Intervalová data, algoritmy a výpočetní složitost

Michal Černý [1] Milan Hladík [2]

[1] Katedra ekonometrie
Fakulta informatiky a statistiky
Vysoká škola ekonomická Praha

[2] Katedra aplikované matematiky
Matematicko-fyzikální fakulta
Univerzita Karlova

ROBUST 2012, Němčičky,
9.-14. září

Intervalová data, algoritmy a výpočetní složitost

Michal Černý [1] Milan Hladík [2]

[1] Katedra ekonometrie
Fakulta informatiky a statistiky
Vysoká škola ekonomická Praha

[2] Katedra aplikované matematiky
Matematicko-fyzikální fakulta
Univerzita Karlova

ROBUST 2012, Němčičky,
9.-14. září

Za pozvání děkujeme J. Antochovi

Úvod

Úvod

- Intervalová data jsou data tvaru $[\underline{x}, \overline{x}]$.

Úvod

- Intervalová data jsou data tvaru $[\underline{x}, \overline{x}]$.
- Vznikají v řadě praktických situací: například
  - při zaokrouhlování a při reprezentaci dat pomocí datových typů s omezeným počtem desetinných míst;

Úvod

- Intervalová data jsou data tvaru $[\underline{x}, \overline{x}]$.
- Vznikají v řadě praktických situací: například
  - při zaokrouhlování a při reprezentaci dat pomocí datových typů s omezeným počtem desetinných míst;
  - obecně: při ztrátě informace, například
    - při kategorizaci dat, při utajování konkrétních hodnot, při některých typech cenzorování, při diskretizaci spojitých dat;

Úvod

- Intervalová data jsou data tvaru $[\underline{x}, \overline{x}]$.
- Vznikají v řadě praktických situací: například
  - při zaokrouhlování a při reprezentaci dat pomocí datových typů s omezeným počtem desetinných míst;
  - obecně: při ztrátě informace, například
    - při kategorizaci dat, při utajování konkrétních hodnot, při některých typech cenzorování, při diskretizaci spojitých dat;
  - při nestabilitě dat, například
    - při nestabilitě fyzikálních „konstant“,
    - při pohybu sledované veličiny uvnitř období (např. u denních burzovních dat je vhodné brát v úvahu, že uvnitř dne se hodnoty mění);

Úvod

- Intervalová data jsou data tvaru $[\underline{x}, \overline{x}]$.
- Vznikají v řadě praktických situací: například
  - při zaokrouhlování a při reprezentaci dat pomocí datových typů s omezeným počtem desetinných míst;
  - obecně: při ztrátě informace, například
    - při kategorizaci dat, při utajování konkrétních hodnot, při některých typech cenzorování, při diskretizaci spojitých dat;
  - při nestabilitě dat, například
    - při nestabilitě fyzikálních „konstant“,
    - při pohybu sledované veličiny uvnitř období (např. u denních burzovních dat je vhodné brát v úvahu, že uvnitř dne se hodnoty mění);
  - při expertní predikci:
    - *expert, poučen letitou zkušeností, pohlédne z okna a praví: „příští rok splníme plán na 120% – 140%...“;*

Úvod

- Intervalová data jsou data tvaru $[\underline{x}, \overline{x}]$.
- Vznikají v řadě praktických situací: například
  - při zaokrouhlování a při reprezentaci dat pomocí datových typů s omezeným počtem desetinných míst;
  - obecně: při ztrátě informace, například
    - při kategorizaci dat, při utajování konkrétních hodnot, při některých typech cenzorování, při diskretizaci spojitých dat;
  - při nestabilitě dat, například
    - při nestabilitě fyzikálních „konstant",
    - při pohybu sledované veličiny uvnitř období (např. u denních burzovních dat je vhodné brát v úvahu, že uvnitř dne se hodnoty mění);
  - při expertní predikci:
    - *expert, poučen letitou zkušeností, pohlédne z okna a praví: „příští rok splníme plán na 120% – 140%. . . ";*
  - v situacích, kdy intervalové predikce jednoho modelu vstupují jako data do druhého modelu:
    - *například — první model generuje intervalovou predikci inflace (= racionální inflační očekávání) pro příští rok; druhý model, vysvětlující spotřebu či investice běžného období, má za regresor inflační očekávání přístího roku.*

Pravděpodobnostní pohled na intervalová data

Pravděpodobnostní pohled na intervalová data

Mějme intervalová data

$$[\underline{x}_1, \overline{x}_1], \ldots, [\underline{x}_t, \overline{x}_t], \ldots, [\underline{x}_n, \overline{x}_n]$$

a předpokládejme, že byla generována dvojicí náhodných procesů $v_t \in \mathbb{R}$, $w_t \geq 0$ takových, že

- v_t generuje středy $\frac{1}{2}[\overline{x}_t + \underline{x}_t]$,
- w_t generuje poloměry $\frac{1}{2}[\overline{x}_t - \underline{x}_t]$.

Pravděpodobnostní pohled na intervalová data

Mějme intervalová data

$$[\underline{x}_1, \overline{x}_1], \ldots, [\underline{x}_t, \overline{x}_t], \ldots, [\underline{x}_n, \overline{x}_n]$$

a předpokládejme, že byla generována dvojicí náhodných procesů $v_t \in \mathbb{R}$, $w_t \geq 0$ takových, že

- v_t generuje středy $\frac{1}{2}[\overline{x}_t + \underline{x}_t]$,
- w_t generuje poloměry $\frac{1}{2}[\overline{x}_t - \underline{x}_t]$.

Nyní je přirozené začít budovat teorii:

- předpokládat vhodná rozdělení v_t a w_t a konstruovat estimátory jejich parametrů;
- uvažovat případy závislých a nezávislých v_t a w_t;

Pravděpodobnostní pohled na intervalová data

Mějme intervalová data

$$[\underline{x}_1, \overline{x}_1], \ldots, [\underline{x}_t, \overline{x}_t], \ldots, [\underline{x}_n, \overline{x}_n]$$

a předpokládejme, že byla generována dvojicí náhodných procesů $v_t \in \mathbb{R}$, $w_t \geq 0$ takových, že

- v_t generuje středy $\frac{1}{2}[\overline{x}_t + \underline{x}_t]$,
- w_t generuje poloměry $\frac{1}{2}[\overline{x}_t - \underline{x}_t]$.

Nyní je přirozené začít budovat teorii:

- předpokládat vhodná rozdělení v_t a w_t a konstruovat estimátory jejich parametrů;
- uvažovat případy závislých a nezávislých v_t a w_t;
- konstruovat rozličné testy;
- formulovat intervalový regresní model atd.

Pravděpodobnostní pohled na intervalová data

Mějme intervalová data

$$[\underline{x}_1, \overline{x}_1], \ldots, [\underline{x}_t, \overline{x}_t], \ldots, [\underline{x}_n, \overline{x}_n]$$

a předpokládejme, že byla generována dvojicí náhodných procesů $v_t \in \mathbb{R}$, $w_t \geq 0$ takových, že

- v_t generuje středy $\frac{1}{2}[\overline{x}_t + \underline{x}_t]$,
- w_t generuje poloměry $\frac{1}{2}[\overline{x}_t - \underline{x}_t]$.

Nyní je přirozené začít budovat teorii:

- předpokládat vhodná rozdělení v_t a w_t a konstruovat estimátory jejich parametrů;
- uvažovat případy závislých a nezávislých v_t a w_t;
- konstruovat rozličné testy;
- formulovat intervalový regresní model atd.

Problém.

Pravděpodobnostní pohled na intervalová data

Mějme intervalová data

$$[\underline{x}_1, \overline{x}_1], \dots, [\underline{x}_t, \overline{x}_t], \dots, [\underline{x}_n, \overline{x}_n]$$

a předpokládejme, že byla generována dvojicí náhodných procesů $v_t \in \mathbb{R}$, $w_t \geq 0$ takových, že

- v_t generuje středy $\frac{1}{2}[\overline{x}_t + \underline{x}_t]$,
- w_t generuje poloměry $\frac{1}{2}[\overline{x}_t - \underline{x}_t]$.

Nyní je přirozené začít budovat teorii:

- předpokládat vhodná rozdělení v_t a w_t a konstruovat estimátory jejich parametrů;
- uvažovat případy závislých a nezávislých v_t a w_t;
- konstruovat rozličné testy;
- formulovat intervalový regresní model atd.

Problém. My se na to díváme jinak.

Algoritmický pohled na intervalová data

Algoritmický pohled na intervalová data

Mějme intervalová data $[\underline{x}_1, \overline{x}_1], \dots, [\underline{x}_t, \overline{x}_t], \dots, [\underline{x}_n, \overline{x}_n]$.

Nepřijímáme o nich žádné pravděpodobnostní předpoklady:

- nepředpokládáme, že intervalová data byla generována některým konkrétním náhodným procesem;
- nepředpokládáme, že nedostupná (nepozorovatelná) hodnota x_t, o níž víme jen $x_t \in [\underline{x}_t, \overline{x}_t]$ a o niž nám ve skutečnosti jde, je náhodnou veličinou nad intervalem $[\underline{x}_t, \overline{x}_t]$.

Proč?

Algoritmický pohled na intervalová data

Mějme intervalová data $[\underline{x}_1, \overline{x}_1], \ldots, [\underline{x}_t, \overline{x}_t], \ldots, [\underline{x}_n, \overline{x}_n]$.

Nepřijímáme o nich žádné pravděpodobnostní předpoklady:

- nepředpokládáme, že intervalová data byla generována některým konkrétním náhodným procesem;
- nepředpokládáme, že nedostupná (nepozorovatelná) hodnota x_t, o níž víme jen $x_t \in [\underline{x}_t, \overline{x}_t]$ a o niž nám ve skutečnosti jde, je náhodnou veličinou nad intervalem $[\underline{x}_t, \overline{x}_t]$.

Proč?

- Nepotřebujeme to. Naším cílem je zkoumat výpočetní složitost algoritmů zpracovávajících intervalová data a výpočetní složitost *problémů*, které se řeší při zpracování intervalových dat.

Algoritmický pohled na intervalová data

Mějme intervalová data $[\underline{x}_1, \overline{x}_1], \ldots, [\underline{x}_t, \overline{x}_t], \ldots, [\underline{x}_n, \overline{x}_n]$.

Nepřijímáme o nich žádné pravděpodobnostní předpoklady:

- nepředpokládáme, že intervalová data byla generována některým konkrétním náhodným procesem;
- nepředpokládáme, že nedostupná (nepozorovatelná) hodnota x_t, o níž víme jen $x_t \in [\underline{x}_t, \overline{x}_t]$ a o niž nám ve skutečnosti jde, je náhodnou veličinou nad intervalem $[\underline{x}_t, \overline{x}_t]$.

Proč?

- Nepotřebujeme to. Naším cílem je zkoumat výpočetní složitost algoritmů zpracovávajících intervalová data a výpočetní složitost *problémů*, které se řeší při zpracování intervalových dat.
- Pro algoritmus jsou data vždy jen pevná čísla $\underline{x}_1, \overline{x}_1, \ldots, \underline{x}_n, \overline{x}_n$.
- Jediné, co potřebujeme předpokládat, je, že jde o racionální čísla.

Algoritmický pohled na intervalová data

Mějme intervalová data $[\underline{x}_1, \overline{x}_1], \ldots, [\underline{x}_t, \overline{x}_t], \ldots, [\underline{x}_n, \overline{x}_n]$.

Nepřijímáme o nich žádné pravděpodobnostní předpoklady:

- nepředpokládáme, že intervalová data byla generována některým konkrétním náhodným procesem;
- nepředpokládáme, že nedostupná (nepozorovatelná) hodnota x_t, o níž víme jen $x_t \in [\underline{x}_t, \overline{x}_t]$ a o niž nám ve skutečnosti jde, je náhodnou veličinou nad intervalem $[\underline{x}_t, \overline{x}_t]$.

Proč?

- Nepotřebujeme to. Naším cílem je zkoumat výpočetní složitost algoritmů zpracovávajících intervalová data a výpočetní složitost *problémů*, které se řeší při zpracování intervalových dat.
- Pro algoritmus jsou data vždy jen pevná čísla $\underline{x}_1, \overline{x}_1, \ldots, \underline{x}_n, \overline{x}_n$.
- Jediné, co potřebujeme předpokládat, je, že jde o racionální čísla.
- Někdy je dokonce příjemné, že nemusíme např. přijímat předpoklad o tvaru rozdělení nepozorovatelné hodnoty x_t nad intervalem $[\underline{x}_t, \overline{x}_t]$.

Značení

Značení

- Intervalová matice $\mathbf{X}$ rozměru $n \times p$ je systém matic
$$\mathbf{X} := [\underline{X}, \overline{X}] := \{X \in \mathbb{R}^{n\times p} : \underline{X} \leq X \leq \overline{X}\},$$
kde nerovnost „$\leq$“ se rozumí *po složkách*.
- Systém intervalových matic rozměru $n \times p$ značíme $\mathbb{IR}^{n\times p}$.
- Intervalové matice značíme $\mathbf{X}, \mathbf{Y}, \ldots$; standardní (pevné) matice značíme $X, Y, \ldots$
- Intervalové vektory ($=$ matice $n \times 1$) značíme $\mathbf{x}, \mathbf{y}, \ldots$; standardní (pevné) vektory značíme $x, y, \ldots$

Motivační příklad na začátek

Motivační příklad na začátek

- Mějme k dispozici intervalová data $(\mathbf{X}, \mathbf{y})$. Uvažme „regresní model" tvaru

$$„\mathbf{y} = \mathbf{X}\beta + \varepsilon“, \tag{1}$$

kde β značí vektor regresních parametrů.

Motivační příklad na začátek

- Mějme k dispozici intervalová data $(\mathbf{X}, \mathbf{y})$. Uvažme „regresní model" tvaru
$$„\mathbf{y} = \mathbf{X}\beta + \varepsilon", \tag{1}$$
kde β značí vektor regresních parametrů.

- Nadále značíme
  - n = počet pozorování,
  - p = počet regresních parametrů.

Motivační příklad na začátek

- Mějme k dispozici intervalová data $(\mathbf{X}, \mathbf{y})$. Uvažme „regresní model“ tvaru
$$„\mathbf{y} = \mathbf{X}\beta + \varepsilon“, \tag{1}$$
kde β značí vektor regresních parametrů.

- Nadále značíme
  - n = počet pozorování,
  - p = počet regresních parametrů.

- Na „model“ (1) můžeme nahlížet jako na systém instancí modelu $y = X\beta + \varepsilon$, kde
  - matice (pozorovaných hodnot) nezávislých proměnných X probíhá intervalovou matici $\mathbf{X}$,
  - vektor (pozorovaných hodnot) závislé proměnné y probíhá intervalový vektor $\mathbf{y}$.
  - jednotlivé složky X a y probíhají $\mathbf{X}$ a $\mathbf{y}$ nezávisle.

Množina B_2

Množina B_2

Zabývejme se množinou

$$B_2 := \{b \in \mathbb{R}^p : X^{\mathrm{T}} X b = X^{\mathrm{T}} y \ \text{ pro některé } X \in \mathbf{X} \text{ a } y \in \mathbf{y}\}. \qquad (2)$$

Množina B_2 je jednou z (více) možných formalizací pojmu *odhad „modelu“* $\mathbf{y} = \mathbf{X}\beta + \varepsilon$ *pomocí nejmenších čtverců* při intervalových datech $(\mathbf{X}, \mathbf{y})$.

Množina B_2

Zabývejme se množinou

$$B_2 := \{b \in \mathbb{R}^p : X^{\mathrm{T}} X b = X^{\mathrm{T}} y \ \text{ pro některé } X \in \mathbf{X} \text{ a } y \in \mathbf{y}\}. \qquad (2)$$

Množina B_2 je jednou z (více) možných formalizací pojmu *odhad „modelu“* $\mathbf{y} = \mathbf{X}\beta + \varepsilon$ *pomocí nejmenších čtverců* při intervalových datech $(\mathbf{X}, \mathbf{y})$.

Dokážeme o množině B_2 něco říci?

Množina B_2

Zabývejme se množinou

$$B_2 := \{b \in \mathbb{R}^p : X^{\mathrm{T}} X b = X^{\mathrm{T}} y \ \text{ pro některé } X \in \mathbf{X} \text{ a } y \in \mathbf{y}\}. \qquad (2)$$

Množina B_2 je jednou z (více) možných formalizací pojmu *odhad „modelu"* $\mathbf{y} = \mathbf{X}\beta + \varepsilon$ *pomocí nejmenších čtverců* při intervalových datech $(\mathbf{X}, \mathbf{y})$.

Dokážeme o množině B_2 něco říci?

- Množina B_2 může být složitá: např. nemusí být konvexní. Obecně je obtížné podat jiný popis množiny B_2 než pomocí definice (2).

Množina B_2

Zabývejme se množinou

$$B_2 := \{b \in \mathbb{R}^p : X^{\mathrm{T}}Xb = X^{\mathrm{T}}y \ \text{ pro některé } X \in \mathbf{X} \text{ a } y \in \mathbf{y}\}. \qquad (2)$$

Množina B_2 je jednou z (více) možných formalizací pojmu *odhad „modelu"* $\mathbf{y} = \mathbf{X}\beta + \varepsilon$ *pomocí nejmenších čtverců* při intervalových datech $(\mathbf{X}, \mathbf{y})$.

Dokážeme o množině B_2 něco říci?

- Množina B_2 může být složitá: např. nemusí být konvexní. Obecně je obtížné podat jiný popis množiny B_2 než pomocí definice (2).
- I tak má smysl se ptát, jak množina B_2 (alespoň přibližně) vypadá: například hledat její aproximaci pomocí jednoduchých geometrických objektů.

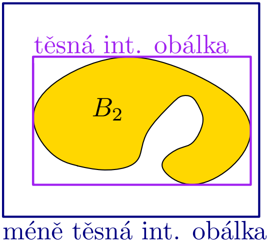
O co se můžeme pokoušet...

O co se můžeme pokoušet...

Můžeme se pokoušet hledat

- intervalovou obálku množiny B_2: je to intervalový vektor $\mathbf{b} \in \mathbb{IR}^p$ splňující $\mathbf{b} \supseteq B_2$;
- těsnou intervalovou obálku množiny B_2: je to intervalová obálka $\mathbf{b}$, pro niž platí, že žádný intervalový vektor $\mathbf{b}' \subsetneqq \mathbf{b}$ nesplňuje $\mathbf{b}' \supseteq B_2$;
- elipsoidovou obálku množiny B_2: je to elipsa $\mathcal{E}$ splňující $\mathcal{E} \supseteq B_2$
- apod.

Poznámka. Zajímavá je především *těsná* intervalová obálka $\mathbf{b} = [\underline{b}, \overline{b}]$: říká, v jakých mezích se mohou pohybovat odhady regresních koeficientů při perturbaci dat (X, y) v intervalech $(\mathbf{X}, \mathbf{y})$.



Položme si mnohem „základnější“ otázku...

Položme si mnohem „základnější“ otázku...

Abychom byli schopni zkonstruovat *vůbec nějakou* obálku množiny

$$B_2 = \{b \in \mathbb{R}^p : X^{\mathrm{T}}Xb = X^{\mathrm{T}}y \ \text{ pro některé } X \in \mathbf{X} \text{ a } y \in \mathbf{y}\},$$

musíme být schopni přinejmenším zodpovědět otázku:

Položme si mnohem „základnější“ otázku…

Abychom byli schopni zkonstruovat *vůbec nějakou* obálku množiny

$$B_2 = \{b \in \mathbb{R}^p : X^{\mathrm{T}}Xb = X^{\mathrm{T}}y \ \text{ pro některé } X \in \mathbf{X} \text{ a } y \in \mathbf{y}\},$$

musíme být schopni přinejmenším zodpovědět otázku:

$$\textbf{je množina } B_2 \textbf{ omezená?} \tag{3}$$

Abychom byli schopni zkonstruovat *vůbec nějakou* obálku množiny

$$B_2 = \{b \in \mathbb{R}^p : X^{\mathrm{T}} X b = X^{\mathrm{T}} y \ \text{ pro některé } X \in \mathbf{X} \text{ a } y \in \mathbf{y}\},$$

musíme být schopni přinejmenším zodpovědět otázku:

$$\textbf{je množina } B_2 \textbf{ omezená?} \tag{3}$$

Lemma. Množina B_2 je neomezená, právě když existuje matice $X \in \mathbf{X}$ neplné sloupcové hodnosti.

$\Longrightarrow$ Hledáme-li algoritmus, který zodpoví otázku (3), hledáme algoritmus, který rozhodne, zdali existuje matice $X \in \mathbf{X}$ neplné sloupcové hodnosti.

Přesnější formulace problému

Přesnější formulace problému

Řešíme problém:

- jsou dány *racionální* matice $\underline{X}, \overline{X}$;
- úkolem je rozhodnout (= říci ANO/NE), zdali existuje matice $X \in \mathbf{X} := [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti. (Není nutně úkolem ji přímo najít, existuje-li.)

Přesnější formulace problému

Řešíme problém:

- jsou dány *racionální* matice $\underline{X}, \overline{X}$;
- úkolem je rozhodnout (= říci ANO/NE), zdali existuje matice $X \in \mathbf{X} := [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti. (Není nutně úkolem ji přímo najít, existuje-li.)

Zkusme hledat *nějaký* algoritmus, který najde správnou odpověď ANO/NE; zatím se neohlížejme na jeho efektivitu (= výpočetní čas).

Přesnější formulace problému

Řešíme problém:

- jsou dány *racionální* matice $\underline{X}, \overline{X}$;
- úkolem je rozhodnout (= říci ANO/NE), zdali existuje matice $X \in \mathbf{X} := [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti. (Není nutně úkolem ji přímo najít, existuje-li.)

Zkusme hledat *nějaký* algoritmus, který najde správnou odpověď ANO/NE; zatím se neohlížejme na jeho efektivitu (= výpočetní čas).

Lemma. Jestliže existuje matice $X \in \mathbf{X}$ neplné sloupcové hodnosti, pak existuje *racionální* matice $X \in \mathbf{X}$ neplné sloupcové hodnosti.

Je problém „$\exists X \in \mathbf{X}$ n.s.h.“ rozhodnutelný?

První nápad: hrubá síla. Systém racionálních matic $X \in \mathbf{X}$ je spočetný. Lze je postupně enumerovat a čekat, zdali narazíme na matici neplné sloupcové hodnosti. Obdržíme tak algoritmus, který není příliš uspokojivý:

(a) Jestliže odpověď je ANO, pak algoritmus skončí a odpoví „`ANO`“;

(b) Jestliže odpověď je NE, pak algoritmus počítá věčně.

Je problém „$\exists X \in \mathbf{X}$ n.s.h.“ rozhodnutelný?

První nápad: hrubá síla. Systém racionálních matic $X \in \mathbf{X}$ je spočetný. Lze je postupně enumerovat a čekat, zdali narazíme na matici neplné sloupcové hodnosti. Obdržíme tak algoritmus, který není příliš uspokojivý:

(a) Jestliže odpověď je ANO, pak algoritmus skončí a odpoví „`ANO`“;

(b) Jestliže odpověď je NE, pak algoritmus počítá věčně.

Prokázali jsme, že problém je rekursivně spočetný.

Je problém „$\exists X \in \mathbf{X}$ n.s.h.“ rozhodnutelný?

První nápad: hrubá síla. Systém racionálních matic $X \in \mathbf{X}$ je spočetný. Lze je postupně enumerovat a čekat, zdali narazíme na matici neplné sloupcové hodnosti. Obdržíme tak algoritmus, který není příliš uspokojivý:

(a) Jestliže odpověď je ANO, pak algoritmus skončí a odpoví „`ANO`“;
(b) Jestliže odpověď je NE, pak algoritmus počítá věčně.

Prokázali jsme, že problém je rekursivně spočetný.

Rádi bychom získali algoritmus, pro který namísto (b) platí

(b') Jestliže odpověď je NE, pak algoritmus skončí a odpoví „`NE`“.

Je problém „$\exists X \in \mathbf{X}$ n.s.h.“ rozhodnutelný?

První nápad: hrubá síla. Systém racionálních matic $X \in \mathbf{X}$ je spočetný. Lze je postupně enumerovat a čekat, zdali narazíme na matici neplné sloupcové hodnosti. Obdržíme tak algoritmus, který není příliš uspokojivý:

(a) Jestliže odpověď je ANO, pak algoritmus skončí a odpoví „`ANO`“;

(b) Jestliže odpověď je NE, pak algoritmus počítá věčně.

Prokázali jsme, že problém je rekursivně spočetný.

Rádi bychom získali algoritmus, pro který namísto (b) platí

(b') Jestliže odpověď je NE, pak algoritmus skončí a odpoví „`NE`“.

Je-li to možné, je náš problém rozhodnutelný (rekursivní); jinak je nerozhodnutelný (nerekursivní).

Je problém „$\exists X \in \mathbf{X}$ n.s.h.“ rozhodnutelný?

První nápad: hrubá síla. Systém racionálních matic $X \in \mathbf{X}$ je spočetný. Lze je postupně enumerovat a čekat, zdali narazíme na matici neplné sloupcové hodnosti. Obdržíme tak algoritmus, který není příliš uspokojivý:

(a) Jestliže odpověď je ANO, pak algoritmus skončí a odpoví „`ANO`“;

(b) Jestliže odpověď je NE, pak algoritmus počítá věčně.

Prokázali jsme, že problém je rekursivně spočetný.

Rádi bychom získali algoritmus, pro který namísto (b) platí

(b') Jestliže odpověď je NE, pak algoritmus skončí a odpoví „`NE`“.

Je-li to možné, je náš problém rozhodnutelný (rekursivní); jinak je nerozhodnutelný (nerekursivní).

Může se stát, že náš problém je nerozhodnutelný? Tedy, že pro něj neexistují žádné algoritmy? Matematika je plná nerozhodnutelných problémů…

Intermezzo: příklady nerozhodnutelných problémů

Intermezzo: příklady nerozhodnutelných problémů

- **Nulové body funkcí.**
  - **Zadání:** funkce $f(x) : \mathbb{R} \longrightarrow \mathbb{R}$ složená z konstant, $+, -, \times, \sin(\cdot)$.
  - **Úkol:** rozhodnout, zdali existuje $x_0 \in \mathbb{R}$ splňující $f(x_0) = 0$.

Intermezzo: příklady nerozhodnutelných problémů

- **Nulové body funkcí.**
  - **Zadání:** funkce $f(x) : \mathbb{R} \longrightarrow \mathbb{R}$ složená z konstant, $+, -, \times, \sin(\cdot)$.
  - **Úkol:** rozhodnout, zdali existuje $x_0 \in \mathbb{R}$ splňující $f(x_0) = 0$.
- **Konvergence integrálu.**
  - **Zadání:** funkce $f(x) : \mathbb{R} \longrightarrow \mathbb{R}$ složená z konstant, $+, -, \times, \div, \sin(\cdot)$.
  - **Úkol:** rozhodnout, zdali $\int_{-\infty}^{\infty} f(x)\, dx$ konverguje.

Intermezzo: příklady nerozhodnutelných problémů

- **Nulové body funkcí.**
  - **Zadání:** funkce $f(x) : \mathbb{R} \longrightarrow \mathbb{R}$ složená z konstant, $+, -, \times, \sin(\cdot)$.
  - **Úkol:** rozhodnout, zdali existuje $x_0 \in \mathbb{R}$ splňující $f(x_0) = 0$.
- **Konvergence integrálu.**
  - **Zadání:** funkce $f(x) : \mathbb{R} \longrightarrow \mathbb{R}$ složená z konstant, $+, -, \times, \div, \sin(\cdot)$.
  - **Úkol:** rozhodnout, zdali $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x$ konverguje.
- **Diofantické rovnice** [tzv. Matijasevičova věta (Matijasevič, 1970); desátý Hilbertův problém (Hilbert, 1900)].
  - **Zadání:** polynom $p(x_1, \ldots, x_9)$ s celočíselnými koeficienty.
  - **Úkol:** rozhodnout, zdali existují $x_1^*, \ldots, x_9^* \in \mathbb{Z}$ splňující $p(x_1^*, \ldots, x_9^*) = 0$.

Intermezzo: příklady nerozhodnutelných problémů

- **Nulové body funkcí.**
  - **Zadání:** funkce $f(x) : \mathbb{R} \longrightarrow \mathbb{R}$ složená z konstant, $+, -, \times, \sin(\cdot)$.
  - **Úkol:** rozhodnout, zdali existuje $x_0 \in \mathbb{R}$ splňující $f(x_0) = 0$.
- **Konvergence integrálu.**
  - **Zadání:** funkce $f(x) : \mathbb{R} \longrightarrow \mathbb{R}$ složená z konstant, $+, -, \times, \div, \sin(\cdot)$.
  - **Úkol:** rozhodnout, zdali $\int_{-\infty}^{\infty} f(x)\, dx$ konverguje.
- **Diofantické rovnice** [tzv. Matijasevičova věta (Matijasevič, 1970); desátý Hilbertův problém (Hilbert, 1900)].
  - **Zadání:** polynom $p(x_1, \ldots, x_9)$ s celočíselnými koeficienty.
  - **Úkol:** rozhodnout, zdali existují $x_1^*, \ldots, x_9^* \in \mathbb{Z}$ splňující $p(x_1^*, \ldots, x_9^*) = 0$.
- **Maticová smrt.**
  - **Zadání:** celočíselné čtvercové matice $A_1, \ldots, A_n$.
  - **Úkol:** rozhodnout, zdali lze z matic $A_1, \ldots, A_n$ násobením (v libovolném pořadí, opakování povoleno) obdržet nulovou matici.

Intermezzo: příklady nerozhodnutelných problémů (pokrač.)

- **Celočíselné programování s kvadratickými omezeními.**
  - *Celočíselné lineární programování* je optimalizační problém

$$\max\{c^{\mathrm{T}}x : \quad Ax \le b, \quad x \in \mathbb{Z}^p\}.$$

  - Připustíme-li vedle lineárních omezení $Ax \le b$ také kvadratická omezení (tj. připustíme-li součiny dvou proměnných), je otázka „je úloha přípustná?“ nerozhodnutelná.

Intermezzo: příklady nerozhodnutelných problémů (pokrač.)

- **Celočíselné programování s kvadratickými omezeními.**
  - *Celočíselné lineární programování* je optimalizační problém

$$\max\{c^{\mathrm{T}}x : \quad Ax \leq b, \quad x \in \mathbb{Z}^p\}.$$

  - Připustíme-li vedle lineárních omezení $Ax \leq b$ také kvadratická omezení (tj. připustíme-li součiny dvou proměnných), je otázka „je úloha přípustná?“ nerozhodnutelná.
- **Dokazatelnost** [tzv. Gödelova věta, (Gödel, 1931)].
  - **Zadání:** tvrzení (= uzavřená formule množinového jazyka) φ.
  - **Úkol:** rozhodnout, zdali tvrzení φ je dokazatelné v teorii množin.

Intermezzo: příklady nerozhodnutelných problémů (pokrač.)

- **Celočíselné programování s kvadratickými omezeními.**
  - *Celočíselné lineární programování* je optimalizační problém

$$\max\{c^{\mathrm{T}}x : \quad Ax \leq b, \quad x \in \mathbb{Z}^p\}.$$

  - Připustíme-li vedle lineárních omezení $Ax \leq b$ také kvadratická omezení (tj. připustíme-li součiny dvou proměnných), je otázka „je úloha připustná?“ nerozhodnutelná.
- **Dokazatelnost** [tzv. Gödelova věta, (Gödel, 1931)].
  - **Zadání:** tvrzení (= uzavřená formule množinového jazyka) φ.
  - **Úkol:** rozhodnout, zdali tvrzení φ je dokazatelné v teorii množin.
- **Halting problem.**
  - **Zadání:** text programu P a data x.
  - **Úkol:** rozhodnout, zdali výpočet programu P nad daty x skončí, anebo zdali počítá věčně.

Intermezzo: příklady nerozhodnutelných problémů (pokrač.)

- **Celočíselné programování s kvadratickými omezeními.**
  - *Celočíselné lineární programování* je optimalizační problém

    $$\max\{c^{\mathrm{T}}x : \quad Ax \leq b, \quad x \in \mathbb{Z}^p\}.$$

  - Připustíme-li vedle lineárních omezení $Ax \leq b$ také kvadratická omezení (tj. připustíme-li součiny dvou proměnných), je otázka „je úloha připustná?“ nerozhodnutelná.
- **Dokazatelnost** [tzv. Gödelova věta, (Gödel, 1931)].
  - **Zadání:** tvrzení (= uzavřená formule množinového jazyka) φ.
  - **Úkol:** rozhodnout, zdali tvrzení φ je dokazatelné v teorii množin.
- **Halting problem.**
  - **Zadání:** text programu P a data x.
  - **Úkol:** rozhodnout, zdali výpočet programu P nad daty x skončí, anebo zdali počítá věčně.
- **Další příklady.** Jsou dány dvě prezentace grup (G_1, R_1) a (G_2, R_2); jsou jimi prezentované grupy isomorfní? Má daná nula-jedničková posloupnost velkou kolmogorovskou složitost? Atd.

Náš problém „$\exists X \in \mathbf{X}$ n.s.h.“ JE rozhodnutelný!

Náš problém „$\exists X \in \mathbf{X}$ n.s.h." JE rozhodnutelný!

Pozorování. Množina B_2 je neomezená

$$\longleftrightarrow \text{existuje } X \in [\underline{X}, \overline{X}] \text{ neplné sloupcové hodnosti}$$

$$\longleftrightarrow (\exists X)[\underline{X} \leq X \leq \overline{X} \ \ \& \ \ \det X^{\mathrm{T}} X = 0]. \tag{4}$$

Náš problém „$\exists X \in \mathbf{X}$ n.s.h." JE rozhodnutelný!

Pozorování. Množina B_2 je neomezená

$$\longleftrightarrow \text{existuje } X \in [\underline{X}, \overline{X}] \text{ neplné sloupcové hodnosti}$$

$$\longleftrightarrow (\exists X)[\underline{X} \leq X \leq \overline{X} \ \ \& \ \ \det X^{\mathrm{T}} X = 0]. \tag{4}$$

Výraz (4) je výraz jazyka teorie (uspořádaných) těles.

Náš problém „$\exists X \in \mathbf{X}$ n.s.h.“ JE rozhodnutelný!

Pozorování. Množina B_2 je neomezená

$$\begin{aligned} &\longleftrightarrow \text{existuje } X \in [\underline{X}, \overline{X}] \text{ neplné sloupcové hodnosti} \\ &\longleftrightarrow (\exists X)[\underline{X} \leq X \leq \overline{X} \ \ \& \ \ \det X^{\mathrm{T}} X = 0]. \end{aligned} \tag{4}$$

Výraz (4) je výraz jazyka teorie (uspořádaných) těles.

Teorie reálně uzavřených těles (RCF, Real Closed Fields) je teorie obsahující

- běžné axiomy teorie uspořádaných těles (tj. vlastnosti $+$, $\times$, $\leq$, 0, 1),
- axiomy zajišťující, že pro každý polynom platí Bolzanova věta. [*Bolzanova věta říká: spojitá funkce $f(x)$ má v intervalu $(\underline{a}, \overline{a})$ nulový bod, kdykoliv platí $f(\underline{a}) < 0$ a $f(\overline{a}) > 0$.*]

Náš problém „$\exists X \in \mathbf{X}$ n.s.h." JE rozhodnutelný!

Pozorování. Množina B_2 je neomezená

$$\longleftrightarrow \text{existuje } X \in [\underline{X}, \overline{X}] \text{ neplné sloupcové hodnosti}$$

$$\longleftrightarrow (\exists X)[\underline{X} \leq X \leq \overline{X} \;\; \& \;\; \det X^{\mathrm{T}} X = 0]. \tag{4}$$

Výraz (4) je výraz jazyka teorie (uspořádaných) těles.

Teorie reálně uzavřených těles (RCF, Real Closed Fields) je teorie obsahující

- běžné axiomy teorie uspořádaných těles (tj. vlastnosti $+$, $\times$, $\leq$, 0, 1),
- axiomy zajišťující, že pro každý polynom platí Bolzanova věta. [*Bolzanova věta říká: spojitá funkce $f(x)$ má v intervalu $(\underline{a}, \overline{a})$ nulový bod, kdykoliv platí $f(\underline{a}) < 0$ a $f(\overline{a}) > 0$.*]

Věta (Tarski, 1956). Teorie **RCF** je úplná.

Náš problém „$\exists X \in \mathbf{X}$ n.s.h.“ JE rozhodnutelný!

Pozorování. Množina B_2 je neomezená

$$\longleftrightarrow \text{ existuje } X \in [\underline{X}, \overline{X}] \text{ neplné sloupcové hodnosti}$$

$$\longleftrightarrow (\exists X)[\underline{X} \leq X \leq \overline{X} \ \ \& \ \ \det X^{\mathrm{T}} X = 0]. \qquad (4)$$

Výraz (4) je výraz jazyka teorie (uspořádaných) těles.

Teorie reálně uzavřených těles (RCF, Real Closed Fields) je teorie obsahující

- běžné axiomy teorie uspořádaných těles (tj. vlastnosti $+$, $\times$, $\leq$, 0, 1),
- axiomy zajišťující, že pro každý polynom platí Bolzanova věta. [*Bolzanova věta říká: spojitá funkce $f(x)$ má v intervalu $(\underline{a}, \overline{a})$ nulový bod, kdykoliv platí $f(\underline{a}) < 0$ a $f(\overline{a}) > 0$.*]

Věta (Tarski, 1956). Teorie **RCF** je úplná.

Důsledek. O pravdivosti tvrzení (4) lze rozhodnout algoritmicky. [*Stačí enumerovat všechny důkazy teorie* **RCF** *a čekat, zdali najdeme důkaz tvrzení (4), anebo jeho negace.*]

Rychlost algoritmů

Nepatrná vada na kráse.

Nepatrná vada na kráse. Metoda založená na Tarského větě vyžaduje extrémní výpočetní čas: až $\approx 2^{2^L}$, kde

$$L := \text{bitová velikost zápisu } \underline{X}, \overline{X}.$$

(Každé racionální číslo v maticích $\underline{X}$ a $\overline{X}$ zapisujeme jako trojici [znaménko, dvojkový zápis čitatele, dvojkový zápis jmenovatele].)

Je možné najít lepší algoritmus?

Rychlost algoritmů

Nepatrná vada na kráse. Metoda založená na Tarského větě vyžaduje extrémní výpočetní čas: až $\approx 2^{2^L}$, kde

$$L := \text{bitová velikost zápisu } \underline{X}, \overline{X}.$$

(Každé racionální číslo v maticích $\underline{X}$ a $\overline{X}$ zapisujeme jako trojici [znaménko, dvojkový zápis čitatele, dvojkový zápis jmenovatele].)

Je možné najít lepší algoritmus?

Věta. Otázka „existuje $X \in [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti?“ je v **NP**.

Důsledek. Existuje algoritmus, který nalezne odpověď v čase $\leq 2^{\text{polynom}(L)}$.

Intermezzo: třída **NP**

„Objektem“ rozumíme konečnou posloupnost nul a jedniček (bitů).

Intermezzo: třída **NP**

„Objektem“ rozumíme konečnou posloupnost nul a jedniček (bitů).

Neformální definice. **NP** je třída otázek tvaru

$$\text{„je pravda, že předložený objekt } x \text{ má vlastnost } \varphi\text{?”}, \tag{5}$$

pro které platí: existuje rychlý algoritmus $\mathrm{PROOF}(x, y)$ s touto vlastností:

- jestliže odpověď na otázku (5) zní ANO, pak existuje krátký objekt y takový, že $\mathrm{PROOF}(x, y) = ANO$,
- jestliže odpověď na otázku (5) zní NE, pak každý krátký objekt y má vlastnost $\mathrm{PROOF}(x, y) = NE$.

Volně řečeno: Otázka je v **NP**, právě když *pozitivní odpověď lze doložit krátkým, snadno ověřitelným důkazem.* (Nemusí ale být snadné takový důkaz efektivně najít.)

Příklady problémů v **NP**

Příklady problémů v **NP**

- **Boolovská splnitelnost.**
  - Otázka: Je daná výroková formule splnitelná?
  - Důkaz prokazující pozitivní odpověď: jedno konkrétní splňující ohodnocení.

Příklady problémů v **NP**

- **Boolovská splnitelnost.**
  - Otázka: Je daná výroková formule splnitelná?
  - Důkaz prokazující pozitivní odpověď: jedno konkrétní splňující ohodnocení.
- **Přípustnost celočíselného lineárního programování.**
  - Otázka: Má daný lineární systém $Ax \leq b$ nějaké celočíselné řešení x?
  - Důkaz prokazující pozitivní odpověď: jedno konkrétní řešení x.

Příklady problémů v **NP**

- **Boolovská splnitelnost.**
  - Otázka: Je daná výroková formule splnitelná?
  - Důkaz prokazující pozitivní odpověď: jedno konkrétní splňující ohodnocení.
- **Přípustnost celočíselného lineárního programování.**
  - Otázka: Má daný lineární systém $Ax \leq b$ nějaké celočíselné řešení x?
  - Důkaz prokazující pozitivní odpověď: jedno konkrétní řešení x.
- **Faktorizace.**
  - Otázka: Je dané přirozené číslo k složené?
  - Důkaz prokazující pozitivní odpověď: netriviální dělitel čísla k.

Příklady problémů v **NP**

- **Boolovská splnitelnost.**
  - Otázka: Je daná výroková formule splnitelná?
  - Důkaz prokazující pozitivní odpověď: jedno konkrétní splňující ohodnocení.
- **Přípustnost celočíselného lineárního programování.**
  - Otázka: Má daný lineární systém $Ax \leq b$ nějaké celočíselné řešení x?
  - Důkaz prokazující pozitivní odpověď: jedno konkrétní řešení x.
- **Faktorizace.**
  - Otázka: Je dané přirozené číslo k složené?
  - Důkaz prokazující pozitivní odpověď: netriviální dělitel čísla k.
- **Jednoduché diofantické rovnice.**
  - Otázka: Má diofantická rovnice $ax^2 + by = c$ nějaké přirozené řešení x, y?
  - Důkaz prokazující pozitivní odpověď: jedno konkrétní řešení x, y.

Co přesněji znamená „krátký“ **NP**-důkaz?

Symbolem size(x) označme bitovou velikost objektu x.

Co přesněji znamená „krátký“ **NP**-důkaz?

Symbolem size(x) označme bitovou velikost objektu x.

Upřesnění. Říkáme-li, že důkaz y prokazující, že platí $\varphi(x)$, je krátký, myslíme tím

$$\text{size}(y) \leq \text{polynom}(\text{size}(x)).$$

Odtud je jasné, že každý problém v **NP** je rozhodnutelný v čase $2^{\text{polynom}(\text{size}(x))}$ — stačí prohledat všechny možné krátké důkazy.

Co přesněji znamená „krátký“ **NP**-důkaz?

Symbolem size(x) označme bitovou velikost objektu x.

Upřesnění. Říkáme-li, že důkaz y prokazující, že platí $\varphi(x)$, je krátký, myslíme tím

$$\text{size}(y) \le \text{polynom}(\text{size}(x)).$$

Odtud je jasné, že každý problém v **NP** je rozhodnutelný v čase $2^{\text{polynom}(\text{size}(x))}$ — stačí prohledat všechny možné krátké důkazy.

Cíl. Chtěli bychom seznam příkladů **NP**-otázek rozšířit:

- Otázka: Existuje $X \in [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti?
- Důkaz prokazující pozitivní odpověď: některá matice $X_0 \in [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti.

Co přesněji znamená „krátký“ **NP**-důkaz?

Symbolem size(x) označme bitovou velikost objektu x.

Upřesnění. Říkáme-li, že důkaz y prokazující, že platí $\varphi(x)$, je krátký, myslíme tím

$$\text{size}(y) \leq \text{polynom}(\text{size}(x)).$$

Odtud je jasné, že každý problém v **NP** je rozhodnutelný v čase $2^{\text{polynom}(\text{size}(x))}$ — stačí prohledat všechny možné krátké důkazy.

Cíl. Chtěli bychom seznam příkladů **NP**-otázek rozšířit:

- Otázka: Existuje $X \in [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti?
- Důkaz prokazující pozitivní odpověď: některá matice $X_0 \in [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti.

Problém. Abychom byli korektní, je třeba prokázat: existuje polynom q takový, že: pro libovolné racionální matice $\underline{X} \leq \overline{X}$ platí, že obsahuje-li $[\underline{X}, \overline{X}]$ matici neplné sloupcové hodnosti, pak vždy také obsahuje racionální matici $X_0^* \in [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti splňující

$$\text{size}(X_0^*) \leq q(\text{size}(\underline{X}) + \text{size}(\overline{X})).$$

Co přesněji znamená „krátký“ **NP**-důkaz?

Symbolem size(x) označme bitovou velikost objektu x.

Upřesnění. Říkáme-li, že důkaz y prokazující, že platí $\varphi(x)$, je krátký, myslíme tím

$$\text{size}(y) \leq \text{polynom}(\text{size}(x)).$$

Odtud je jasné, že každý problém v **NP** je rozhodnutelný v čase $2^{\text{polynom}(\text{size}(x))}$ — stačí prohledat všechny možné krátké důkazy.

Cíl. Chtěli bychom seznam příkladů **NP**-otázek rozšířit:

- Otázka: Existuje $X \in [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti?
- Důkaz prokazující pozitivní odpověď: některá matice $X_0 \in [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti.

Problém. Abychom byli korektní, je třeba prokázat: existuje polynom q takový, že: pro libovolné racionální matice $\underline{X} \leq \overline{X}$ platí, že obsahuje-li $[\underline{X}, \overline{X}]$ matici neplné sloupcové hodnosti, pak vždy také obsahuje racionální matici $X_0^* \in [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti splňující

$$\text{size}(X_0^*) \leq q(\text{size}(\underline{X}) + \text{size}(\overline{X})).$$

To je ovšem těžké. Zkusme to jinak...

Intervalové rovnice

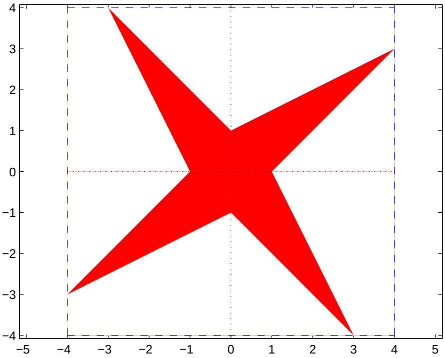
Intervalové rovnice

Definice. Nechť $\mathbf{A} \in \mathbb{IR}^{n \times p}$ a $\mathbf{b} \in \mathbb{IR}^{n}$. Řekneme, že $z_0 \in \mathbb{R}^p$ je řešením systému $\mathbf{A}z = \mathbf{b}$, jestliže existují $A \in \mathbf{A}$ a $b \in \mathbf{b}$ takové, že $Az_0 = b$.

Definice. Nechť $\mathbf{A} \in \mathbb{IR}^{n\times p}$ a $\mathbf{b} \in \mathbb{IR}^{n}$. Řekneme, že $z_0 \in \mathbb{R}^p$ je řešením systému $\mathbf{A}z = \mathbf{b}$, jestliže existují $A \in \mathbf{A}$ a $b \in \mathbf{b}$ takové, že $Az_0 = b$.

Příklad (Barth & Nuding, 1974). Množina řešení intervalové soustavy

$$\begin{pmatrix} [2,4] & [-2,1] \\ [-1,2] & [2,4] \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} [-2,2] \\ [-2,2] \end{pmatrix}$$



Intervalové rovnice

Věta (Oettli-Prager). Označme

$$A^c := \tfrac{1}{2}(\underline{A} + \overline{A}), \qquad A^\Delta := \tfrac{1}{2}(\overline{A} - \underline{A}),$$
$$b^c := \tfrac{1}{2}(\underline{b} + \overline{b}), \qquad b^\Delta := \tfrac{1}{2}(\overline{b} - \underline{b}).$$

Vektor $z \in \mathbb{R}^p$ je řešením systému $\mathbf{A}z = \mathbf{b}$, právě když platí

$$|A^c z - b^c| \leq A^\Delta |z| + b^\Delta.$$

Intervalové rovnice

Věta (Oettli-Prager). Označme

$$A^c := \tfrac{1}{2}(\underline{A} + \overline{A}), \qquad A^\Delta := \tfrac{1}{2}(\overline{A} - \underline{A}),$$
$$b^c := \tfrac{1}{2}(\underline{b} + \overline{b}), \qquad b^\Delta := \tfrac{1}{2}(\overline{b} - \underline{b}).$$

Vektor $z \in \mathbb{R}^p$ je řešením systému $\mathbf{A}z = \mathbf{b}$, právě když platí

$$|A^c z - b^c| \le A^\Delta |z| + b^\Delta.$$

Důsledek. Nechť $s \in \{\pm 1\}^p$. Nechť $\mathbb{R}^p_s$ označuje orthant $\{x \in \mathbb{R}^p : \mathrm{diag}(s)x \ge 0\}$. Množina řešení systému $\mathbf{A}z = \mathbf{b}$ ležících v orthantu $\mathbb{R}^p_s$ je tvaru

$$\{z \in \mathbb{R}^p : \ (A^c - A^\Delta \mathrm{diag}(s))z \le \underline{b},$$
$$(-A^c - A^\Delta \mathrm{diag}(s))z \le -\overline{b}, \ \mathrm{diag}(s)z \ge 0\}.$$

Náš problém „$\exists X \in \mathbf{X}$ n.s.h." je v **NP**

Náš problém „$\exists X \in \mathbf{X}$ n.s.h." je v **NP**

Pozorování. Množina B_2 je neomezená

$$\longleftrightarrow \text{existuje } X \in [\underline{X}, \overline{X}] \text{ neplné sloupcové hodnosti}$$

$$\longleftrightarrow \text{intervalový systém } \mathbf{X}z = 0 \text{ má nenulové řešení.} \qquad (6)$$

Náš problém „$\exists X \in \mathbf{X}$ n.s.h.“ je v **NP**

Pozorování. Množina B_2 je neomezená

$$\longleftrightarrow \text{ existuje } X \in [\underline{X}, \overline{X}] \text{ neplné sloupcové hodnosti}$$

$$\longleftrightarrow \text{ intervalový systém } \mathbf{X}z = 0 \text{ má nenulové řešení.} \tag{6}$$

Podle důsledku Oettliho-Pragerovy věty lze podmínku (6) testovat pro každý orthant zvlášť: v daném orthantu $s \in \{\pm 1\}^p$ můžeme existenci nenulového vektoru v polyedru

$$(X^c - X^\Delta \operatorname{diag}(s))z \leq 0, \ (-X^c - X^\Delta \operatorname{diag}(s))z \leq 0, \ \operatorname{diag}(s)z \geq 0 \tag{7}$$

testovat pomocí lineárního programování. A lineární programování je rychlá (= polynomiální) procedura — stačí užít elipsoidový algoritmus či vhodnou metodu vnitřního bodu.

Náš problém „$\exists X \in \mathbf{X}$ n.s.h.“ je v **NP**

Pozorování. Množina B_2 je neomezená

$\longleftrightarrow$ existuje $X \in [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti

$\longleftrightarrow$ intervalový systém $\mathbf{X}z = 0$ má nenulové řešení. (6)

Podle důsledku Oettliho-Pragerovy věty lze podmínku (6) testovat pro každý orthant zvlášť: v daném orthantu $s \in \{\pm 1\}^p$ můžeme existenci nenulového vektoru v polyedru

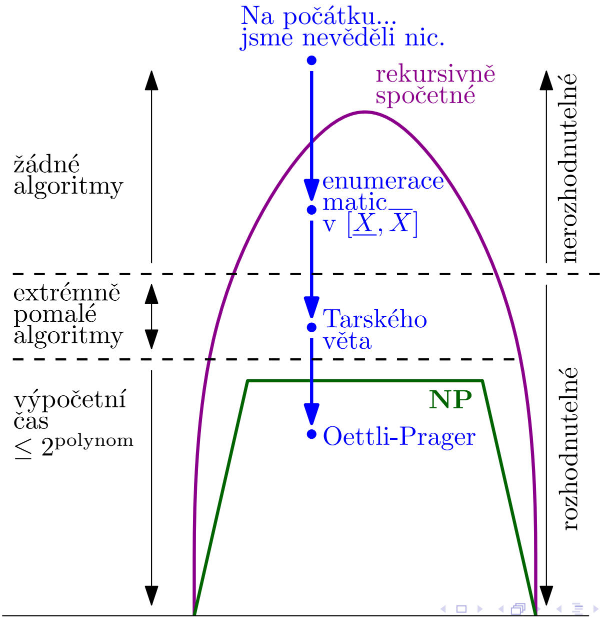
$$(X^c - X^\Delta \operatorname{diag}(s))z \leq 0, \ (-X^c - X^\Delta \operatorname{diag}(s))z \leq 0, \ \operatorname{diag}(s)z \geq 0 \quad (7)$$

testovat pomocí lineárního programování. A lineární programování je rychlá (= polynomiální) procedura — stačí užít elipsoidový algoritmus či vhodnou metodu vnitřního bodu.

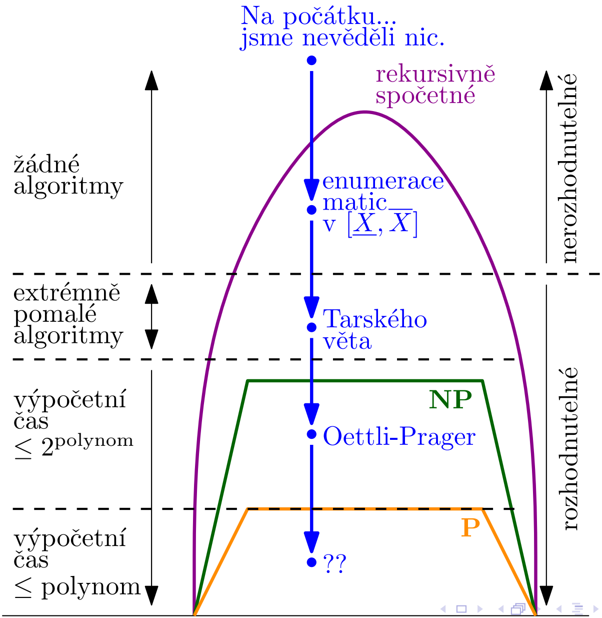
Dosáhli jsme cíle. Naše otázka je opravdu v **NP**:

- Otázka: Existuje $X \in [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti?
- Důkaz prokazující pozitivní odpověď: vektor $s \in \{\pm 1\}^p$ takový, že orthant $\mathbb{R}^p_s$ obsahuje nenulové řešení systému (7).

Shrnutí: čeho jsme zatím dosáhli



Pokračujme: čeho bychom chtěli dosáhnout



NP-těžkost

Klademe si otázku, zdali by se podařilo najít rychlý algoritmus pro náš problém „existuje $X \in [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti?“.

Rychlý algoritmus = algoritmus pracující v čase $\leq$ polynom(size(x)).

NP-těžkost

Klademe si otázku, zdali by se podařilo najít rychlý algoritmus pro náš problém „existuje $X \in [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti?".

Rychlý algoritmus = algoritmus pracující v čase $\leq$ polynom(size(x)).

Ukážeme, že tento cíl je nerealistický.

NP-těžkost

Klademe si otázku, zdali by se podařilo najít rychlý algoritmus pro náš problém „existuje $X \in [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti?“.

Rychlý algoritmus = algoritmus pracující v čase $\leq$ polynom(size(x)).

Ukážeme, že tento cíl je nerealistický.

Neformální definice. Problém A je **NP**-těžký, jestliže pro něj platí implikace: jestliže problém A má polynomiální algoritmus, pak i každý problém v **NP** má polynomiální algoritmus.

To například znamená: má-li problém A polynomiální algoritmus, pak

- o splnitelnosti výrokové formule s n proměnnými dokážeme rozhodovat rychle (tj. bez konstrukce tabulky pravdivostních hodnot, jež má velikost 2^n),
- úlohy celočíselného lineárního programování dokážeme řešit rychle,
- problém obchodního cestujícího dokážeme řešit rychle atd.

Špatné zprávy

Hypotéza. Jestliže problém A je **NP**-těžký, pak nemá rychlý algoritmus (= není ve třídě **P**).

- *Jde o obtížný otevřený problém známý jako otázka* $\mathbf{P} =^? \mathbf{NP}$*. Obecně panuje konsensus, že tvrzení platí; nadále jej berme za axiom.*

Věta. Problém „existuje $X \in [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti?" je **NP**-těžký.

Špatné zprávy

Hypotéza. Jestliže problém A je **NP**-těžký, pak nemá rychlý algoritmus (= není ve třídě **P**).

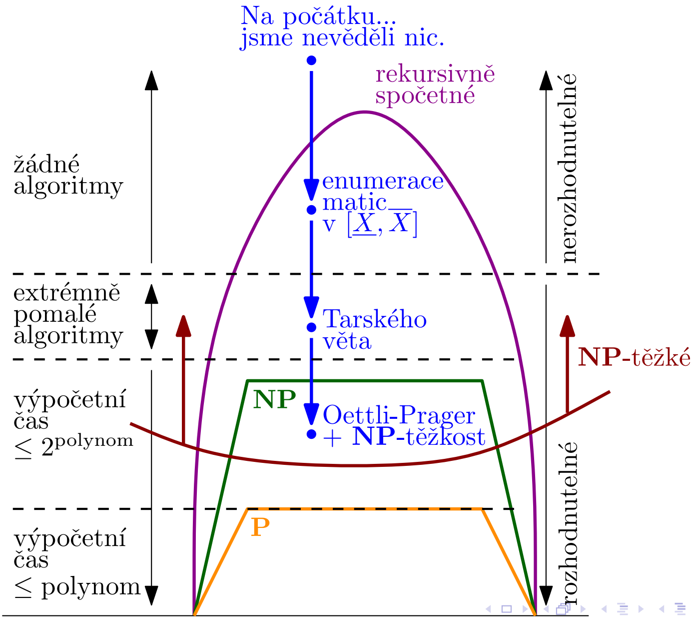
- *Jde o obtížný otevřený problém známý jako otázka* $\mathbf{P} =^? \mathbf{NP}$. *Obecně panuje konsensus, že tvrzení platí; nadále jej berme za axiom.*

Věta. Problém „existuje $X \in [\underline{X}, \overline{X}]$ neplné sloupcové hodnosti?“ je **NP**-těžký.

Poznámky.

- Věta říká, že náš problém nemá polynomiální algoritmy. Může se ale stát, že má algoritmy jen mírně horší než polynomiální a podstatně lepší než exponenciální, např. pracující v čase $n^{\log\log n}$.
- Někdy se formuluje silnější hypotéza, která říká, že *NP*-těžké problémy nemají lepší než *exponenciální* algoritmy. Na pravdivosti této hypotézy ovšem není shoda. Nic to nemění na tom, že lepší než exponenciální algoritmy neznáme.

Co jsme zjistili



Co vyplývá z **NP**-těžkosti

- Nemůžeme čekat, že se nám podaří zkonstruovat rychlé algoritmy pro konstrukci obálek množiny B_2 (ať už jde o těsné intervalové obálky, méně těsné intervalové obálky či elipsoidové obálky).
- Není ale cílem propadnout skepsi. Jsme v podobné situaci, kterou známe z celočíselného lineárního programování: sice víme, že obecně rychlé algoritmy neexistují, ale v konkrétních případech můžeme být úspěšní. (V případě celočíselného lineárního programování lze za příklad uvést totálně unimodulární programy.)
- I u naší otázky je zajímavé začít zkoumat speciální případy problému, které rychlé algoritmy mají.

Co vyplývá z **NP**-těžkosti

- Nemůžeme čekat, že se nám podaří zkonstruovat rychlé algoritmy pro konstrukci obálek množiny B_2 (ať už jde o těsné intervalové obálky, méně těsné intervalové obálky či elipsoidové obálky).
- Není ale cílem propadnout skepsi. Jsme v podobné situaci, kterou známe z celočíselného lineárního programování: sice víme, že obecně rychlé algoritmy neexistují, ale v konkrétních případech můžeme být úspěšní. (V případě celočíselného lineárního programování lze za příklad uvést totálně unimodulární programy.)
- I u naší otázky je zajímavé začít zkoumat speciální případy problému, které rychlé algoritmy mají.

Pohleďme na tři pozoruhodné speciální případy, kdy úspěšně dokážeme učinit krok

$$\mathbf{NP} \downarrow \mathbf{P}.$$

Speciální případ I: omezený počet regresních parametrů

Speciální případ I: omezený počet regresních parametrů

Připomeňme důsledek Oettliho-Pragerovy věty: množina B_2 je neomezená, právě když pro některé $s \in \{\pm 1\}^p$ má lineární systém

$$(X^c - X^\Delta \operatorname{diag}(s))z \leq 0, \; (-X^c - X^\Delta \operatorname{diag}(s))z \leq 0, \; \operatorname{diag}(s)z \geq 0$$

nenulové řešení.

Speciální případ I: omezený počet regresních parametrů

Připomeňme důsledek Oettliho-Pragerovy věty: množina B_2 je neomezená, právě když pro některé $s \in \{\pm 1\}^p$ má lineární systém

$$(X^c - X^\Delta \operatorname{diag}(s))z \leq 0, \; (-X^c - X^\Delta \operatorname{diag}(s))z \leq 0, \; \operatorname{diag}(s)z \geq 0$$

nenulové řešení.

Důsledek. Problém lze řešit v čase

$$2^p \times (\text{čas na lineární programování}). \tag{8}$$

Speciální případ I: omezený počet regresních parametrů

Připomeňme důsledek Oettliho-Pragerovy věty: množina B_2 je neomezená, právě když pro některé $s \in \{\pm 1\}^p$ má lineární systém

$$(X^c - X^\Delta \operatorname{diag}(s))z \leq 0, \ (-X^c - X^\Delta \operatorname{diag}(s))z \leq 0, \ \operatorname{diag}(s)z \geq 0$$

nenulové řešení.

Důsledek. Problém lze řešit v čase

$$2^p \times (\text{čas na lineární programování}). \tag{8}$$

Omezíme-li se na regresní modely s $\leq p_0$ regresními parametry, kde p_0 je pevná konstanta, pak i 2^{p_0} je pevná konstanta, a tudíž (8) je polynom. Ukázali jsme:

Věta. Pro libovolné p_0 platí: omezíme-li se jen na regresní modely s $\leq p_0$ regresními parametry, pak rozhodování o omezenosti množiny B_2 je v **P**.

Speciální případ II: nezáporné regresní parametry

Speciální případ II: nezáporné regresní parametry

Reformulujme jinak důsledek plynoucí z (důsledku) Oettliho--Pragerovy věty. Problém lze řešit v čase

$$\underbrace{(\text{počet orthantů prostoru parametrů } \mathbb{R}^p)}_{=2^p} \times (\text{čas na lineární programování}).$$

Řekněme, že (z teorie) víme, že hodnoty regresních parametrů jsou nezáporné. Pak stačí prozkoumat *jediný* orthant prostoru parametrů. Obdržíme výpočetní čas

$$1 \times (\text{čas na lineární programování}).$$

To je polynom. Ukázali jsme:

Věta. Je-li prostorem parametrů jen nezáporný orthant $\mathbb{R}^p$, pak rozhodování o omezenosti množiny B_2 je v **P**.

Speciální případ III: množina B_2 při $\underline{X} = \overline{X}$

Mějme k dispozici data $(X, \mathbf{y})$, kde X má plnou sloupcovou hodnost. Pak

$$B_2 = \{(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y : \; y \in \mathbf{y}\}.$$

Speciální případ III: množina B_2 při $\underline{X} = \overline{X}$

Mějme k dispozici data $(X, \mathbf{y})$, kde X má plnou sloupcovou hodnost. Pak

$$B_2 = \{(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y : \ y \in \mathbf{y}\}.$$

Geometricky: B_2 je obraz „krychle“ $\mathbf{y} \subseteq \mathbb{R}^n$ v prostoru parametrů $\mathbb{R}^p$ při lineárním zobrazení $v \mapsto (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}v$.

Speciální případ III: množina B_2 při $\underline{X} = \overline{X}$

Mějme k dispozici data $(X, \mathbf{y})$, kde X má plnou sloupcovou hodnost. Pak

$$B_2 = \{(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y : \ y \in \mathbf{y}\}.$$

Geometricky: B_2 je obraz „krychle“ $\mathbf{y} \subseteq \mathbb{R}^n$ v prostoru parametrů $\mathbb{R}^p$ při lineárním zobrazení $v \mapsto (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}v$.

- Množina B_2 je konvexní polyedr speciální struktury, tzv. zonotop. Má zajímavé geometrické vlastnosti, např. každá jeho stěna (libovolné dimenze) je středově symetrická.

Speciální případ III: množina B_2 při $\underline{X} = \overline{X}$

Mějme k dispozici data $(X, \mathbf{y})$, kde X má plnou sloupcovou hodnost. Pak

$$B_2 = \{(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y : \ y \in \mathbf{y}\}.$$

Geometricky: B_2 je obraz „krychle“ $\mathbf{y} \subseteq \mathbb{R}^n$ v prostoru parametrů $\mathbb{R}^p$ při lineárním zobrazení $\upsilon \mapsto (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}\upsilon$.

- Množina B_2 je konvexní polyedr speciální struktury, tzv. zonotop. Má zajímavé geometrické vlastnosti, např. každá jeho stěna (libovolné dimenze) je středově symetrická.
- Snadno zkonstruujeme těsnou intervalovou obálku, například pomocí lineárních programů

  $$\max\{\pm z_i : \ z = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y, \ \underline{y} \le y \le \overline{y}, \ y \in \mathbb{R}^n\}, \quad i = 1, \dots, p.$$

  (jde to dokonce ještě snadněji).

Speciální případ III: množina B_2 při $\underline{X} = \overline{X}$

Mějme k dispozici data $(X, \mathbf{y})$, kde X má plnou sloupcovou hodnost. Pak

$$B_2 = \{(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y : \; y \in \mathbf{y}\}.$$

Geometricky: B_2 je obraz „krychle" $\mathbf{y} \subseteq \mathbb{R}^n$ v prostoru parametrů $\mathbb{R}^p$ při lineárním zobrazení $\upsilon \mapsto (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}\upsilon$.

- Množina B_2 je konvexní polyedr speciální struktury, tzv. zonotop. Má zajímavé geometrické vlastnosti, např. každá jeho stěna (libovolné dimenze) je středově symetrická.
- Snadno zkonstruujeme těsnou intervalovou obálku, například pomocí lineárních programů

  $$\max\{\pm z_i : \;\; z = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y, \; \underline{y} \le y \le \overline{y}, \;\; y \in \mathbb{R}^n\}, \quad i = 1, \ldots, p.$$

  (jde to dokonce ještě snadněji).
- Umíme konstruovat i elipsoidové obálky.

Speciální případ III: množina B_2 při $\underline{X} = \overline{X}$

Mějme k dispozici data $(X, \mathbf{y})$, kde X má plnou sloupcovou hodnost. Pak

$$B_2 = \{(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y : \ y \in \mathbf{y}\}.$$

Geometricky: B_2 je obraz „krychle“ $\mathbf{y} \subseteq \mathbb{R}^n$ v prostoru parametrů $\mathbb{R}^p$ při lineárním zobrazení $v \mapsto (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}v$.

- Množina B_2 je konvexní polyedr speciální struktury, tzv. zonotop. Má zajímavé geometrické vlastnosti, např. každá jeho stěna (libovolné dimenze) je středově symetrická.
- Snadno zkonstruujeme těsnou intervalovou obálku, například pomocí lineárních programů

  $$\max\{\pm z_i : \ \ z = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y, \ \underline{y} \le y \le \overline{y}, \ \ y \in \mathbb{R}^n\}, \quad i = 1, \ldots, p.$$

  (jde to dokonce ještě snadněji).
- Umíme konstruovat i elipsoidové obálky.
- Umíme toho mnohem více.

Konec motivačního příkladu.

Ještě několik příkladů

Nechť symboly $x_1, \ldots, x_n$ označují data. Řekněme pro příklad, že jde o výběr z $N(\mu, \sigma^2)$.

Ještě několik příkladů

Nechť symboly $x_1, \dots, x_n$ označují data. Řekněme pro příklad, že jde o výběr z $N(\mu, \sigma^2)$.

Z algebraického pohledu je jistá statistika $\mathcal{S}$ jen *funkcí dat* $x_1, \dots, x_n$.

Ještě několik příkladů

Nechť symboly $x_1, \ldots, x_n$ označují data. Řekněme pro příklad, že jde o výběr z $N(\mu, \sigma^2)$.

Z algebraického pohledu je jistá statistika $\mathcal{S}$ jen *funkcí dat* $x_1, \ldots, x_n$.

Řekněme dále, že pozorování $x_1, \ldots, x_n$ nejsou k dispozici; máme k dispozici jen intervaly

$$\mathbf{x}_i = [\underline{x}_i, \overline{x}_i], \qquad i = 1, \ldots, n$$

splňující $x_1 \in \mathbf{x}_1, \ldots, x_n \in \mathbf{x}_n$.

Ještě několik příkladů

Nechť symboly $x_1, \dots, x_n$ označují data. Řekněme pro příklad, že jde o výběr z $N(\mu, \sigma^2)$.

Z algebraického pohledu je jistá statistika $\mathcal{S}$ jen *funkcí dat* $x_1, \dots, x_n$.

Řekněme dále, že pozorování $x_1, \dots, x_n$ nejsou k dispozici; máme k dispozici jen intervaly

$$\mathbf{x}_i = [\underline{x}_i, \overline{x}_i], \qquad i = 1, \dots, n$$

splňující $x_1 \in \mathbf{x}_1, \dots, x_n \in \mathbf{x}_n$.

Zkoumejme hodnoty

$$\overline{\mathcal{S}} := \sup\{\mathcal{S}(x_1, \dots, x_n) : x_1 \in \mathbf{x}_1, \dots, x_n \in \mathbf{x}_n\},$$
$$\underline{\mathcal{S}} := \inf\{\mathcal{S}(x_1, \dots, x_n) : x_1 \in \mathbf{x}_1, \dots, x_n \in \mathbf{x}_n\}.$$

Ještě několik příkladů

Nechť symboly $x_1, \ldots, x_n$ označují data. Řekněme pro příklad, že jde o výběr z $N(\mu, \sigma^2)$.

Z algebraického pohledu je jistá statistika $\mathcal{S}$ jen *funkcí dat* $x_1, \ldots, x_n$.

Řekněme dále, že pozorování $x_1, \ldots, x_n$ nejsou k dispozici; máme k dispozici jen intervaly

$$\mathbf{x}_i = [\underline{x}_i, \overline{x}_i], \qquad i = 1, \ldots, n$$

splňující $x_1 \in \mathbf{x}_1, \ldots, x_n \in \mathbf{x}_n$.

Zkoumejme hodnoty

$$\overline{\mathcal{S}} := \sup\{\mathcal{S}(x_1, \ldots, x_n) : x_1 \in \mathbf{x}_1, \ldots, x_n \in \mathbf{x}_n\},$$
$$\underline{\mathcal{S}} := \inf\{\mathcal{S}(x_1, \ldots, x_n) : x_1 \in \mathbf{x}_1, \ldots, x_n \in \mathbf{x}_n\}.$$

Obecně je výpočet $\overline{\mathcal{S}}$ a $\underline{\mathcal{S}}$ algoritmicky neřešitelný (nerekursivní). Ve speciálních případech je situace lepší.

Ještě několik příkladů — pokračování

„**P**“ značí „polynomiální čas“, „**NPH**“ značí „**NP**-těžký problém“.

Ještě několik příkladů — pokračování

„**P**“ značí „polynomiální čas“, „**NPH**“ značí „**NP**-těžký problém“.

- **Výběrový průměr** $\widehat{\mu}$
  - $\overline{\widehat{\mu}}$ . . . **P**
  - $\underline{\widehat{\mu}}$ . . . **P**

Ještě několik příkladů — pokračování

„**P**“ značí „polynomiální čas“, „**NPH**“ značí „**NP**-těžký problém“.

- **Výběrový průměr $\widehat{\mu}$**
  - $\overline{\widehat{\mu}}$ . . . **P**
  - $\underline{\widehat{\mu}}$ . . . **P**
- **Výběrový rozptyl $\widehat{\sigma^2}$**
  - je-li μ známé: $\overline{\widehat{\sigma^2}}$ . . . **P**
  - je-li μ známé: $\underline{\widehat{\sigma^2}}$ . . . **P**

Ještě několik příkladů — pokračování

„**P**“ značí „polynomiální čas“, „**NPH**“ značí „**NP**-těžký problém“.

- **Výběrový průměr $\widehat{\mu}$**
  - $\overline{\widehat{\mu}}$ . . . **P**
  - $\underline{\widehat{\mu}}$ . . . **P**
- **Výběrový rozptyl $\widehat{\sigma^2}$**
  - je-li μ známé: $\overline{\widehat{\sigma^2}}$ . . . **P**
  - je-li μ známé: $\underline{\widehat{\sigma^2}}$ . . . **P**
  - je-li μ neznámé: $\overline{\widehat{\sigma^2}}$ . . . **NPH**
  - je-li μ neznámé: $\underline{\widehat{\sigma^2}}$ . . . **P**

Ještě několik příkladů — pokračování

„**P**“ značí „polynomiální čas“, „**NPH**“ značí „**NP**-těžký problém“.

- **Výběrový průměr** $\widehat{\mu}$
  - $\overline{\widehat{\mu}}$ ... **P**
  - $\underline{\widehat{\mu}}$ ... **P**
- **Výběrový rozptyl** $\widehat{\sigma^2}$
  - je-li μ známé: $\overline{\widehat{\sigma^2}}$ ... **P**
  - je-li μ známé: $\underline{\widehat{\sigma^2}}$ ... **P**
  - je-li μ neznámé: $\overline{\widehat{\sigma^2}}$ ... **NPH**
  - je-li μ neznámé: $\underline{\widehat{\sigma^2}}$ ... **P**
- **t-poměr** t
  - $\overline{t}$ ... **P**
  - $\underline{t}$ ... **NPH**

Ještě několik příkladů — pokračování

„**P**“ značí „polynomiální čas“, „**NPH**“ značí „**NP**-těžký problém“.

- **Výběrový průměr** $\widehat{\mu}$
  - $\overline{\widehat{\mu}}$ . . . **P**
  - $\underline{\widehat{\mu}}$ . . . **P**
- **Výběrový rozptyl** $\widehat{\sigma^2}$
  - je-li μ známé: $\overline{\widehat{\sigma^2}}$ . . . **P**
  - je-li μ známé: $\underline{\widehat{\sigma^2}}$ . . . **P**
  - je-li μ neznámé: $\overline{\widehat{\sigma^2}}$ . . . **NPH**
  - je-li μ neznámé: $\underline{\widehat{\sigma^2}}$ . . . **P**
- **t-poměr** t
  - $\overline{t}$ . . . **P**
  - $\underline{t}$ . . . **NPH**
- **F-poměr** F
  - $\overline{F}$ . . . **NPH**
  - $\underline{F}$ . . . **NPH**

Residuální hodnoty modelu: úvod

Uvažme opět systém instancí regresního modelu

$$y = X\beta + \varepsilon,$$

kde $X \in \mathbf{X}$ a $y \in \mathbf{y}$, a reziduální hodnotu

$$\mathcal{R}_k = \min_{\beta \in \mathbb{R}^p} \|X\beta - y\|_k.$$

Residuální hodnoty modelu: úvod

Uvažme opět systém instancí regresního modelu

$$y = X\beta + \varepsilon,$$

kde $X \in \mathbf{X}$ a $y \in \mathbf{y}$, a reziduální hodnotu

$$\mathcal{R}_k = \min_{\beta \in \mathbb{R}^p} \|X\beta - y\|_k.$$

L_k-normy.

- $\|x\|_2 := \sqrt{x^{\mathrm{T}} x}$, metoda nejmenších čtverců,
- $\|x\|_1 := \sum_{i=1}^{n} |x_i|$, metoda LAD, Least Absolute Deviations
- $\|x\|_\infty := \max_{i=1}^{n} |x_i|$, čebyševovská aproximace.

Residuální hodnoty modelu: úvod

Uvažme opět systém instancí regresního modelu

$$y = X\beta + \varepsilon,$$

kde $X \in \mathbf{X}$ a $y \in \mathbf{y}$, a reziduální hodnotu

$$\mathcal{R}_k = \min_{\beta \in \mathbb{R}^p} \|X\beta - y\|_k.$$

L_k-normy.

- $\|x\|_2 := \sqrt{x^{\mathrm{T}} x}$, metoda nejmenších čtverců,
- $\|x\|_1 := \sum_{i=1}^n |x_i|$, metoda LAD, Least Absolute Deviations
- $\|x\|_\infty := \max_{i=1}^n |x_i|$, čebyševovská aproximace.

Problém. Zajímá nás rozsah možných hodnot $\mathcal{R}_k$, tj.

$$\underline{\mathcal{R}}_k = \min_{X \in \mathbf{X}, y \in \mathbf{y}} \min_{\beta \in \mathbb{R}^p} \|X\beta - y\|_k,$$

$$\overline{\mathcal{R}}_k = \max_{X \in \mathbf{X}, y \in \mathbf{y}} \min_{\beta \in \mathbb{R}^p} \|X\beta - y\|_k.$$

Residuální hodnoty modelu: úvod

Připomeňme:

- $\min_{\beta\in\mathbb{R}^p} \|X\beta - y\|_2$ se dá vyjádřit explicitně

$$\beta = (X^T X)^{-1} X^T y,$$

- $\min_{\beta\in\mathbb{R}^p} \|X\beta - y\|_\infty$ se dá zredukovat na lineární program

$$\min t \ \text{ subject to } \ X\beta - y \le te, \ -X\beta + y \le te, \ t \ge 0,$$

- $\min_{\beta\in\mathbb{R}^p} \|X\beta - y\|_1$ se dá zredukovat na lineární program

$$\min e^{\mathrm{T}} w \ \text{ subject to } \ X\beta - y \le w, \ -X\beta + y \le w, \ w \ge 0.$$

Residuální hodnoty modelu: přehled výsledků

Třídy problémů: polynomiální (P) vs. **NP**-těžké (NPH)

	I	II	III	IV	V
m	obecné	obecné	pevné	obecné	pevné
X	intervalové	intervalové	intervalové	reálné	reálné
y	intervalové	intervalové	intervalové	intervalové	intervalové
β	obecné	nezáporné	obecné	obecné	obecné
$\overline{\mathcal{R}}_1$	NPH	NPH	P	P	P
$\underline{\mathcal{R}}_1$	NPH	P	P	P	P
$\overline{\mathcal{R}}_2$	NPH	NPH	NPH	NPH	NPH
$\underline{\mathcal{R}}_2$	NPH	P	P	P	P
$\overline{\mathcal{R}}_\infty$	NPH	NPH	P	P	P
$\underline{\mathcal{R}}_\infty$	NPH	P	P	P	P

Residuální hodnoty modelu: výpočet $\underline{\mathcal{R}}$ pro $\beta \geq 0$

Věta

Buď β a priori nezáporné (= prostor parametrů je omezen na nezáporný ortant $\mathbb{R}^p$).

(a) **L_∞-norma:**

$$\underline{\mathcal{R}}_\infty = \min\{t : \underline{X}\beta - \overline{y} \leq te,\ -\overline{X}\beta + \underline{y} \leq te,\ t \geq 0\}.$$

(b) **L_1-norma:**

$$\underline{\mathcal{R}}_1 = \min\{e^{\mathrm{T}} w : \underline{X}\beta - \overline{y} \leq w,\ -\overline{X}\beta + \underline{y} \leq w,\ w \geq 0\}.$$

(b) **L_2-norma:**

$$\underline{\mathcal{R}}_2^2 = \min\{\|w\|_2^2 : \underline{X}\beta - \overline{y} \leq w \leq \overline{X}\beta - \underline{y},\ \beta \geq 0\}.$$

Residuální hodnoty modelu: výpočet $\underline{\mathcal{R}}$ pro p fixní

Idea. Rozděl prostor parametrů $\mathbb{R}^p \ni \beta$ na jednotlivé ortanty, na nich to umíme řešit efektivně.

Residuální hodnoty modelu: výpočet $\underline{\mathcal{R}}$ pro p fixní

Idea. Rozděl prostor parametrů $\mathbb{R}^p \ni \beta$ na jednotlivé ortanty, na nich to umíme řešit efektivně.

Věta

(a) L_∞**-norma:**

$$\underline{\mathcal{R}}_\infty = \min_{s \in \{\pm 1\}^p} R_s^\infty,$$

kde

$$\begin{aligned} R_s^\infty = \min\{t : \quad & (X^c - X^\Delta \operatorname{diag}(s))\beta - \overline{y} \leq te, \\ & (-X^c - X^\Delta \operatorname{diag}(s))\beta + \underline{y} \leq -te, \\ & \operatorname{diag}(s)\beta \geq 0, \ t \geq 0\}. \end{aligned}$$

(b) L_1**,** L_2**-norma:** *analogicky*

Residuální hodnoty modelu: výpočet $\underline{\mathcal{R}}$ pro p fixní

Idea. Rozděl prostor parametrů $\mathbb{R}^p \ni \beta$ na jednotlivé ortanty, na nich to umíme řešit efektivně.

Věta

(a) L_∞**-norma:**

$$\underline{\mathcal{R}}_\infty = \min_{s \in \{\pm 1\}^p} R_s^\infty,$$

kde

$$R_s^\infty = \min\{t : \begin{aligned}[t] &(X^c - X^\Delta \operatorname{diag}(s))\beta - \overline{y} \le te, \\ &(-X^c - X^\Delta \operatorname{diag}(s))\beta + \underline{y} \le -te, \\ &\operatorname{diag}(s)\beta \ge 0,\ t \ge 0\}. \end{aligned}$$

(b) L_1, L_2**-norma:** *analogicky*

Výpočetní složitost:

- $2^p \cdot$ (lineární program) pro L_∞ a L_1-normu
- $2^p \cdot$ (konvexní kvadratický program) pro L_2-normu

Residuální hodnoty modelu: výpočet $\overline{\mathcal{R}}$ pro p fixní

Věta

Pro p pevné, $\overline{\mathcal{R}}_\infty$ *and* $\overline{\mathcal{R}}_1$ *jsou spočítatelné v polynomiálním čase.*

Residuální hodnoty modelu: výpočet $\overline{\mathcal{R}}$ pro p fixní

Věta

Pro p pevné, $\overline{\mathcal{R}}_\infty$ and $\overline{\mathcal{R}}_1$ jsou spočítatelné v polynomiálním čase.

Důkaz.

1. Musíme projít všechny možné báze určitého lineárního programu (těch je maximálně $\binom{2n+1}{p+1}$).
2. Pro každou bázi řešíme intervalovou soustavu rovnic řádu $p+1$. $\square$

Residuální hodnoty modelu: výpočet $\overline{\mathcal{R}}$ pro p fixní

Věta

Pro p pevné, $\overline{\mathcal{R}}_\infty$ and $\overline{\mathcal{R}}_1$ jsou spočítatelné v polynomiálním čase.

Důkaz.

1. Musíme projít všechny možné báze určitého lineárního programu (těch je maximálně $\binom{2n+1}{p+1}$).
2. Pro každou bázi řešíme intervalovou soustavu rovnic řádu $p+1$. □

Důsledek.

Úloha je sice v P, ale časová složitost algoritmu je shora omezená

$$(4n)^{p+1} \cdot (\text{lineární program})$$

Residuální hodnoty modelu: výpočet $\overline{\mathcal{R}}$ pro p fixní

Věta

Pro p pevné, $\overline{\mathcal{R}}_\infty$ and $\overline{\mathcal{R}}_1$ jsou spočítatelné v polynomiálním čase.

Důkaz.

1. Musíme projít všechny možné báze určitého lineárního programu (těch je maximálně $\binom{2n+1}{p+1}$).
2. Pro každou bázi řešíme intervalovou soustavu rovnic řádu $p+1$. $\square$

Důsledek.
Úloha je sice v P, ale časová složitost algoritmu je shora omezená

$$(4n)^{p+1} \cdot (\text{lineární program})$$

Tedy: ne vždy v P znamená rychle!

Množina hodnot estimátoru

Problém. Zajímá nás množina hodnot estimátoru $\arg\min_{\beta'} \|X\beta' - y\|_k$, tj. množina

$$B_k := \left\{ \beta \in \mathbb{R}^p : \beta \in \arg\min_{\beta' \in \mathbb{R}^p} \|X\beta' - y\|_k \text{ pro nějaké } y \in \mathbf{y}, X \in \mathbf{X} \right\}.$$

Množina hodnot estimátoru

Problém. Zajímá nás množina hodnot estimátoru $\arg\min_{\beta'} \|X\beta' - y\|_k$, tj. množina

$$B_k := \left\{ \beta \in \mathbb{R}^p : \beta \in \arg\min_{\beta' \in \mathbb{R}^p} \|X\beta' - y\|_k \text{ pro nějaké } y \in \mathbf{y}, X \in \mathbf{X} \right\}.$$

Cheme najít co nejtěsnější intervalovou obálku $\mathbf{b} \in \mathbb{IR}^m$ tak, že $B_k \subseteq \mathbf{b}$.

Množina hodnot estimátoru

Problém. Zajímá nás množina hodnot estimátoru $\arg\min_{\beta'} \|X\beta' - y\|_k$, tj. množina

$$B_k := \left\{ \beta \in \mathbb{R}^p : \beta \in \arg\min_{\beta' \in \mathbb{R}^p} \|X\beta' - y\|_k \text{ pro nějaké } y \in \mathbf{y}, X \in \mathbf{X} \right\}.$$

Cheme najít co nejtěsnější intervalovou obálku $\mathbf{b} \in \mathbb{IR}^m$ tak, že $B_k \subseteq \mathbf{b}$.

Velmi těžký problém!

Množina hodnot estimátoru

Problém. Zajímá nás množina hodnot estimátoru $\arg\min_{\beta'} \|X\beta' - y\|_k$, tj. množina

$$B_k := \left\{ \beta \in \mathbb{R}^p : \beta \in \arg\min_{\beta' \in \mathbb{R}^p} \|X\beta' - y\|_k \text{ pro nějaké } y \in \mathbf{y}, X \in \mathbf{X} \right\}.$$

Cheme najít co nejtěsnější intervalovou obálku $\mathbf{b} \in \mathbb{IR}^m$ tak, že $B_k \subseteq \mathbf{b}$.

Velmi těžký problém!

Věta

1. *Pro jakoukoli normu, testovat neomezenost B_k je co-**NP**-těžké.*
2. *Pro jakoukoli normu, obalit B_k do intervalu je **NP**-těžké i s předepsanou chybou.*

Estimátor s L_2-normou

1. způsob: Řešíme intervalovou soustavu (řádu p)

$$\mathbf{X}^T\mathbf{X}\beta = \mathbf{X}^T\mathbf{y}.$$

Estimátor s L_2-normou

1. způsob: Řešíme intervalovou soustavu (řádu p)

$$\mathbf{X}^T\mathbf{X}\beta = \mathbf{X}^T\mathbf{y}.$$

Díky tzv. problému závislosti je výsledek velmi nadhodnocený.

Estimátor s L_2-normou

1. způsob: Řešíme intervalovou soustavu (řádu p)

$$\mathbf{X}^T\mathbf{X}\beta = \mathbf{X}^T\mathbf{y}.$$

Díky tzv. problému závislosti je výsledek velmi nadhodnocený.

2. způsob: Řešíme intervalovou soustavu (řádu $n+p$)

$$\begin{pmatrix} 0_p & \mathbf{X}^T \\ \mathbf{X} & I_n \end{pmatrix} \begin{pmatrix} \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{y} \end{pmatrix}.$$

Estimátor s L_2-normou

1. způsob: Řešíme intervalovou soustavu (řádu p)

$$\mathbf{X}^T\mathbf{X}\beta = \mathbf{X}^T\mathbf{y}.$$

Díky tzv. problému závislosti je výsledek velmi nadhodnocený.

2. způsob: Řešíme intervalovou soustavu (řádu $n+p$)

$$\begin{pmatrix} 0_p & \mathbf{X}^T \\ \mathbf{X} & I_n \end{pmatrix} \begin{pmatrix} \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{y} \end{pmatrix}.$$

- vždy lepší řešení
- ale soustava je větší

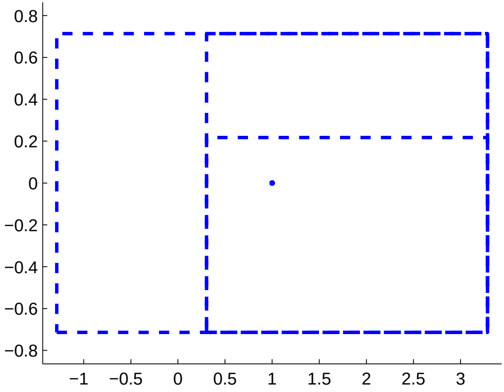
Příklad:

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ [0,2] & 2 & 3 & 4 & 5 \end{pmatrix}^{\mathrm{T}}, \quad \mathbf{y} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix}^{\mathrm{T}}.$$

Estimátor s L_2-normou

Příklad:

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ [0,2] & 2 & 3 & 4 & 5 \end{pmatrix}^{\mathrm{T}}, \quad \mathbf{y} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix}^{\mathrm{T}}.$$

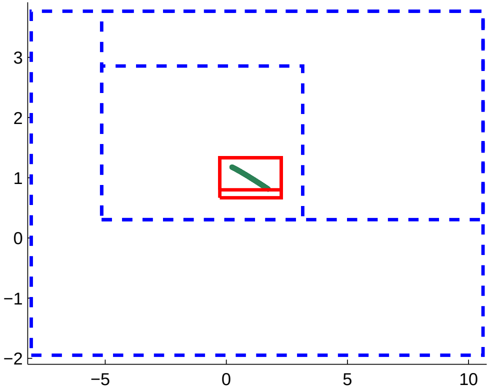


Příklad:

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ [0,2] & 2 & 3 & 4 & 5 \end{pmatrix}^{\mathrm{T}}, \quad \mathbf{y} = \begin{pmatrix} 2 & 3 & 4 & 5 & 6 \end{pmatrix}^{\mathrm{T}}.$$

Příklad:

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ [0,2] & 2 & 3 & 4 & 5 \end{pmatrix}^{\mathrm{T}}, \quad \mathbf{y} = \begin{pmatrix} 2 & 3 & 4 & 5 & 6 \end{pmatrix}^{\mathrm{T}}.$$



Estimátor s L_1 a L_∞-normou: bázická stabilita

Pro L_1 a L_∞-normu lze úlohu vyjádřit ve tvaru

$$\min c^{\mathrm{T}} u \quad \text{subject to} \quad Au \le b, \tag{$*$}$$

kde $A \in \mathbf{A}$ a $b \in \mathbf{b}$.

Estimátor s L_1 a L_∞-normou: bázická stabilita

Pro L_1 a L_∞-normu lze úlohu vyjádřit ve tvaru

$$\min c^{\mathrm{T}} u \quad \text{subject to} \quad Au \le b, \tag{$*$}$$

kde $A \in \mathbf{A}$ a $b \in \mathbf{b}$.

Pěkná podtřída úloh: **bázicky stabilní.**

Estimátor s L_1 a L_∞-normou: bázická stabilita

Pro L_1 a L_∞-normu lze úlohu vyjádřit ve tvaru

$$\min c^{\mathrm{T}} u \;\; \text{subject to} \;\; Au \leq b, \tag{$*$}$$

kde $A \in \mathbf{A}$ a $b \in \mathbf{b}$.

Pěkná podtřída úloh: **bázicky stabilní.**

Definice. Úloha $(*)$ je bázicky stabilní, pokud existuje báze jsoucí optimální pro všechna $A \in \mathbf{A}$ a $b \in \mathbf{b}$.

Estimátor s L_1 a L_∞-normou: bázická stabilita

Pro L_1 a L_∞-normu lze úlohu vyjádřit ve tvaru

$$\min c^{\mathrm{T}} u \quad \text{subject to} \quad Au \leq b, \tag{$*$}$$

kde $A \in \mathbf{A}$ a $b \in \mathbf{b}$.

Pěkná podtřída úloh: **bázicky stabilní.**

Definice. Úloha $(*)$ je bázicky stabilní, pokud existuje báze jsoucí optimální pro všechna $A \in \mathbf{A}$ a $b \in \mathbf{b}$.

Věta

*Ověřování bázické stability je co-**NP**-těžké.*

Proč je bázická stabilita užitečná?

Buď úloha bázicky stabilní s bází B. Pak:

- Množina optimálních řešení je rovna množině řešení intervalové soustavy $\mathbf{A}_B u = \mathbf{b}_B$.

Proč je bázická stabilita užitečná?

Buď úloha bázicky stabilní s bází B. Pak:

- Množina optimálních řešení je rovna množině řešení intervalové soustavy $\mathbf{A}_B u = \mathbf{b}_B$.

  To je známý problém, mnoho metod, lze přesně pro malé rozměry (obecně **NP**-těžké).

Proč je bázická stabilita užitečná?

Buď úloha bázicky stabilní s bází B. Pak:

- Množina optimálních řešení je rovna množině řešení intervalové soustavy $\mathbf{A}_B u = \mathbf{b}_B$.

  To je známý problém, mnoho metod, lze přesně pro malé rozměry (obecně **NP**-těžké).

- Víme-li, že řešení je a priori nezáporné, pak množina optimálních řešení je konvexní polyedr popsaný

$$\underline{A}_B u \leq \overline{b}_B, \ \overline{A}_B u \geq \underline{b}_B, \ u \geq 0.$$

Testování bázické stability

Báze B je optimální pro dané $A \in \mathbf{A}$ a $b \in \mathbf{b}$, pokud platí:

1. regularita: A_B je regulární;
2. přípustnost: lineární systém $A_B u = b_B$, $A_N u \leq b_N$ je přípustný;
3. optimalita: řešení soustavy $A_B^{\mathrm{T}} v = c$ je nezáporné.

Testování bázické stability

Báze B je optimální pro dané $A \in \mathbf{A}$ a $b \in \mathbf{b}$, pokud platí:

1. regularita: A_B je regulární;
2. přípustnost: lineární systém $A_B u = b_B$, $A_N u \leq b_N$ je přípustný;
3. optimalita: řešení soustavy $A_B^{\mathrm{T}} v = c$ je nezáporné.

Zobecnění na intervaly:

1. regularita: implicitně v následujícím;
2. najdi obálku $\mathbf{u}$ pro řešení $\mathbf{A}_B u = \mathbf{b}_B$ a otestuj $\mathbf{A}_N \mathbf{u} \leq \mathbf{b}_N$;
3. optimalita: řešení soustavy $A_B^{\mathrm{T}} v = c$ je nezáporné.

Testování bázické stability

Báze B je optimální pro dané $A \in \mathbf{A}$ a $b \in \mathbf{b}$, pokud platí:

1. regularita: A_B je regulární;
2. přípustnost: lineární systém $A_B u = b_B$, $A_N u \leq b_N$ je přípustný;
3. optimalita: řešení soustavy $A_B^T v = c$ je nezáporné.

Zobecnění na intervaly:

1. regularita: implicitně v následujicím;
2. najdi obálku $\mathbf{u}$ pro řešení $\mathbf{A}_B u = \mathbf{b}_B$ a otestuj $\mathbf{A}_N \mathbf{u} \leq \mathbf{b}_N$;
3. optimalita: řešení soustavy $A_B^T v = c$ je nezáporné.

Ilustrace bázické stability.

- L_1-regresní přímka jako klasifikator, který rozděluje daná data (x, y) na dvě třídy, nad přímkou a pod přímkou
- bázická stabilita = stejná klasifikace pro libovolný výběr (x, y) z $(\mathbf{x}, \mathbf{y})$

Závěr

Michale, napadá tě něco chytrého na závěr?

Závěr

Michale, napadá tě něco chytrého na závěr?
Michal: Ne. . .

Závěr

Michale, napadá tě něco chytrého na závěr?
Michal: Ne. . .

Co jsme chtěli sdělit:

- intervaly se běžně vyskytují a je třeba s nimi počítat

Závěr

Michale, napadá tě něco chytrého na závěr?
Michal: Ne...

Co jsme chtěli sdělit:

- intervaly se běžně vyskytují a je třeba s nimi počítat
- náš přístup: vyšetřit nejhorší a nejlepší případ

Závěr

Michale, napadá tě něco chytrého na závěr?
Michal: Ne...

Co jsme chtěli sdělit:

- intervaly se běžně vyskytují a je třeba s nimi počítat
- náš přístup: vyšetřit nejhorší a nejlepší případ
- obecně může těžké tyto případy spočítat

Závěr

Michale, napadá tě něco chytrého na závěr?
Michal: Ne...

Co jsme chtěli sdělit:

- intervaly se běžně vyskytují a je třeba s nimi počítat
- náš přístup: vyšetřit nejhorší a nejlepší případ
- obecně může těžké tyto případy spočítat
- ale za určitých předpokladů to může být snadné

Závěr

Michale, napadá tě něco chytrého na závěr?
Michal: Ne...

Co jsme chtěli sdělit:

- intervaly se běžně vyskytují a je třeba s nimi počítat
- náš přístup: vyšetřit nejhorší a nejlepší případ
- obecně může těžké tyto případy spočítat
- ale za určitých předpokladů to může být snadné

Je zde možné pracovat na řadě problémů: např.

- konstruovat dobré obálky pro B_k;
- zkoumat hodnoty výrazu $(X^{\mathrm{T}}\Omega^{-1}X)^{-1}X^{\mathrm{T}}\Omega^{-1}y$, kde matice Ω je intervalová apod.

Závěr

Michale, napadá tě něco chytrého na závěr?
Michal: Ne…

Co jsme chtěli sdělit:

- intervaly se běžně vyskytují a je třeba s nimi počítat
- náš přístup: vyšetřit nejhorší a nejlepší případ
- obecně může těžké tyto případy spočítat
- ale za určitých předpokladů to může být snadné

Je zde možné pracovat na řadě problémů: např.

- konstruovat dobré obálky pro B_k;
- zkoumat hodnoty výrazu $(X^{\mathrm{T}}\Omega^{-1}X)^{-1}X^{\mathrm{T}}\Omega^{-1}y$, kde matice Ω je intervalová apod.

Děkujeme za pozornost; budeme vděčni za dotazy a za „feedback“.