_Ys
```

### 1.4.4 Příklad 4

Jako závěrečný ilustrační příklad uvažujme chemickou reakci



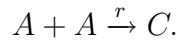
při níž z látek  $A$  a  $B$  vzniká látka  $C$ . Rychlost reakce  $r$  je definována jako (viz [7]):

$$r = -\frac{1}{a} \cdot \frac{d[A]}{dt} = -\frac{1}{b} \cdot \frac{d[B]}{dt} = \frac{1}{c} \cdot \frac{d[C]}{dt},$$

kde symboly v hranatých závorkách značí koncentraci příslušné látky, a platí pro ni Guldbergův-Waagův zákon (viz [1]):

$$r = k \cdot [A]^\alpha \cdot [B]^\beta. \quad (1.59)$$

Konstanta  $k$  z rovnice (1.59) se nazývá rychlostní konstanta. Koeficienty  $\alpha, \beta$  jsou obecně různé od stochiometrických koeficientů  $a, b$  v rovnici chemické reakce a musí být určeny experimentálně. Pro jednoduchost zvolme reakci



a předpokládejme, že  $\alpha = 1, \beta = 1$ . Rychlostní konstanta je tepelně závislá a platí pro ni Arrheniova rovnice (viz [17])

$$k = X_1 \cdot e^{-\frac{X_2}{T}}. \quad (1.60)$$

Nechť v našem případě  $X_1 \sim U(8, 97 \cdot 10^6; 3, 59 \cdot 10^7) \text{ dm}^3 \text{ mol s}^{-1}$ ,  $X_2 \sim U(0, 1000) \text{ K}^{-1}$  a  $T = 300 \text{ K}$ . Pro počáteční koncentraci  $Y_0$  látky  $A$  v čase  $t = 0 \text{ s}$  platí:  $Y_0 \sim U(1, 0; 1, 2) \text{ mol dm}^{-3}$ . Pro hodnotu koncentrace  $Y(t)$  látky  $A$  v čase  $t$  dostáváme modelovou rovnici

$$\frac{dY(t)}{dt} = -2 \cdot k \cdot Y^2(t), \quad Y(0) = Y_0. \quad (1.61)$$

Díky jednoduchosti modelové rovnice 1.61 jsme schopni určit její řešení, a získat tak model

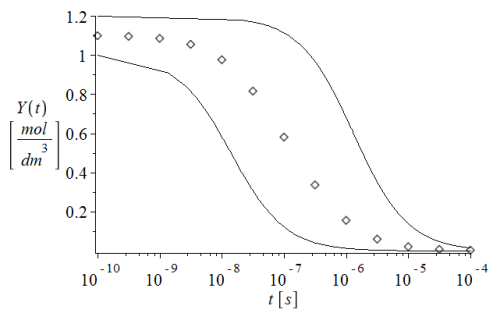
$$Y(t) = \left( 2 \cdot X_1 \cdot e^{-\frac{X_2}{T}} \cdot t + \frac{1}{Y_0} \right)^{-1}. \quad (1.62)$$

## Analýza neurčitosti

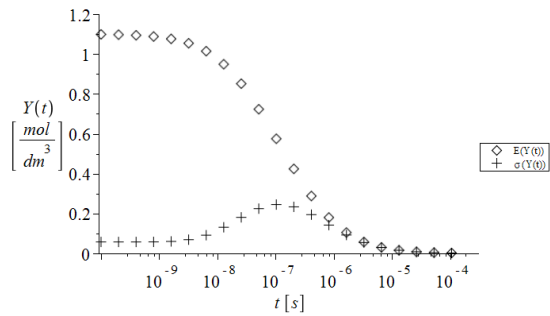
Řešení modelu je závislé na čase. K prvotnímu nahlédnutí, jak se řešení chová, můžeme vykreslit grafy hraničních situací, tj. maximální  $Y(t)$  a minimální  $Y(t)$  pro daná  $t$ . Maximální  $Y(t)$  dostaneme maximalizací veličin  $X_2$  a  $Y_0$  a minimalizací veličiny  $X_1$  (viz předpis řešení (1.62)). Analogicky získáme minimální  $Y(t)$ . Řešení se tak pro libovolné realizace veličin  $X_1, X_2, Y_0$  pohybuje mezi křivkami na obrázku 1.5.(a). Na tomtéž grafu jsou také odhady střední hodnoty (v některých časových bodech).

Od počátku až do času  $t = 10^{-4}$  je zřejmý rozptyl hodnot způsobený neurčitostmi v parametrech  $X_1, X_2$  a  $Y_0$ . Dále od tohoto časového bodu převažuje vliv času  $t$  na hodnotu  $Y(t)$  a neurčitosti v parametrech tak mohou být zanedbány. Zaměříme se proto pouze na vyšetření intervalu  $t \in [10^{-10}, 10^{-4}]$ , jak je tomu i na obrázku 1.5.

Na grafu 1.5.(b) jsou zobrazeny odhady střední hodnoty a směrodatné odchylky řešení  $Y(t)$ . Můžeme pozorovat, že největší směrodatné odchylky dosahuje  $Y(t)$  kolem času  $t = 10^{-7}$ . K tomuto grafu lze dospět následujícími zdrojovými kódy.



(a) Maximální a minimální řešení (1.62)



(b) Odhad střední hodnoty a směrodatné odchylky řešení (1.62)

Obrázek 1.5: Grafy analýzy neurčitosti ilustračního příkladu 4.

## Maple

```

M := 10^5;
with(Statistics):
X[1] := RandomVariable(Uniform(8.97*10^6, 3.59*10^7));
X[2] := RandomVariable(Uniform(0, 1000));
Y0 := RandomVariable(Uniform(1, 1.2));

X1s := Sample(X[1], M): # M realizaci veliciny X[1]
X2s := Sample(X[2], M): # M realizaci veliciny X[2]
Y0s := Sample(Y0, M): # M realizaci veliciny X[3]
Ys := Vector(M):

for i to M do Ys[i] := Y0s[i]/(2*X1s[i]*exp(-(1/300)*X2s[i])*t*Y0s[i]+1) end do:

# vypočet str. hodnot a smerodat. odchylek
for i from 1 to 21 do
t := 10^(-10)*2^(i-1);
mean_Ys[i] := Mean(Ys);
dev_Ys[i] := StandardDeviation(Ys);
end do:

t := 't':

# "příprava" bodu k vykreslení
str_hodnoty := seq([10^(-10)*2^(i-1), mean_Ys[i]], i = 1 .. 21);
odchylky := seq([10^(-10)*2^(i-1), dev_Ys[i]], i = 1 .. 21);

# vykreslení
plot1 := plots[pointplot]([str_hodnoty]);
plot2 := plots[pointplot]([odchylky], symbol = cross);
plots[display]({plot1, plot2, plots[semilogplot](0, view = [10^(-10) .. 3*10^(-4), 0 .. 1.2],
labels = [typeset(t*[s]), typeset(Y(t)*[mol/dm^3])])});

```

## Matlab

```

M=100000;
% M realizaci jednotlivych velicin
X1s = unifrnd(8.97*10^6, 3.59*10^7, [M 1]);
X2s = unifrnd(0, 1000, [M 1]);
Y0s = unifrnd(1, 1.2, [M 1]);

mean_Ys=zeros(1,21,'double');
dev_Ys=zeros(1,21,'double');
time=zeros(1,21,'double');

% vypočet strednich hodnot a smerodatnych odchylek
for i=1:21
t(i) = 10^(-10)*2^(i-1);
Ys = Y0s./(2*X1s.*exp(-(1/300)*X2s)*t(i).*Y0s+1);
mean_Ys(i) = mean(Ys);
dev_Ys(i) = std(Ys);
end

semilogx(t,mean_Ys, '.'); % vykreslení str. hodnot
hold on
semilogx(t,dev_Ys, '+'); % vykreslení std. odchylek
xlabel('t-[s]');
ylabel('Y(t)-[mol/dm^3]');

```

## Lokální analýza citlivosti

Lokální citlivosti bychom získali přímo z definice díky předpisu řešení (1.62). Tentokrát jsou citlivosti závislé nejen na všech parametrech  $X_1, X_2, Y_0$ , ale též na čase  $t$ .

## Globální analýza citlivosti

Analýzu globální citlivosti je třeba vyšetřit opět v závislosti na čase. Vzhledem k tomu, že se nám nepodaří určit globální citlivosti analyticky, musíme je počítat numericky pro konkrétní časové body. Aplikujme tentokrát metodu FAST konkrétně na bod  $t = 10^{-7}$ , kde dosahuje směrodatná odchylka (a potažmo i rozptyl) řešení nejvyšších hodnot.

Nejprve je třeba zvolit transformační funkce  $G_i$  a úhlové frekvence  $\omega_i$  (pro  $i = 1, 2, 3$ ). Využijme dříve uvedené funkce  $G_i(s) = \frac{1}{2} + \frac{1}{\pi} \cdot \arcsin(\sin(\omega_i \cdot s))$ . Jelikož tato funkce transformuje interval  $(-\pi, \pi)$  na interval  $(0, 1)$ , musíme ji pro naše potřeby upravit. Všechny náhodné veličiny v tomto příkladu mají rovnoměrné rozdělení, takže stačí zapsanou funkci  $G_i$  vynásobit délkou příslušného intervalu neurčitosti a přičíst dolní mez tohoto intervalu. Dostaneme tedy

$$G_1(s) = 2,693 \cdot 10^7 \cdot \left( \frac{1}{2} + \frac{1}{\pi} \cdot \arcsin(\sin(\omega_1 \cdot s)) \right) + 8,97 \cdot 10^6,$$

$$G_2(s) = 1000 \cdot \left( \frac{1}{2} + \frac{1}{\pi} \cdot \arcsin(\sin(\omega_2 \cdot s)) \right),$$

$$G_3(s) = 0,2 \cdot \left( \frac{1}{2} + \frac{1}{\pi} \cdot \arcsin(\sin(\omega_3 \cdot s)) \right) + 1.$$

Úhlové frekvence  $\omega_i$  nechť splňují  $\omega_1 = 11, \omega_2 = 21, \omega_3 = 31$ . Transformovali jsme tak modelovou rovnici (1.62) na rovnici

$$Y(t) = \left( 2 \cdot G_1(S) \cdot e^{-\frac{G_2(S)}{T}} \cdot t + \frac{1}{G_3(S)} \right)^{-1}, \quad (1.63)$$

kde  $S \sim U(-\pi, \pi)$ . Počet generovaných bodů ponechme stejný, tj.  $M = 10^5$ . Nechť dále počet harmonických složek  $N = 6$ . Získáme tak:

$$\widehat{E}(Y(10^{-7})) \doteq 0,5801; \quad \widehat{V}(Y(10^{-7})) \doteq 0,0599;$$

$$\widehat{V}_1 \doteq 0,0070; \quad \widehat{V}_2 \doteq 0,0513; \quad \widehat{V}_3 \doteq 0,0008;$$

$$\widehat{S}_1 \doteq 0,1177; \quad \widehat{S}_2 \doteq 0,8573; \quad \widehat{S}_3 \doteq 0,0128.$$

Pro výpočet lze použít zobrazené zdrojové kódy.

## Maple

```

M := 10^5; N:=6;
with(Statistics):
S := RandomVariable(Uniform(-Pi, Pi)):
Ss := Sample(S, M):

# transformacni funkce
G[1](s):=2.693*10^(7)*(1/(2)+1/(Pi)*arcsin(sin(omega[1]*s)))+8.97*10^(6):
G[2](s):=1000*(1/(2)+1/(Pi)*arcsin(sin(omega[2]*s))):
G[3](s):=0.2*(1/(2)+1/(Pi)*arcsin(sin(omega[3]*s)))+1:

# uhlove frekvence
omega[1] := 11; omega[2] := 21; omega[3] := 31:

Ys := Vector(M):
for i from 1 to M do
  Ys[i] := evalf(G[3](Ss[i])/(2*G[1](Ss[i])*exp(-1/300)*G[2](Ss[i]))*10^(-7)*G[3](Ss[i]+1))
end do:

f[0] := Mean(Ys); # odhad stredni hodnoty
V[0] := Variance(Ys); # odhad rozptylu

# vypocet Fourierovych koeficientu
for j from 1 to N do
  fk[j] := add(Ys[i]*cos(j*omega[1]*Ss[i]), i = 1 .. M)/M:
  fk[j+N] := add(Ys[i]*sin(j*omega[1]*Ss[i]), i = 1 .. M)/M:
  fk2[j] := add(Ys[i]*cos(j*omega[2]*Ss[i]), i = 1 .. M)/M:
  fk2[j+N] := add(Ys[i]*sin(j*omega[2]*Ss[i]), i = 1 .. M)/M:
  fk3[j] := add(Ys[i]*cos(j*omega[3]*Ss[i]), i = 1 .. M)/M:
  fk3[j+N] := add(Ys[i]*sin(j*omega[3]*Ss[i]), i = 1 .. M)/M:
end do:

# citlivost parametru X[1]
V[1] := 2*add(fk[i]^2, i = 1 .. 12);
S[1] = V[1]/V[0];

# citlivost parametru X[2]
V[2] := 2*add(fk2[i]^2, i = 1 .. 12);
S[2] = V[2]/V[0];

# citlivost parametru Y[0]
V[3] := 2*add(fk3[i]^2, i = 1 .. 12);
S[3] = V[3]/V[0];

```

## Matlab

```

M=100000; N=6;
omega1=11; omega2=21; omega3=31; % uhlove frekvence

Ss = unifrnd(-pi, pi, [M 1]);
x1 = 2.693*10^7*(.5 + asin(sin(omega1*Ss))/pi)+8.97*10^6;
x2 = 1000*(.5 + asin(sin(omega2*Ss))/pi);
y0 = .2*(.5 + asin(sin(omega3*Ss))/pi)+1;

% modelova rovnice
Y = y0./(2*x1.*exp(-x2/300)*10^(-7).*y0 + 1);

mean_Y = mean(Y) % odhad stredni hodnoty
var_Y = var(Y) % odhad rozptylu

fk=zeros(1,2*N,'double');
fk2=zeros(1,2*N,'double');
fk3=zeros(1,2*N,'double');

% vypocet Fourierovych koeficientu
for j=1:N
  fk(j) = sum(Y.*cos(j*omega1*Ss))/M;
  fk(j+N) = sum(Y.*sin(j*omega1*Ss))/M;
  fk2(j) = sum(Y.*cos(j*omega2*Ss))/M;
  fk2(j+N) = sum(Y.*sin(j*omega2*Ss))/M;
  fk3(j) = sum(Y.*cos(j*omega3*Ss))/M;
  fk3(j+N) = sum(Y.*sin(j*omega3*Ss))/M;
end

% citlivost parametru X[1]
V1 = 2*sum(fk.^2)
S1 = V1/var_Y

% citlivost parametru X[2]
V2 = 2*sum(fk2.^2)
S2 = V2/var_Y

% citlivost parametru Y[0]
V3 = 2*sum(fk3.^2)
S3 = V3/var_Y

```

## 1.5 Příklady k procvičení

**Příklad 1.6:** Uvažujme model  $Y = X_1 + X_2$ , kde  $X_1 \sim N(1, 4)$ ,  $X_2 \sim N(2, 9)$ . Určete  $E(Y)$ ,  $V(Y)$  a indexy globální citlivosti prvního řádu.<sup>11</sup>

**Příklad 1.7:** Nechť

$$Y = 2^{X_1} \cdot X_2,$$

kde  $X_1 \sim U(0, 3)$ ,  $X_2 \sim U(0, 3)$ . Proveďte analýzu neurčitosti i analýzu citlivosti řešení uvedeného modelu, tj. určete  $E(Y)$ ,  $V(Y)$ ,  $p(y)$  a indexy globální i lokální citlivosti prvního řádu.

**Příklad 1.8:** Uvažujme model

$$Y = \frac{X_1^4}{X_2^2},$$

kde  $X_1 \sim U(\frac{1}{2}, \frac{3}{2})$ ,  $X_2 \sim U(\frac{1}{2}, \frac{3}{2})$ . Určete  $E(Y)$ ,  $V(Y)$  a indexy globální i lokální citlivosti prvního řádu.

**Příklad 1.9:** Nechť

$$Y = \sum_{i=1}^8 X_i,$$

kde  $X_1 \sim N(0, 1)$ ,  $X_2 \sim N(2, 4)$ ,  $X_3 \sim N(1, 9)$ ,  $X_4 \sim N(1, 25)$ ,  $X_5 \sim N(3, 1)$ ,  $X_6 \sim N(4, 1)$ ,  $X_7 \sim N(1, 4)$ ,  $X_8 \sim N(5, 25)$ . Určete, které tři parametry nejméně ovlivňují rozptyl řešení  $Y$ .

**Příklad 1.10:** Uvažujme model  $Y = \sin(X_1) + 7 \cdot \sin^2(X_2) + \frac{1}{10} \cdot X_3 \cdot \sin(X_1)$  z třetího ilustračního příkladu se stejně rozdělenými parametry  $X_1, X_2, X_3$ . Vypočítejte indexy globální citlivosti prvního řádu pomocí metody FAST.

**Příklad 1.11:** Vraťte se k modelu chemické rovnice ze čtvrtého ilustračního příkladu. Vypočítejte indexy globální citlivosti prvního řádu pomocí Sobol'ovy metody. Dále vytvořte lineární modely (a to jak „klasické“, tak pro přiřazené pořadí) pro různé časové body v intervalu  $[10^{-10}, 10^{-4}]$  a zkoumejte závislost koeficientu determinace na čase.

---

<sup>11</sup> $N(\mu, \sigma^2)$  značí normální rozdělení se střední hodnotou  $\mu$  a rozptylem  $\sigma^2$ .

# Literatura

- [1] Burda, J.: Chemická kinetika [online]. [cit 2011-09-25]. Dostupný z WWW: <<http://physics.mff.cuni.cz/kchfo/burda/0Ch/Part09.pdf>>.
- [2] Drog, M.: Dealing with uncertainties. Springer (2009)
- [3] Homma, T., Saltelli, A.: Importance measures in global sensitivity analysis of nonlinear models. Reliability Engineering and System Safety (1996)
- [4] Horová, I., Zelinka, J.: Numerické metody. Masarykova univerzita, Brno (2004)
- [5] Hřebíček, J., Pospíšil, Z., Urbánek, J.: Úvod do matematického modelování s využitím Maple. CERM, Brno (2010)
- [6] Hušková, M.: Bayesovské metody. Karlova univerzita, Praha (1985)
- [7] IUPAC: Gold Book [online]. [cit 2011-09-25]. Dostupný z WWW: <<http://goldbook.iupac.org/R05156.html>>.
- [8] Komárek, A.: Bayesovské metody (druhá část) [online]. [cit 2011-09-01]. Dostupný z WWW: <[http://www.karlin.mff.cuni.cz/~komarek/vyuka/2010\\_11/nstp021/NSTP021-Bayes-Komarek-tisk.pdf](http://www.karlin.mff.cuni.cz/~komarek/vyuka/2010_11/nstp021/NSTP021-Bayes-Komarek-tisk.pdf)>.
- [9] Luo, Y., Yang, X.: A multimedia environmental model of chemical distribution: Fate, transport, and uncertainty analysis Chemosphere, 66 (8), p. 1396 – 1407 (2007)
- [10] Muehlenstaedt, T., Roustant, O., Carraro, L., Kuhnt, S.: Data-driven Kriging models based on FANOVA-decomposition [online]. [cit 2011-09-09]. Dostupný z WWW: <[http://hal.archives-ouvertes.fr/docs/00/53/77/81/PDF/Data-Driven\\_Kriging\\_models\\_based\\_on\\_FANOVA\\_decomposition.pdf](http://hal.archives-ouvertes.fr/docs/00/53/77/81/PDF/Data-Driven_Kriging_models_based_on_FANOVA_decomposition.pdf)>.
- [11] Qian, S. S., Stow, C. A., Borsuk, M. E.: On Monte Carlo methods for Bayesian inference [online]. [cit 2011-09-01]. Dostupný z WWW: <<http://www.namland.com/journal/science.pdf>>.
- [12] Roustant, O.: Analytical computation of Sobol indices with a Gaussian process metamodel [online]. [cit 2011-09-09]. Dostupný z WWW: <[http://www.emse.fr/~roustant/Documents/Analytical%20computation%20of%20Sobol%20indices\\_kriging%20mean.pdf](http://www.emse.fr/~roustant/Documents/Analytical%20computation%20of%20Sobol%20indices_kriging%20mean.pdf)>.
- [13] Saltelli, A., Chan, K., Scott, E. M.: Sensitivity analysis. Wiley (2000)
- [14] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S.: Global sensitivity analysis. Wiley (2008)

- [15] Saltelli, A., Tarantola, S., Campolongo, F.: Sensitivity Analysis as an Ingredient of Modeling. Statistical Science (2000)
- [16] Shonkwiler, R. W., Herod, J.: Mathematical Biology. An Introduction with Maple and Matlab. Springer (2009)
- [17] Wikipedia: Arrhenius equation [online]. [cit 2011-09-25]. Dostupný z WWW: <[http://en.wikipedia.org/wiki/Arrhenius\\_equation](http://en.wikipedia.org/wiki/Arrhenius_equation)>.
- [18] Wikipedia: Shannonův teorém [online]. [cit 2011-09-18]. Dostupný z WWW: <[http://cs.wikipedia.org/wiki/Shannon%C5%AFv\\_teor%C3%A9m](http://cs.wikipedia.org/wiki/Shannon%C5%AFv_teor%C3%A9m)>.
- [19] Wikipedia: Spearman's rank correlation coefficient [online]. [cit 2011-09-08]. Dostupný z WWW: <[http://en.wikipedia.org/wiki/Spearman%27s\\_rank\\_correlation\\_coefficient](http://en.wikipedia.org/wiki/Spearman%27s_rank_correlation_coefficient)>.