Numerikus algoritmusok

Hajnal Péter

2021. tavasz

Bináris optimum keresés egyváltozós függvény esetében

Bináris optimum keresés egyváltozós függvény esetében

A kiindulási probléma

Legyen f az $\mathbb{R}$ egy konvex részhalmazán értelmezett konvex függvény. Szeretnénk meghatározni a minimumát.

Bináris optimum keresés egyváltozós függvény esetében

A kiindulási probléma

Legyen f az $\mathbb{R}$ egy konvex részhalmazán értelmezett konvex függvény. Szeretnénk meghatározni a minimumát.

- Tegyük fel, hogy kiindulásként ismert egy E_0 zárt intervallum (kompakt konvex halmaz $\mathbb{R}$-ben), amely tartalmaz egy x^* optimális helyet.

Bináris optimum keresés egyváltozós függvény esetében

A kiindulási probléma

Legyen f az $\mathbb{R}$ egy konvex részhalmazán értelmezett konvex függvény. Szeretnénk meghatározni a minimumát.

- Tegyük fel, hogy kiindulásként ismert egy E_0 zárt intervallum (kompakt konvex halmaz $\mathbb{R}$-ben), amely tartalmaz egy x^* optimális helyet.

- Tegyük fel, ha ismerünk egy $x_0 \in \mathrm{dom}\ f$ helyet, akkor $\mathrm{dom}\ f \cap \{x \in R : x \leq x_0\}$ és $\mathrm{dom}\ f \cap \{x \in R : x \geq x_0\}$ halmazok egyikét ki tudjuk zárni x^* keresésénél.

Bináris optimum keresés egyváltozós függvény esetében

A kiindulási probléma

Legyen f az $\mathbb{R}$ egy konvex részhalmazán értelmezett konvex függvény. Szeretnénk meghatározni a minimumát.

- Tegyük fel, hogy kiindulásként ismert egy E_0 zárt intervallum (kompakt konvex halmaz $\mathbb{R}$-ben), amely tartalmaz egy x^* optimális helyet.
- Tegyük fel, ha ismerünk egy $x_0 \in \operatorname{dom} f$ helyet, akkor $\operatorname{dom} f \cap \{x \in R : x \leq x_0\}$ és $\operatorname{dom} f \cap \{x \in R : x \geq x_0\}$ halmazok egyikét ki tudjuk zárni x^* keresésénél.
- A legegyszerűbb módszer egy optimum hely meghatározására, illetve kellő közelítésére a bináris keresés.

Bináris keresés

Bináris keresés

Bináris keresés

Bináris keresés

Bináris keresés

// E_0 egy közel optimumhelyét keressük

Bináris keresés

Bináris keresés

// E_0 egy közel optimumhelyét keressük

(F) // Felezés

Bináris keresés

Bináris keresés

// E_0 egy közel optimumhelyét keressük

(F) // Felezés

Megvizsgáljuk az E_0 intervallum o_0 középpontját.

Bináris keresés

Bináris keresés

// E_0 egy közel optimumhelyét keressük

(F) // Felezés

Megvizsgáljuk az E_0 intervallum o_0 középpontját.

- Ha ez a középpont nem esik dom f-ba, akkor áttérünk arra a félintervallumra, amelyik tartalmazza dom f-et.

Bináris keresés

Bináris keresés

// E_0 egy közel optimumhelyét keressük

(F) // Felezés

Megvizsgáljuk az E_0 intervallum o_0 középpontját.

- Ha ez a középpont nem esik dom f-ba, akkor áttérünk arra a félintervallumra, amelyik tartalmazza dom f-et. Ez lesz E'.

Bináris keresés

Bináris keresés

// E_0 egy közel optimumhelyét keressük

(F) // Felezés

Megvizsgáljuk az E_0 intervallum o_0 középpontját.

- Ha ez a középpont nem esik dom f-ba, akkor áttérünk arra a félintervallumra, amelyik tartalmazza dom f-et. Ez lesz E'.
- Ha ez a középpont dom f-be esik, akkor megvizsgáljuk f viselkedését a középpontban: melyik félintervallumban kell tovább keresni.

Bináris keresés

Bináris keresés

// E_0 egy közel optimumhelyét keressük

(F) // Felezés

Megvizsgáljuk az E_0 intervallum o_0 középpontját.

- Ha ez a középpont nem esik dom f-ba, akkor áttérünk arra a félintervallumra, amelyik tartalmazza dom f-et. Ez lesz E'.
- Ha ez a középpont dom f-be esik, akkor megvizsgáljuk f viselkedését a középpontban: melyik félintervallumban kell tovább keresni. Ez lesz E'.

Bináris keresés

Bináris keresés

// E_0 egy közel optimumhelyét keressük

(F) // Felezés
Megvizsgáljuk az E_0 intervallum o_0 középpontját.

- Ha ez a középpont nem esik dom f-ba, akkor áttérünk arra a félintervallumra, amelyik tartalmazza dom f-et. Ez lesz E'.
- Ha ez a középpont dom f-be esik, akkor megvizsgáljuk f viselkedését a középpontban: melyik félintervallumban kell tovább keresni. Ez lesz E'.

(L) // Leállási feltétel

Bináris keresés

Bináris keresés

// E_0 egy közel optimumhelyét keressük

(F) // Felezés

Megvizsgáljuk az E_0 intervallum o_0 középpontját.

- Ha ez a középpont nem esik dom f-ba, akkor áttérünk arra a félintervallumra, amelyik tartalmazza dom f-et. Ez lesz E'.
- Ha ez a középpont dom f-be esik, akkor megvizsgáljuk f viselkedését a középpontban: melyik félintervallumban kell tovább keresni. Ez lesz E'.

(L) // Leállási feltétel

Mindkét esetben biztosak lehetünk, hogy E' tartalmazza keresett x^*-t.

Bináris keresés

Bináris keresés

// E_0 egy közel optimumhelyét keressük

(F) // Felezés
Megvizsgáljuk az E_0 intervallum o_0 középpontját.

- Ha ez a középpont nem esik dom f-ba, akkor áttérünk arra a félintervallumra, amelyik tartalmazza dom f-et. Ez lesz E'.
- Ha ez a középpont dom f-be esik, akkor megvizsgáljuk f viselkedését a középpontban: melyik félintervallumban kell tovább keresni. Ez lesz E'.

(L) // Leállási feltétel
Mindkét esetben biztosak lehetünk, hogy E' tartalmazza keresett x^*-t.
- Ha E' rövid, akkor középpontja kellő pontosságú közelítése x^*-nak

Bináris keresés

Bináris keresés

// E_0 egy közel optimumhelyét keressük

(F) // Felezés
Megvizsgáljuk az E_0 intervallum o_0 középpontját.

- Ha ez a középpont nem esik dom f-ba, akkor áttérünk arra a félintervallumra, amelyik tartalmazza dom f-et. Ez lesz E'.
- Ha ez a középpont dom f-be esik, akkor megvizsgáljuk f viselkedését a középpontban: melyik félintervallumban kell tovább keresni. Ez lesz E'.

(L) // Leállási feltétel
Mindkét esetben biztosak lehetünk, hogy E' tartalmazza keresett x^*-t.

- Ha E' rövid, akkor középpontja kellő pontosságú közelítése x^*-nak
- Ha E' hosszú, akkor az új intervallummal ($E_0 \leftarrow E$) vissza (F)-hez.

Bináris keresés magasabb dimenzióban?

Bináris keresés magasabb dimenzióban?

- Mik töltsék be az intervallumok szerepét?

Bináris keresés magasabb dimenzióban?

- Mik töltsék be az intervallumok szerepét?
- Téglák, szabályos téglák?

Bináris keresés magasabb dimenzióban?

- Mik töltsék be az intervallumok szerepét?
- Téglák, szabályos téglák?
- Gömbök?

Bináris keresés magasabb dimenzióban?

- Mik töltsék be az intervallumok szerepét?
- Téglák, szabályos téglák?
- Gömbök?
- Felezni kell!

Bináris keresés magasabb dimenzióban?

- Mik töltsék be az intervallumok szerepét?
- Téglák, szabályos téglák?
- Gömbök?
- Felezni kell! Felezés után

Bináris keresés magasabb dimenzióban?

- Mik töltsék be az intervallumok szerepét?
- Téglák, szabályos téglák?
- Gömbök?
- Felezni kell! Felezés után
  - elcsúnyulás?

Bináris keresés magasabb dimenzióban?

- Mik töltsék be az intervallumok szerepét?
- Téglák, szabályos téglák?
- Gömbök?
- Felezni kell! Felezés után
  - elcsúnyulás?
  - szépítés, aminek ára nagyobbodás?

Bináris keresés magasabb dimenzióban?

- Mik töltsék be az intervallumok szerepét?
- Téglák, szabályos téglák?
- Gömbök?
- Felezni kell! Felezés után
  - elcsúnyulás?
  - szépítés, aminek ára nagyobbodás?
- Az intervallumok szerepét az ellipszoidok töltik be.

Bináris keresés magasabb dimenzióban?

- Mik töltsék be az intervallumok szerepét?
- Téglák, szabályos téglák?
- Gömbök?
- Felezni kell! Felezés után
  - elcsúnyulás?
  - szépítés, aminek ára nagyobbodás?
- Az intervallumok szerepét az ellipszoidok töltik be. Jelentős ötlet, áttörés.

Bináris keresés magasabb dimenzióban?

- Mik töltsék be az intervallumok szerepét?
- Téglák, szabályos téglák?
- Gömbök?
- Felezni kell! Felezés után
  - elcsúnyulás?
  - szépítés, aminek ára nagyobbodás?

• Az intervallumok szerepét az ellipszoidok töltik be. Jelentős ötlet, áttörés. $\rightarrow$ Ellipszoid módszer.

Ellipszoid módszer története

Ellipszoid módszer története

- Az ellipszoid módszert konvex nem differenciálható függvények optimalizálására Shor, Yudin, Nemirovskiy fejlesztette ki az 1970-es években.

Ellipszoid módszer története

- Az ellipszoid módszert konvex nem differenciálható függvények optimalizálására Shor, Yudin, Nemirovskiy fejlesztette ki az 1970-es években.
- Kahchian 1979-ben alkalmazta az LP feladatra.

Ellipszoid módszer története

- Az ellipszoid módszert konvex nem differenciálható függvények optimalizálására Shor, Yudin, Nemirovskiy fejlesztette ki az 1970-es években.
- Kahchian 1979-ben alkalmazta az LP feladatra.
- A gyakorlati teljesítménye rosszabb volt a szimplex módszerénél.

Ellipszoid módszer története

- Az ellipszoid módszert konvex nem differenciálható függvények optimalizálására Shor, Yudin, Nemirovskiy fejlesztette ki az 1970-es években.
- Kahchian 1979-ben alkalmazta az LP feladatra.
- A gyakorlati teljesítménye rosszabb volt a szimplex módszerénél.
- Azonban elméletileg rendkívül jelentős volt.

Ellipszoid módszer története

- Az ellipszoid módszert konvex nem differenciálható függvények optimalizálására Shor, Yudin, Nemirovskiy fejlesztette ki az 1970-es években.
- Kahchian 1979-ben alkalmazta az LP feladatra.
- A gyakorlati teljesítménye rosszabb volt a szimplex módszerénél.
- Azonban elméletileg rendkívül jelentős volt.

Ellipszoid módszer története

- Az ellipszoid módszert konvex nem differenciálható függvények optimalizálására Shor, Yudin, Nemirovskiy fejlesztette ki az 1970-es években.
- Kahchian 1979-ben alkalmazta az LP feladatra.
- A gyakorlati teljesítménye rosszabb volt a szimplex módszerénél.
- Azonban elméletileg rendkívül jelentős volt.

(i) Egyrészt az LP feladat problémáját $\mathcal{P}$-be helyezte.

Ellipszoid módszer története

- Az ellipszoid módszert konvex nem differenciálható függvények optimalizálására Shor, Yudin, Nemirovskiy fejlesztette ki az 1970-es években.
- Kahchian 1979-ben alkalmazta az LP feladatra.
- A gyakorlati teljesítménye rosszabb volt a szimplex módszerénél.
- Azonban elméletileg rendkívül jelentős volt.
  (i) Egyrészt az LP feladat problémáját $\mathcal{P}$-be helyezte.
  (ii) Másrészt az algoritmus alkalmazhatósága jóval szélesebbé vált: nem tette fel a lineáris egyenletrendszer explicit felírását.

Az ellipszoid módszer története II

Science Watch

The Flu Virus Presents New Enigmas

Mathematician Is Obscure No More

An Approach to Difficult Problems

Ellipszis a síkon, kör/gömb n-dimenzióban

Ellipszis a síkon, kör/gömb n-dimenzióban

Definíció (középiskola)

Origó középpontú, kétdimenziós ellipszoid, szabályos állású ellipszis az, melynek egyenlete $\frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1$.

Ellipszis a síkon, kör/gömb n-dimenzióban

Definíció (középiskola)

Origó középpontú, kétdimenziós ellipszoid, szabályos állású ellipszis az, melynek egyenlete $\frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1$.

Tétel (középiskola/egyetem?)

Az ellipszis a kör affin képe.

Ellipszis a síkon, kör/gömb n-dimenzióban

Definíció (középiskola)

Origó középpontú, kétdimenziós ellipszoid, szabályos állású ellipszis az, melynek egyenlete $\frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1$.

Tétel (középiskola/egyetem?)

Az ellipszis a kör affin képe.

Definíció (egyetem)

Egység sugarú, origó középpontú, n-dimenziós gömb:

$$B(1, \underline{0}) = \{x \in \mathbb{R}^n : x^T \cdot x \leq 1\}.$$

Ellipszis a síkon, kör/gömb n-dimenzióban

Definíció (középiskola)

Origó középpontú, kétdimenziós ellipszoid, szabályos állású ellipszis az, melynek egyenlete $\frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1$.

Tétel (középiskola/egyetem?)

Az ellipszis a kör affin képe.

Definíció (egyetem)

Egység sugarú, origó középpontú, n-dimenziós gömb:

$$B(1, \underline{0}) = \{x \in \mathbb{R}^n : x^T \cdot x \leq 1\}.$$

Egység sugarú, o középpontú, n-dimenziós gömb:

$$B(1, o) = \{x \in \mathbb{R}^n : (x - o)^T \cdot (x - o) \leq 1\}.$$

Ellipszoid: az ellipszis n-dimenzióban

Ellipszoid: az ellipszis n-dimenzióban

Definíció (egyetem)

$\varphi : \mathbb{R}^n \to \mathbb{R}^n$ transzformáció affin transzformáció, ha $\varphi : x \mapsto A \cdot x + b$ képlettel írható le, ahol $A \in \mathbb{R}^{n \times n}$ nem elfajuló mátrix.

Ellipszoid: az ellipszis n-dimenzióban

Definíció (egyetem)

$\varphi : \mathbb{R}^n \to \mathbb{R}^n$ transzformáció affin transzformáció, ha $\varphi : x \mapsto A \cdot x + b$ képlettel írható le, ahol $A \in \mathbb{R}^{n \times n}$ nem elfajuló mátrix.

- Az origót fixen hagyó affin transzformáció $x \mapsto A \cdot x$ képlettel írható le.

Ellipszoid: az ellipszis n-dimenzióban

Definíció (egyetem)

$\varphi : \mathbb{R}^n \to \mathbb{R}^n$ transzformáció affin transzformáció, ha $\varphi : x \mapsto A \cdot x + b$ képlettel írható le, ahol $A \in \mathbb{R}^{n \times n}$ nem elfajuló mátrix.

- Az origót fixen hagyó affin transzformáció $x \mapsto A \cdot x$ képlettel írható le.
- Az egységgömb affin képe:

$$\begin{aligned} A \cdot B(1, \underline{0}) &= \{A \cdot x : \quad x^T \cdot x \leq 1\} \\ &= \{A \cdot x : \quad (Ax)^T \cdot (A^T)^{-1} \cdot A^{-1} \cdot Ax \leq 1\} \\ &= \{y : \quad y^T \cdot (A \cdot A^T)^{-1} \cdot y \leq 1\} \end{aligned}$$

Ellipszoid: az ellipszis n-dimenzióban

Definíció (egyetem)

$\varphi : \mathbb{R}^n \to \mathbb{R}^n$ transzformáció affin transzformáció, ha $\varphi : x \mapsto A \cdot x + b$ képlettel írható le, ahol $A \in \mathbb{R}^{n\times n}$ nem elfajuló mátrix.

- Az origót fixen hagyó affin transzformáció $x \mapsto A \cdot x$ képlettel írható le.
- Az egységgömb affin képe:

$$\begin{aligned} A \cdot B(1, \underline{0}) &= \{A \cdot x : \quad x^T \cdot x \le 1\} \\ &= \{A \cdot x : \quad (Ax)^T \cdot (A^T)^{-1} \cdot A^{-1} \cdot Ax \le 1\} \\ &= \{y : \quad y^T \cdot (A \cdot A^T)^{-1} \cdot y \le 1\} \end{aligned}$$

- Legyen $P := A \cdot A^T$. Ekkor P szimmetrikus és pozitív definit.

Ellipszoid: az ellipszis n-dimenzióban

Definíció (egyetem)

$\varphi : \mathbb{R}^n \to \mathbb{R}^n$ transzformáció affin transzformáció, ha $\varphi : x \mapsto A \cdot x + b$ képlettel írható le, ahol $A \in \mathbb{R}^{n \times n}$ nem elfajuló mátrix.

- Az origót fixen hagyó affin transzformáció $x \mapsto A \cdot x$ képlettel írható le.
- Az egységgömb affin képe:

$$\begin{aligned} A \cdot B(1, \underline{0}) &= \{A \cdot x : \quad x^T \cdot x \le 1\} \\ &= \{A \cdot x : \quad (Ax)^T \cdot (A^T)^{-1} \cdot A^{-1} \cdot Ax \le 1\} \\ &= \{y : \quad y^T \cdot (A \cdot A^T)^{-1} \cdot y \le 1\} \end{aligned}$$

- Legyen $P := A \cdot A^T$. Ekkor P szimmetrikus és pozitív definit.

Ellipszoid (n-dimenzióban, teljes dimenziós)

$\mathcal{E}(P, o) = \{x : \ (x - o)^T \cdot P^{-1} \cdot (x - o) \le 1\}$, ahol $P \in \mathcal{S}^n_{++}$, $o \in \mathbb{R}^n$.

Egy kis geometria

Egy kis geometria

Mivel P szimmetrikus, így felírható a következő alakban: $P = Q^T \cdot \Lambda \cdot Q$, ahol $\Lambda = diag(\lambda_1, \ldots \lambda_n)$. $\lambda_i > 0$, mivel P pozitív definit. Ezen λ_i-k a féltengelyek hosszainak négyzetei ($\lambda_1 = a^2, \ldots, a_n^2$).

Egy kis geometria

Mivel P szimmetrikus, így felírható a következő alakban: $P = Q^T \cdot \Lambda \cdot Q$, ahol $\Lambda = diag(\lambda_1, \ldots \lambda_n)$. $\lambda_i > 0$, mivel P pozitív definit. Ezen λ_i-k a féltengelyek hosszainak négyzetei ($\lambda_1 = a^2, \ldots, a_n^2$).

Az egységsugarú, n-dimenziós gömb térfogata

$Vol(B(1,\underline{0})) = Vol(B(1,o)) = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)} = \beta_n$, egy n-től függő konstans, ahol $\Gamma(x) = \int_0^\infty e^{-t} \cdot t^{x-1}\, dt$, ha $x > 0$. Speciálisan $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, $\Gamma(1) = 1$, $\Gamma(k+1) = k\Gamma(k)$, így $k \in \mathbb{N}$ esetén $\Gamma(k+1) = k!$.

Egy kis geometria

Mivel P szimmetrikus, így felírható a következő alakban: $P = Q^T \cdot \Lambda \cdot Q$, ahol $\Lambda = diag(\lambda_1, \dots \lambda_n)$. $\lambda_i > 0$, mivel P pozitív definit. Ezen λ_i-k a féltengelyek hosszainak négyzetei ($\lambda_1 = a^2, \dots, a_n^2$).

Az egységsugarú, n-dimenziós gömb térfogata

$Vol(B(1, \underline{0})) = Vol(B(1, o)) = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)} = \beta_n$, egy n-től függő konstans, ahol $\Gamma(x) = \int_0^\infty e^{-t} \cdot t^{x-1}\, dt$, ha $x > 0$. Speciálisan $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, $\Gamma(1) = 1$, $\Gamma(k+1) = k\Gamma(k)$, így $k \in \mathbb{N}$ esetén $\Gamma(k+1) = k!$.

Tétel

Ha T mérhető test, $A \in \mathbb{R}^{n \times n}$, akkor $Vol(A \cdot T) = |\det A| \cdot Vol(T)$.

Egy kis geometria

Mivel P szimmetrikus, így felírható a következő alakban: $P = Q^T \cdot \Lambda \cdot Q$, ahol $\Lambda = diag(\lambda_1, \dots \lambda_n)$. $\lambda_i > 0$, mivel P pozitív definit. Ezen λ_i-k a féltengelyek hosszainak négyzetei ($\lambda_1 = a^2, \dots, a_n^2$).

Az egységsugarú, n-dimenziós gömb térfogata

$Vol(B(1, \underline{0})) = Vol(B(1, o)) = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)} = \beta_n$, egy n-től függő konstans, ahol $\Gamma(x) = \int_0^\infty e^{-t} \cdot t^{x-1}\, dt$, ha $x > 0$. Speciálisan $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, $\Gamma(1) = 1$, $\Gamma(k+1) = k\Gamma(k)$, így $k \in \mathbb{N}$ esetén $\Gamma(k+1) = k!$.

Tétel

Ha T mérhető test, $A \in \mathbb{R}^{n \times n}$, akkor $Vol(A \cdot T) = |\det A| \cdot Vol(T)$.

Következmény

$Vol(\mathcal{E}(P, o)) = \sqrt{\det P} \cdot \beta_n = \sqrt{\det(A \cdot A^\mathsf{T})} \cdot \beta_n = \det(A) \cdot \beta_n$.

Ellipszoidok és optimalizálás I

Ellipszoidok és optimalizálás I

Minimalizáljuk	$c^T \cdot x$-t
Feltéve, hogy	$x \in \mathcal{E}(P, o)$,

ahol $c \neq \underline{0}$.

Ellipszoidok és optimalizálás I

Minimalizáljuk	$c^T \cdot x$-t
Feltéve, hogy	$x \in \mathcal{E}(P, o)$,

ahol $c \neq \underline{0}$.

$$\begin{aligned}
&\min\{c^T \cdot x : (x-o)^T \cdot P^{-1} \cdot (x-o) \leq 1\} \\
&\quad = min\{c^T \cdot P^{\frac{1}{2}} \cdot y : (P^{\frac{1}{2}} \cdot y - o)^T \cdot P^{-1} \cdot (P^{\frac{1}{2}} \cdot y - o) \leq 1\} \\
&\quad = min\{c^T \cdot P^{\frac{1}{2}} \cdot y : \left(P^{\frac{1}{2}} \cdot (y - P^{-\frac{1}{2}} \cdot o)\right)^T \cdot P^{-1} \cdot P^{\frac{1}{2}} \cdot (y - P^{-\frac{1}{2}} \cdot o) \leq 1\} \\
&\quad = min\{c^T \cdot P^{\frac{1}{2}} \cdot y : (y - P^{-\frac{1}{2}} \cdot o)^T \cdot P^{\frac{1}{2}} \cdot P^{-1} \cdot P^{\frac{1}{2}} \cdot (y - P^{-\frac{1}{2}} \cdot o) \leq 1\} \\
&\quad = min\{c^T \cdot P^{\frac{1}{2}} \cdot y : (y - P^{-\frac{1}{2}} \cdot o)^T \cdot (y - P^{-\frac{1}{2}} \cdot o) \leq 1\},
\end{aligned}$$

Ellipszoidok és optimalizálás I

Minimalizáljuk	$c^T \cdot x$-t
Feltéve, hogy	$x \in \mathcal{E}(P, o)$,

ahol $c \neq \underline{0}$.

$$\begin{aligned}
&\min\{c^T \cdot x : (x-o)^T \cdot P^{-1} \cdot (x-o) \leq 1\} \\
&\quad = min\{c^T \cdot P^{\frac{1}{2}} \cdot y : (P^{\frac{1}{2}} \cdot y - o)^T \cdot P^{-1} \cdot (P^{\frac{1}{2}} \cdot y - o) \leq 1\} \\
&\quad = min\{c^T \cdot P^{\frac{1}{2}} \cdot y : \left(P^{\frac{1}{2}} \cdot (y - P^{-\frac{1}{2}} \cdot o)\right)^T \cdot P^{-1} \cdot P^{\frac{1}{2}} \cdot (y - P^{-\frac{1}{2}} \cdot o) \leq 1\} \\
&\quad = min\{c^T \cdot P^{\frac{1}{2}} \cdot y : (y - P^{-\frac{1}{2}} \cdot o)^T \cdot P^{\frac{1}{2}} \cdot P^{-1} \cdot P^{\frac{1}{2}} \cdot (y - P^{-\frac{1}{2}} \cdot o) \leq 1\} \\
&\quad = min\{c^T \cdot P^{\frac{1}{2}} \cdot y : (y - P^{-\frac{1}{2}} \cdot o)^T \cdot (y - P^{-\frac{1}{2}} \cdot o) \leq 1\},
\end{aligned}$$

ahol az első egyenlőségnél az $x = P^{\frac{1}{2}} \cdot y$ helyettesítést alkalmaztuk,

Ellipszoidok és optimalizálás I

Minimalizáljuk	$c^T \cdot x$-t
Feltéve, hogy	$x \in \mathcal{E}(P, o)$,

ahol $c \neq \underline{0}$.

$$
\begin{aligned}
&\min\{c^T \cdot x : (x-o)^T \cdot P^{-1} \cdot (x-o) \leq 1\} \\
&\quad = min\{c^T \cdot P^{\frac{1}{2}} \cdot y : (P^{\frac{1}{2}} \cdot y - o)^T \cdot P^{-1} \cdot (P^{\frac{1}{2}} \cdot y - o) \leq 1\} \\
&\quad = min\{c^T \cdot P^{\frac{1}{2}} \cdot y : \left(P^{\frac{1}{2}} \cdot (y - P^{-\frac{1}{2}} \cdot o)\right)^T \cdot P^{-1} \cdot P^{\frac{1}{2}} \cdot (y - P^{-\frac{1}{2}} \cdot o) \leq 1\} \\
&\quad = min\{c^T \cdot P^{\frac{1}{2}} \cdot y : (y - P^{-\frac{1}{2}} \cdot o)^T \cdot P^{\frac{1}{2}} \cdot P^{-1} \cdot P^{\frac{1}{2}} \cdot (y - P^{-\frac{1}{2}} \cdot o) \leq 1\} \\
&\quad = min\{c^T \cdot P^{\frac{1}{2}} \cdot y : (y - P^{-\frac{1}{2}} \cdot o)^T \cdot (y - P^{-\frac{1}{2}} \cdot o) \leq 1\},
\end{aligned}
$$

ahol az első egyenlőségnél az $x = P^{\frac{1}{2}} \cdot y$ helyettesítést alkalmaztuk, $P^{\frac{1}{2}} = Q^T \cdot diag(\sqrt{\lambda_1}, \ldots \sqrt{\lambda_n}) \cdot Q$, és $\left(P^{\frac{1}{2}}\right)^T = P^{\frac{1}{2}}$, melyet a harmadik egyenlőségnél használtunk ki.

Ellipszoidok és optimalizálás I (folytatás)

Ellipszoidok és optimalizálás I (folytatás)

A helyettesítéssel egy egységsugarú gömbre vonatkozó minimumfeladatot kaptunk, melynek megoldását könnyen meghatározhatjuk:

Ellipszoidok és optimalizálás I (folytatás)

A helyettesítéssel egy egységsugarú gömbre vonatkozó minimumfeladatot kaptunk, melynek megoldását könnyen meghatározhatjuk: a gömb középpontjától a célfüggvény vektorával ellentétes irányba mozduljunk el egy egységnyivel, vagyis

Ellipszoidok és optimalizálás I (folytatás)

A helyettesítéssel egy egységsugarú gömbre vonatkozó minimumfeladatot kaptunk, melynek megoldását könnyen meghatározhatjuk: a gömb középpontjától a célfüggvény vektorával ellentétes irányba mozduljunk el egy egységnyivel, vagyis

$$y^* = P^{-\frac{1}{2}} \cdot o - \frac{P^{\frac{1}{2}} \cdot c}{|P^{\frac{1}{2}} \cdot c|}.$$

Ellipszoidok és optimalizálás I (folytatás)

A helyettesítéssel egy egységsugarú gömbre vonatkozó minimumfeladatot kaptunk, melynek megoldását könnyen meghatározhatjuk: a gömb középpontjától a célfüggvény vektorával ellentétes irányba mozduljunk el egy egységnyivel, vagyis

$$y^* = P^{-\frac{1}{2}} \cdot o - \frac{P^{\frac{1}{2}} \cdot c}{|P^{\frac{1}{2}} \cdot c|}.$$

Visszatranszformálva ezt az ellipszoidra, x-re kapjuk, hogy

$$x^* = o - \frac{P \cdot c}{|P^{\frac{1}{2}} \cdot c|}.$$

Ellipszoidok és optimalizálás II

Ellipszoidok és optimalizálás II

Optimalizálási alapkérdés

Adott egy $\mathcal{E}(P, o)$ ellipszoid és az ellipszoid o középpontját a határán tartalmazó $\mathcal{F} = \{x : \quad g^T \cdot x \leq o^T \cdot x\}$ zárt féltér. Foglaljuk az $\mathcal{E} \cap \mathcal{F}$ félellipszoidot minimális térfogatú ellipszoidba!

Ellipszoidok és optimalizálás II

Optimalizálási alapkérdés

Adott egy $\mathcal{E}(P, o)$ ellipszoid és az ellipszoid o középpontját a határán tartalmazó $\mathcal{F} = \{x : \quad g^T \cdot x \leq o^T \cdot x\}$ zárt féltér. Foglaljuk az $\mathcal{E} \cap \mathcal{F}$ félellipszoidot minimális térfogatú ellipszoidba!

Tétel (Löwner—John)

Ha K kompakt, konvex, akkor létezik olyan minimális térfogatú $\mathcal{E}$ ellipszoid, hogy $K \subseteq \mathcal{E}$.

Ellipszoidok és optimalizálás II

Optimalizálási alapkérdés

Adott egy $\mathcal{E}(P, o)$ ellipszoid és az ellipszoid o középpontját a határán tartalmazó $\mathcal{F} = \{x : \quad g^T \cdot x \leq o^T \cdot x\}$ zárt féltér. Foglaljuk az $\mathcal{E} \cap \mathcal{F}$ félellipszoidot minimális térfogatú ellipszoidba!

Tétel (Löwner—John)

Ha K kompakt, konvex, akkor létezik olyan minimális térfogatú $\mathcal{E}$ ellipszoid, hogy $K \subseteq \mathcal{E}$.

Azaz a fenti feladat a félellipszoid Löwner—John-ellipszoidjának meghatározását kéri.

Ellipszoidok és optimalizálás II

Optimalizálási alapkérdés

Adott egy $\mathcal{E}(P, o)$ ellipszoid és az ellipszoid o középpontját a határán tartalmazó $\mathcal{F} = \{x : \quad g^T \cdot x \leq o^T \cdot x\}$ zárt féltér. Foglaljuk az $\mathcal{E} \cap \mathcal{F}$ félellipszoidot minimális térfogatú ellipszoidba!

Tétel (Löwner—John)

Ha K kompakt, konvex, akkor létezik olyan minimális térfogatú $\mathcal{E}$ ellipszoid, hogy $K \subseteq \mathcal{E}$.

Azaz a fenti feladat a félellipszoid Löwner—John-ellipszoidjának meghatározását kéri.

Általános alakzatra ez nehéz, néha kezelhetetlen feladat.

Ellipszoidok és optimalizálás II

Optimalizálási alapkérdés

Adott egy $\mathcal{E}(P, o)$ ellipszoid és az ellipszoid o középpontját a határán tartalmazó $\mathcal{F} = \{x : \quad g^T \cdot x \leq o^T \cdot x\}$ zárt féltér. Foglaljuk az $\mathcal{E} \cap \mathcal{F}$ félellipszoidot minimális térfogatú ellipszoidba!

Tétel (Löwner—John)

Ha K kompakt, konvex, akkor létezik olyan minimális térfogatú $\mathcal{E}$ ellipszoid, hogy $K \subseteq \mathcal{E}$.

Azaz a fenti feladat a félellipszoid Löwner—John-ellipszoidjának meghatározását kéri.

Általános alakzatra ez nehéz, néha kezelhetetlen feladat.

Félellipszoid esetén, azonban szép analitikus formula adható a megoldásra.

Egy egyszerű eset

Egy egyszerű eset

- Legyen $\mathcal{E} = B(1, 0)$ egységgömb, $\mathcal{F} = \{x_1 \geq 0\}$.

Egy egyszerű eset

- Legyen $\mathcal{E} = B(1, 0)$ egységgömb, $\mathcal{F} = \{x_1 \geq 0\}$. Először ezt az esetet vizsgáljuk.

Egy egyszerű eset

- Legyen $\mathcal{E} = B(1, 0)$ egységgömb, $\mathcal{F} = \{x_1 \geq 0\}$. Először ezt az esetet vizsgáljuk.
- Metszetükre mint felső félgömbre hivatkozunk.

Egy egyszerű eset

- Legyen $\mathcal{E} = B(1,0)$ egységgömb, $\mathcal{F} = \{x_1 \geq 0\}$. Először ezt az esetet vizsgáljuk.
- Metszetükre mint felső félgömbre hivatkozunk. Ide esik az $e_1 = (1,0,0,\dots)$ pont, amire mint a gömb „északi pólusa" hivatkozunk.

Egy egyszerű eset

- Legyen $\mathcal{E} = B(1,0)$ egységgömb, $\mathcal{F} = \{x_1 \geq 0\}$. Először ezt az esetet vizsgáljuk.

- Metszetükre mint felső félgömbre hivatkozunk. Ide esik az $e_1 = (1,0,0,\ldots)$ pont, amire mint a gömb „északi pólusa" hivatkozunk.

- Az e_1 irányú origón átmenő egyenes körüli fogatások a felső félgömb szimmetriái.

Egy egyszerű eset

- Legyen $\mathcal{E} = B(1,0)$ egységgömb, $\mathcal{F} = \{x_1 \geq 0\}$. Először ezt az esetet vizsgáljuk.

- Metszetükre mint felső félgömbre hivatkozunk. Ide esik az $e_1 = (1, 0, 0, \dots)$ pont, amire mint a gömb „északi pólusa" hivatkozunk.

- Az e_1 irányú origón átmenő egyenes körüli fogatások a felső félgömb szimmetriái.

- Azon ellipszoidok közül keressük meg a legkisebb térfogatút, amelyek tartalmazzák a felső félgömböt és a fenti forgásszimmetriájuk megvan.

Az egyszerű eset „versenyzői”

Az egyszerű eset „versenyzői”

- Az ilyen, félgömb köré írt ellipszoidoknál az e_1 irányú „függőleges” fél-tengelyhossz a legyen.

Az egyszerű eset „versenyzői”

- Az ilyen, félgömb köré írt ellipszoidoknál az e_1 irányú „függőleges” fél-tengelyhossz a legyen. Az összes többi tengely azonos hosszúságú (b).

Az egyszerű eset „versenyzői”

- Az ilyen, félgömb köré írt ellipszoidoknál az e_1 irányú „függőleges” fél-tengelyhossz a legyen. Az összes többi tengely azonos hosszúságú (b).
- Középpontja az x_1 koordinátatengely egy pontja, vagyis $(m, 0, 0, \dots)$.

Az egyszerű eset „versenyzői”

- Az ilyen, félgömb köré írt ellipszoidoknál az e_1 irányú „függőleges” fél-tengelyhossz a legyen. Az összes többi tengely azonos hosszúságú (b).
- Középpontja az x_1 koordinátatengely egy pontja, vagyis $(m, 0, 0, \dots)$.
- Ezek alapján az ellipszoid egyenlete:

$$\frac{(x_1 - m)^2}{a^2} + \frac{\sum_{i=2} x_i^2}{b^2} \leq 1$$

Az egyszerű eset „versenyzői"

- Az ilyen, félgömb köré írt ellipszoidoknál az e_1 irányú „függőleges" fél-tengelyhossz a legyen. Az összes többi tengely azonos hosszúságú (b).
- Középpontja az x_1 koordinátatengely egy pontja, vagyis $(m, 0, 0, \dots)$.
- Ezek alapján az ellipszoid egyenlete:

$$\frac{(x_1 - m)^2}{a^2} + \frac{\sum_{i=2} x_i^2}{b^2} \leq 1$$

- A forgásszimmetria miatt az ellipszoid felületén lesz a gömb „északi pólusa" $(1, 0, 0, \dots)$ pontja, illetve az $x_1 = 0$ síkba eső főkör.

Az egyszerű eset „versenyzői"

- Az ilyen, félgömb köré írt ellipszoidoknál az e_1 irányú „függőleges" fél-tengelyhossz a legyen. Az összes többi tengely azonos hosszúságú (b).
- Középpontja az x_1 koordinátatengely egy pontja, vagyis $(m, 0, 0, \dots)$.
- Ezek alapján az ellipszoid egyenlete:

$$\frac{(x_1 - m)^2}{a^2} + \frac{\sum_{i=2} x_i^2}{b^2} \leq 1$$

- A forgásszimmetria miatt az ellipszoid felületén lesz a gömb „északi pólusa" $(1, 0, 0, \dots)$ pontja, illetve az $x_1 = 0$ síkba eső főkör.
- Ebből látható, hogy

$$\frac{(1 - m)^2}{a^2} = 1, \qquad \frac{m^2}{a^2} + \frac{1}{b^2} = 1,$$

amiből következik, hogy $a = 1 - m$ és $b = \frac{1-m}{\sqrt{1-2m}}$.

Az egyszerű eset optimuma: a számolás

Az egyszerű eset optimuma: a számolás

- Ennek az ellipszoidnak a térfogatát szeretnénk m függvényében minimalizálni.

$$Vol(\mathcal{E}) = \sqrt{\det P} \cdot \beta_n = a \cdot b^{n-1} \cdot \beta_n = (1-m)^n \cdot \frac{1}{(1-2m)^{\frac{n-1}{2}}} \cdot \beta_n$$

β_n konstans, így a maradék kifejezést kell minimalizálni.

Az egyszerű eset optimuma: a számolás

- Ennek az ellipszoidnak a térfogatát szeretnénk m függvényében minimalizálni.

$$Vol(\mathcal{E}) = \sqrt{\det P} \cdot \beta_n = a \cdot b^{n-1} \cdot \beta_n = (1-m)^n \cdot \frac{1}{(1-2m)^{\frac{n-1}{2}}} \cdot \beta_n$$

β_n konstans, így a maradék kifejezést kell minimalizálni.

- Deriválás elvégzése, majd átalakítások után kapjuk, hogy a derivált

$$\frac{(1-2m)^{(\frac{n-1}{2})-1}(1-m)^{n-1}}{(1-2m)^{n-1}}\left((n+1)m-1\right).$$

Az egyszerű eset optimuma: a számolás

- Ennek az ellipszoidnak a térfogatát szeretnénk m függvényében minimalizálni.

$$Vol(\mathcal{E}) = \sqrt{\det P} \cdot \beta_n = a \cdot b^{n-1} \cdot \beta_n = (1-m)^n \cdot \frac{1}{(1-2m)^{\frac{n-1}{2}}} \cdot \beta_n$$

β_n konstans, így a maradék kifejezést kell minimalizálni.

- Deriválás elvégzése, majd átalakítások után kapjuk, hogy a derivált

$$\frac{(1-2m)^{(\frac{n-1}{2})-1}(1-m)^{n-1}}{(1-2m)^{n-1}}\left((n+1)m-1\right).$$

- Ezt 0-val egyenlővé téve (szélsőérték létezésének szükséges feltétele) az $m = \frac{1}{n+1}$ egyenlőséget (ahol a szükséges feltétel elegendősége is könnyen átlátható).

Az egyszerű eset optimuma: a tétel

Az egyszerű eset optimuma: a tétel

Tétel

Az egyszerű eset optimuma: a tétel

Tétel

• A felső félgömböt magában foglaló, forgásszimmetrikus ellipszoidok között a legkisebb térfogatú paraméterei

$$P = \begin{pmatrix} \frac{n^2}{(n+1)^2} & 0 & \dots & 0 & 0 \\ 0 & \frac{n^2}{n^2-1} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{n^2}{n^2-1} & 0 \\ 0 & 0 & \dots & 0 & \frac{n^2}{n^2-1} \end{pmatrix} = \frac{n^2}{n^2-1}\left(I - \frac{2}{n+1} e_1 e_1^T\right).$$

$$o = \left(\frac{1}{n+1}, 0, \dots, 0, 0\right)$$

paraméterekkel leírt ellipszoid.

Az egyszerű eset optimuma: az előrelépés

Az egyszerű eset optimuma: az előrelépés

A félgömb köré írt ellipszis térfogata a gömb térfogatához képest:

$$\mathrm{vol}\left(\mathcal{E}\left(P, \frac{1}{n+1}\right)\right) \leq \sqrt{\left(1+\frac{1}{n-1}\right)^{n-1}\left(1-\frac{1}{n+1}\right)^{n+1}}\,\mathrm{vol}(\mathcal{B}(1,0)).$$

Az egyszerű eset optimuma: az előrelépés

A félgömb köré írt ellipszis térfogata a gömb térfogatához képest:

$$\mathrm{vol}\left(\mathcal{E}\left(P, \frac{1}{n+1}\right)\right) \leq \sqrt{\left(1+\frac{1}{n-1}\right)^{n-1}\left(1-\frac{1}{n+1}\right)^{n+1}}\mathrm{vol}(\mathcal{B}(1,0)).$$

Ez egy „összeesés”.

Az egyszerű eset optimuma: az előrelépés

A félgömb köré írt ellipszis térfogata a gömb térfogatához képest:

$$\mathrm{vol}\left(\mathcal{E}\left(P, \frac{1}{n+1}\right)\right) \leq \sqrt{\left(1+\frac{1}{n-1}\right)^{n-1}\left(1-\frac{1}{n+1}\right)^{n+1}}\mathrm{vol}(\mathcal{B}(1,0)).$$

Ez egy „összeesés”. Ezt bizonyítás nélkül közöljük.

Az egyszerű eset optimuma: az előrelépés

A félgömb köré írt ellipszis térfogata a gömb térfogatához képest:

$$\mathrm{vol}\left(\mathcal{E}\left(P, \frac{1}{n+1}\right)\right) \leq \sqrt{\left(1+\frac{1}{n-1}\right)^{n-1}\left(1-\frac{1}{n+1}\right)^{n+1}} \mathrm{vol}(\mathcal{B}(1,0)).$$

Ez egy „összeesés”. Ezt bizonyítás nélkül közöljük.

Lemma

$$\left(1+\frac{1}{n-1}\right)^{n-1}\left(1-\frac{1}{n+1}\right)^{n+1} \leq e^{-1/2n} < 1.$$

Egy kis geometria

Egy kis geometria

Matematikai intuíció azt sugalja, hogy a fenti megtalált ellipszoid az összes ellipszoid között a legkisebb térfogatú. Ez valóban így van, de ezt nem igazoljuk.

Egy kis geometria

Matematikai intuíció azt sugalja, hogy a fenti megtalált ellipszoid az összes ellipszoid között a legkisebb térfogatú. Ez valóban így van, de ezt nem igazoljuk.

Tétel

A fent leírt ellipszoid a felső félgömb Löwner—John-ellipszoidja.

Egy kis geometria

Matematikai intuíció azt sugalja, hogy a fenti megtalált ellipszoid az összes ellipszoid között a legkisebb térfogatú. Ez valóban így van, de ezt nem igazoljuk.

Tétel

A fent leírt ellipszoid a felső félgömb Löwner—John-ellipszoidja.

Vegyük észre, hogy a felsőfélgömb esetén ha a megtalált köréírt ellipszoidot középpontjából $n+1$-ed részére lekicisnyítjük, akkor egy beírt ellipszoidot kapunk. Ez nem véletlen.

Egy kis geometria

Matematikai intuíció azt sugalja, hogy a fenti megtalált ellipszoid az összes ellipszoid között a legkisebb térfogatú. Ez valóban így van, de ezt nem igazoljuk.

Tétel

A fent leírt ellipszoid a felső félgömb Löwner—John-ellipszoidja.

Vegyük észre, hogy a felsőfélgömb esetén ha a megtalált köréírt ellipszoidot középpontjából $n+1$-ed részére lekicisnyítjük, akkor egy beírt ellipszoidot kapunk. Ez nem véletlen.

Löwner—John-tétel

Adott K kompakt, konvex test. A köréírható minimális térfogatú $\mathcal{E}$ ellipszoid (a test Löwner—John-ellipszoidja) olyan, hogy

$$\frac{1}{n+1}\cdot\mathcal{E}\subseteq K\subseteq\mathcal{E}.$$

Az általános eset

Az általános eset

Update formulák

Vizsgáljuk az $\mathcal{E}(P,o) := \{x : (x-o)^T P^{-1}(x-o)\}$ ellipszoid $\mathcal{F}(g,o) := \{x : g^T x \leq g^T o\}$ féltérrel vett metszetét:

$$\mathcal{E}(P,o)/g := \{x : (x-o)^T P^{-1}(x-o)\} \cap \{x : g^T x \leq g^T o\}$$

Vezessük be a következő jelölést: $\widetilde{g} = \frac{1}{\sqrt{g^T P g}} g$. Legyen

$$o^+ = o - \frac{1}{n+1} P\widetilde{g}, \qquad P^+ = \frac{n^2}{n^2-1}\left(P - \frac{2}{n+1} P\widetilde{g}\widetilde{g}^T\right).$$

Ekkor

$$\mathcal{E}(P,o)/g \subset \mathcal{E}(P^+,o^+),$$

$$vol(\mathcal{E}(P^+,o^+)) = \left(1+\frac{1}{n-1}\right)^{n-1}\left(1-\frac{1}{n+1}\right)^{n+1} \cdot vol(\mathcal{E}(P,o)).$$

Update szabály bizonyítása: a séma

Update szabály bizonyítása: a séma

- Affin transzformációk sorozatával az $\mathcal{E}(P,o)$, $\mathcal{F}(g,o)$ párt $\mathcal{B}(1,\vec{0}) = \mathcal{E}(I,\vec{0})$, $\mathcal{F}(e_1,\vec{0}) = \{x : x_1 \geq 0\}$ párba visszük át.

Update szabály bizonyítása: a séma

- Affin transzformációk sorozatával az $\mathcal{E}(P, o)$, $\mathcal{F}(g, o)$ párt $\mathcal{B}(1, \vec{0}) = \mathcal{E}(I, \vec{0})$, $\mathcal{F}(e_1, \vec{0}) = \{x : x_1 \geq 0\}$ párba visszük át.
- Metszetük köré a korábban megkonstruált ellipszoidot írjuk.

Update szabály bizonyítása: a séma

- Affin transzformációk sorozatával az $\mathcal{E}(P,o)$, $\mathcal{F}(g,o)$ párt $\mathcal{B}(1,\vec{0}) = \mathcal{E}(I,\vec{0})$, $\mathcal{F}(e_1,\vec{0}) = \{x : x_1 \geq 0\}$ párba visszük át.

- Metszetük köré a korábban megkonstruált ellipszoidot írjuk.

- Az affin transzformációk inverzeit alkalmazva visszatérünk az eredeti félellipszoidhoz.

Update szabály bizonyítása: a séma

- Affin transzformációk sorozatával az $\mathcal{E}(P, o)$, $\mathcal{F}(g, o)$ párt $\mathcal{B}(1, \vec{0}) = \mathcal{E}(I, \vec{0})$, $\mathcal{F}(e_1, \vec{0}) = \{x : x_1 \geq 0\}$ párba visszük át.

- Metszetük köré a korábban megkonstruált ellipszoidot írjuk.

- Az affin transzformációk inverzeit alkalmazva visszatérünk az eredeti félellipszoidhoz.

- Kapjuk a tétel állításában szepelő köréírt ellipszoidot.

Update szabály bizonyítása: a séma

- Affin transzformációk sorozatával az $\mathcal{E}(P, o)$, $\mathcal{F}(g, o)$ párt $\mathcal{B}(1, \vec{0}) = \mathcal{E}(I, \vec{0})$, $\mathcal{F}(e_1, \vec{0}) = \{x : x_1 \geq 0\}$ párba visszük át.

- Metszetük köré a korábban megkonstruált ellipszoidot írjuk.

- Az affin transzformációk inverzeit alkalmazva visszatérünk az eredeti félellipszoidhoz.

- Kapjuk a tétel állításában szepelő köréírt ellipszoidot.

- A hosszadalmas számolás lényege, hogy az ellipszoid algoritmus programjához kellenek a számolás végén kapott képletek.

Update szabály bizonyítása: az eltolás

Update szabály bizonyítása: az eltolás

Az első transzformáció legyen egy eltolás.

$$A_1: \quad x = x' + o \qquad / \qquad x' = x - o.$$

Update szabály bizonyítása: az eltolás

Az első transzformáció legyen egy eltolás.

$$A_1: \quad x = x' + o \qquad / \qquad x' = x - o.$$

Az x' koordinátarendszerben ellipszoidunk középpontja már az origó, félterünk határa áthalad az origón

$$\mathcal{E} = \{x' : {x'}^T P^{-1} x'\}, \qquad \mathcal{F}\{x' : g^T x' \leq 0\}$$

Update szabály bizonyítása: a kikerekítés

Update szabály bizonyítása: a kikerekítés

A második transzformáció gömbbé kerekíti az ellipszoidot (a középpont az origóban marad).

$$A_2: \quad x' = P^{1/2}x'' \qquad / \qquad x'' = P^{-1/2}x'.$$

Update szabály bizonyítása: a kikerekítés

A második transzformáció gömbbé kerekíti az ellipszoidot (a középpont az origóban marad).

$$A_2: \quad x' = P^{1/2}x'' \qquad / \qquad x'' = P^{-1/2}x'.$$

Az x'' koordinátarendszerben ellipszoidunk már origó középpontú, félterünk határa áthalad az origón

$$\mathcal{E} = \{x'' : {x''}^T x'' \leq 1\},$$

$$\mathcal{F} = \{x'' : (P^{1/2}g)^T x'' \leq 0\} = \{x'' : \frac{1}{\sqrt{g^T P g}}(P^{1/2}g)^T x'' \leq 0\}.$$

Update szabály bizonyítása: a kikerekítés

A második transzformáció gömbbé kerekíti az ellipszoidot (a középpont az origóban marad).

$$A_2: \quad x' = P^{1/2}x'' \qquad / \qquad x'' = P^{-1/2}x'.$$

Az x'' koordinátarendszerben ellipszoidunk már origó középpontú, félterünk határa áthalad az origón

$$\mathcal{E} = \{x'' : {x''}^T x'' \leq 1\},$$
$$\mathcal{F} = \{x'' : (P^{1/2}g)^T x'' \leq 0\} = \{x'' : \frac{1}{\sqrt{g^T P g}}(P^{1/2}g)^T x'' \leq 0\}.$$

Az utolsó átalakítás csupán a féltér normálvektorát írta át egységvektorrá, a féltér mint koordináta-halmaz megmaradt. Ami hátramaradt az a féltér normálvektorának $-e_1$-be forgatása (ezért kellett az egységvektorra áttérnünk).

Update szabály bizonyítása: a forgatás

Update szabály bizonyítása: a forgatás

A harmadik transzformáció egy origó középpontú forgatás (így gömbünk invariáns lesz), amely a $\frac{1}{\sqrt{g^T P g}}(P^{1/2}g)^T$ vektort $-e_1$-be viszi.

Update szabály bizonyítása: a forgatás

A harmadik transzformáció egy origó középpontú forgatás (így gömbünk invariáns lesz), amely a $\frac{1}{\sqrt{g^T P g}}(P^{1/2}g)^T$ vektort $-e_1$-be viszi.

Erre több lehetőség van, Mindegy melyiket választjuk, mátrixa legyen a Q ortogonális mátrix.

$$A_3: \quad x'' = Qx''' \qquad / \qquad x''' = Q^{-1}x'' = Q^T x''.$$

Update szabály bizonyítása: a forgatás

A harmadik transzformáció egy origó középpontú forgatás (így gömbünk invariáns lesz), amely a $\frac{1}{\sqrt{g^T P g}}(P^{1/2}g)^T$ vektort $-e_1$-be viszi.

Erre több lehetőség van, Mindegy melyiket választjuk, mátrixa legyen a Q ortogonális mátrix.

$$A_3 : \quad x'' = Qx''' \qquad / \qquad x''' = Q^{-1}x'' = Q^T x''.$$

Az x''' koordinátarendszerben elértük célunkat:

$$\mathcal{E} = \{x''' : x'''^T x''' \leq 1\}, \qquad \mathcal{F} = \{x''' : -e_1{}^T x''' \leq 0\} = \{x''' : x_1''' \geq 0\}.$$

Update szabály bizonyítása: a forgatás

A harmadik transzformáció egy origó középpontú forgatás (így gömbünk invariáns lesz), amely a $\frac{1}{\sqrt{g^T P g}}(P^{1/2}g)^T$ vektort $-e_1$-be viszi.

Erre több lehetőség van, Mindegy melyiket választjuk, mátrixa legyen a Q ortogonális mátrix.

$$A_3: \quad x'' = Qx''' \qquad / \qquad x''' = Q^{-1}x'' = Q^T x''.$$

Az x''' koordinátarendszerben elértük célunkat:

$$\mathcal{E} = \{x''' : {x'''}^T x''' \leq 1\}, \qquad \mathcal{F} = \{x''' : -{e_1}^T x''' \leq 0\} = \{x''' : x_1''' \geq 0\}.$$

Célunk eléréséhez

$$\frac{1}{\sqrt{g^T P g}}(P^{1/2}g)^T Q = -e_1^T, \text{ azaz } Qe_1 = -\frac{1}{\sqrt{g^T P g}} P^{1/2} g = -P^{1/2}\widetilde{g}$$

kellett.

Update szabály bizonyítása: az egyszerű eset alkalmazása

Update szabály bizonyítása: az egyszerű eset alkalmazása

Az egyszerű eset alapján az új koordinátarendszerben fel tudjuk írni a félgömböt tartalmazó optimális ellipszoidot leíró formulát:

Update szabály bizonyítása: az egyszerű eset alkalmazása

Az egyszerű eset alapján az új koordinátarendszerben fel tudjuk írni a félgömböt tartalmazó optimális ellipszoidot leíró formulát:

$$\mathcal{K} = \{x''' : (x''' - \frac{1}{n+1}e_1)^T \left(\frac{n^2}{n^2-1}(I - \frac{2}{n+1}e_1 e_1^T)\right)^{-1} \cdot \\ \cdot (x''' - \frac{1}{n+1}e_1) \leq 1\}$$

Update szabály bizonyítása: az egyszerű eset alkalmazása

Az egyszerű eset alapján az új koordinátarendszerben fel tudjuk írni a félgömböt tartalmazó optimális ellipszoidot leíró formulát:

$$\mathcal{K} = \{x''' : (x''' - \frac{1}{n+1} e_1)^T \left(\frac{n^2}{n^2-1} (I - \frac{2}{n+1} e_1 e_1^T) \right)^{-1} \cdot \\ \cdot (x''' - \frac{1}{n+1} e_1) \leq 1\}$$

Ezt kell rendre vissza írni a bevezetett koordinátarendszereken keresztül az eredetibe.

Update szabály bizonyítása: az egyszerű eset alkalmazása

Az egyszerű eset alapján az új koordinátarendszerben fel tudjuk írni a félgömböt tartalmazó optimális ellipszoidot leíró formulát:

$$\mathcal{K} = \{x''' : (x''' - \frac{1}{n+1}e_1)^T \left(\frac{n^2}{n^2-1}(I - \frac{2}{n+1}e_1 e_1^T) \right)^{-1} \cdot \\ \cdot (x''' - \frac{1}{n+1}e_1) \leq 1\}$$

Ezt kell rendre vissza írni a bevezetett koordinátarendszereken keresztül az eredetibe.

Az első lépés az x'' koodinátarendszerbe való átírás:

Update szabály bizonyítása: visszaforgatás

$$
\begin{aligned}
\mathcal{K} = \{x'' : (Q^T x'' - \frac{1}{n+1} e_1)^T \left(\frac{n^2}{n^2-1} (I - \frac{2}{n+1} e_1 e_1^T) \right)^{-1} \cdot \\
\cdot (Q^T x'' - \frac{1}{n+1} e_1) \leq 1\} \\
= \{x'' : (Q^T (x'' - \frac{1}{n+1} Q e_1))^T \left(\frac{n^2}{n^2-1} (I - \frac{2}{n+1} e_1 e_1^T) \right)^{-1} \cdot \\
\cdot Q^T (x'' - \frac{1}{n+1} Q e_1) \leq 1\} \\
= \{x'' : (x'' - \frac{1}{n+1} Q e_1)^T Q \left(\frac{n^2}{n^2-1} (I - \frac{2}{n+1} e_1 e_1^T) \right)^{-1} Q^T \cdot \\
\cdot (x'' - \frac{1}{n+1} Q e_1) \leq 1\}
\end{aligned}
$$

Update szabály bizonyítása: visszaforgatás II

$$\mathcal{K} = \{x'' : (x'' - \frac{1}{n+1}Qe_1)^T \left(Q\left(\frac{n^2}{n^2-1}(I - \frac{2}{n+1}e_1e_1^T\right)^{-1} Q^T\right) \cdot$$
$$\cdot (x'' - \frac{1}{n+1}Qe_1) \leq 1\}$$
$$= \{x'' : (x'' - \frac{1}{n+1}Qe_1)^T \left(Q\left(\frac{n^2}{n^2-1}(I - \frac{2}{n+1}e_1e_1^T\right) Q^T\right)^{-1} \cdot$$
$$\cdot (x'' - \frac{1}{n+1}Qe_1) \leq 1\}$$
$$= \{x'' : (x'' - \frac{1}{n+1}Qe_1)^T \left(\frac{n^2}{n^2-1}\left(I - \frac{2}{n+1}Qe_1e_1^TQ^T\right)\right)^{-1} \cdot$$
$$\cdot (x'' - \frac{1}{n+1}Qe_1) \leq 1\}$$

Update szabály bizonyítása: visszaforgatás III

$$\mathcal{K} = \{x'' : (x'' + \frac{1}{n+1}P^{1/2}\widetilde{g})^T \left(\frac{n^2}{n^2-1}\left(I - \frac{2}{n+1}(-P^{1/2}\widetilde{g})(-P^{1/2}\widetilde{g})^T\right)\right)^{-1} \cdot$$

$$\cdot (x'' + \frac{1}{n+1}P^{1/2}\widetilde{g}) \leq 1\}$$

$$= \{x'' : (x'' + \frac{1}{n+1}P^{1/2}\widetilde{g})^T \left(\frac{n^2}{n^2-1}\left(I - \frac{2}{n+1}P^{1/2}\widetilde{g}\widetilde{g}^T P^{1/2}\right)\right)^{-1} \cdot$$

$$(x'' + \frac{1}{n+1}P^{1/2}\widetilde{g}) \leq 1\}.$$

Update szabály bizonyítása: visszalapítás

Update szabály bizonyítása: visszalapítás

Az x' koodinátarendszerbe való átírás:

$$\mathcal{K} = \{x' : (P^{-1/2}x' + \frac{1}{n+1}P^{1/2}\widetilde{g})^T \cdot \left(\frac{n^2}{n^2-1}\left(I - \frac{2}{n+1}P^{1/2}\widetilde{g}\widetilde{g}^T P^{1/2}\right)\right)^{-1} \cdot (P^{-1/2}x' + \frac{1}{n+1}P^{1/2}\widetilde{g}) \leq 1\}$$

$$\mathcal{K} = \{x' : (P^{-1/2}(x' + \frac{1}{n+1}P\widetilde{g}))^T \left(\frac{n^2}{n^2-1}\left(I - \frac{2}{n+1}P^{1/2}\widetilde{g}\widetilde{g}^T P^{1/2}\right)\right)^{-1} \cdot P^{-1/2}(x' + \frac{1}{n+1}P\widetilde{g}) \leq 1\}$$

Update szabály bizonyítása: visszalapítás II

$$\mathcal{K} = \{x' : (x' + \frac{1}{n+1}P\widetilde{g})^T \cdot P^{-1/2}\left(\frac{n^2}{n^2-1}\left(I - \frac{2}{n+1}P^{1/2}\widetilde{g}\widetilde{g}^T P^{1/2}\right)\right)^{-1} P^{-1/2} \cdot (x' + \frac{1}{n+1}P\widetilde{g}) \leq 1\}$$

$$\mathcal{K} = \{x' : (x' + \frac{1}{n+1}P\widetilde{g})^T \left(\frac{n^2}{n^2-1}P^{1/2}\left(I - \frac{2}{n+1}P^{1/2}\widetilde{g}\widetilde{g}^T P^{1/2}\right)P^{1/2}\right)^{-1} \cdot (x' + \frac{1}{n+1}P\widetilde{g}) \leq 1\}$$

$$\mathcal{K} = \{x' : (x' + \frac{1}{n+1}P\widetilde{g})^T \left(\frac{n^2}{n^2-1}\left(P - \frac{2}{n+1}P\widetilde{g}\widetilde{g}^T P\right)\right)^{-1} \cdot (x' + \frac{1}{n+1}P\widetilde{g}) \leq 1\}$$

Update szabály bizonyításának vége: visszatolás

Update szabály bizonyításának vége: visszatolás

Végül az eredeti x koodinátarendszerben:

$$\mathcal{K} = \{x : (x - o + \frac{1}{n+1} P\widetilde{g})^T \left(\frac{n^2}{n^2-1} \left(P - \frac{2}{n+1} P\widetilde{g}\widetilde{g}^T P \right) \right)^{-1} \cdot$$
$$\cdot (x - o + \frac{1}{n+1} P\widetilde{g}) \leq 1\}$$

Update szabály bizonyításának vége: visszatolás

Végül az eredeti x koodinátarendszerben:

$$\mathcal{K} = \{x : (x-o+\frac{1}{n+1}P\widetilde{g})^T \left(\frac{n^2}{n^2-1}\left(P - \frac{2}{n+1}P\widetilde{g}\widetilde{g}^T P\right)\right)^{-1} \cdot \\ \cdot (x - o + \frac{1}{n+1}P\widetilde{g}) \leq 1\}$$

Azaz a kapott körülírt ellipszoid középpontja

$$o^+ = o - \frac{1}{n+1}P\widetilde{g},$$

Update szabály bizonyításának vége: visszatolás

Végül az eredeti x koodinátarendszerben:

$$\mathcal{K} = \{x : (x - o + \frac{1}{n+1} P\widetilde{g})^T \left(\frac{n^2}{n^2-1} \left(P - \frac{2}{n+1} P\widetilde{g}\widetilde{g}^T P \right) \right)^{-1} \cdot (x - o + \frac{1}{n+1} P\widetilde{g}) \leq 1\}$$

Azaz a kapott körülírt ellipszoid középpontja

$$o^+ = o - \frac{1}{n+1} P\widetilde{g},$$

mátrixa

$$P^+ = \frac{n^2}{n^2-1} \left(P - \frac{2}{n+1} P\widetilde{g}\widetilde{g}^T P \right),$$

Update szabály bizonyításának vége: visszatolás

Végül az eredeti x koodinátarendszerben:

$$\mathcal{K} = \{x : (x - o + \frac{1}{n+1}P\widetilde{g})^T \left(\frac{n^2}{n^2-1}\left(P - \frac{2}{n+1}P\widetilde{g}\widetilde{g}^T P\right)\right)^{-1} \cdot (x - o + \frac{1}{n+1}P\widetilde{g}) \leq 1\}$$

Azaz a kapott körülírt ellipszoid középpontja

$$o^+ = o - \frac{1}{n+1}P\widetilde{g},$$

mátrixa

$$P^+ = \frac{n^2}{n^2-1}\left(P - \frac{2}{n+1}P\widetilde{g}\widetilde{g}^T P\right),$$

ahogy bizonyítani kellett.

Szünet

Az ellipszoidokon alapuló algoritmusok: Az input

Az ellipszoidokon alapuló algoritmusok: Az input

Szeparációs orákulum

Legyen $\mathcal{L}$ egy zárt konvex halmaz $\mathbb{R}^d$-ben.

Az ellipszoidokon alapuló algoritmusok: Az input

Szeparációs orákulum

Legyen $\mathcal{L}$ egy zárt konvex halmaz $\mathbb{R}^d$-ben. Azt mondjuk, hogy $\mathcal{L}$ rendelkezik egy hatékony szeparációs algoritmussal/orákulummal, ha van olyan polinomiális algoritmus ami egy $x \in \mathbb{Q}^d$-beli pontról megmondja, hogy $x \in \mathcal{L}$ vagy ad egy hipersíkot, amely egyik nyilt félterébe esik x, másik zárt félterébe esik $\mathcal{L}$.

Az ellipszoidokon alapuló algoritmusok: Az input

Szeparációs orákulum

Legyen $\mathcal{L}$ egy zárt konvex halmaz $\mathbb{R}^d$-ben. Azt mondjuk, hogy $\mathcal{L}$ rendelkezik egy hatékony szeparációs algoritmussal/orákulummal, ha van olyan polinomiális algoritmus ami egy $x \in \mathbb{Q}^d$-beli pontról megmondja, hogy $x \in \mathcal{L}$ vagy ad egy hipersíkot, amely egyik nyilt félterébe esik x, másik zárt félterébe esik $\mathcal{L}$.

- A definícióban szereplő polinomiálitás kérdéses.

Az ellipszoidokon alapuló algoritmusok: Az input

Szeparációs orákulum

Legyen $\mathcal{L}$ egy zárt konvex halmaz $\mathbb{R}^d$-ben. Azt mondjuk, hogy $\mathcal{L}$ rendelkezik egy hatékony szeparációs algoritmussal/orákulummal, ha van olyan polinomiális algoritmus ami egy $x \in \mathbb{Q}^d$-beli pontról megmondja, hogy $x \in \mathcal{L}$ vagy ad egy hipersíkot, amely egyik nyilt félterébe esik x, másik zárt félterébe esik $\mathcal{L}$.

- A definícióban szereplő polinomiálitás kérdéses.

- Az algoritmushoz egy $\mathcal{L}$ halmaz és egy x pont szükséges. $\mathcal{L}$ kódolásának hosszát tisztázni kell.

Az ellipszoidokon alapuló algoritmusok: Az input, példák

Az ellipszoidokon alapuló algoritmusok: Az input, példák

Példa

$\mathcal{L}$: egy explicit módon felírt LP feladat lehetséges megoldásainak halmaza.

Az ellipszoidokon alapuló algoritmusok: Az input, példák

Példa

$\mathcal{L}$: egy explicit módon felírt LP feladat lehetséges megoldásainak halmaza.

$\mathcal{L}$ nyilván rendelkezik hatékony szeparációs algoritmussal, ha $\mathcal{L}$ méretét mátrixával definiáljuk.

Az ellipszoidokon alapuló algoritmusok: Az input, példák

Példa

$\mathcal{L}$: egy explicit módon felírt LP feladat lehetséges megoldásainak halmaza.

$\mathcal{L}$ nyilván rendelkezik hatékony szeparációs algoritmussal, ha $\mathcal{L}$ méretét mátrixával definiáljuk.

Ekkor „van időnk” az összes definiáló egyenlőtlenség ellenőrzésére.

Az ellipszoidokon alapuló algoritmusok: Az input, példák

Példa

$\mathcal{L}$: egy explicit módon felírt LP feladat lehetséges megoldásainak halmaza.

$\mathcal{L}$ nyilván rendelkezik hatékony szeparációs algoritmussal, ha $\mathcal{L}$ méretét mátrixával definiáljuk.

Ekkor „van időnk" az összes definiáló egyenlőtlenség ellenőrzésére.

Példa

Az $\mathcal{MP}(G)$ párosítási politóp is rendelkezik hatékony szeparációs algoritmussal akkor is, ha kódolása a kiinduló gráffal történik.

Az ellipszoidokon alapuló algoritmusok: Az input, példák

Példa

$\mathcal{L}$: egy explicit módon felírt LP feladat lehetséges megoldásainak halmaza.

$\mathcal{L}$ nyilván rendelkezik hatékony szeparációs algoritmussal, ha $\mathcal{L}$ méretét mátrixával definiáljuk.

Ekkor „van időnk” az összes definiáló egyenlőtlenség ellenőrzésére.

Példa

Az $\mathcal{MP}(G)$ párosítási politóp is rendelkezik hatékony szeparációs algoritmussal akkor is, ha kódolása a kiinduló gráffal történik.

Itt is ismert explicit leírás (Edmonds-féle politóp tétel).

Az ellipszoidokon alapuló algoritmusok: Az input, példák

Példa

$\mathcal{L}$: egy explicit módon felírt LP feladat lehetséges megoldásainak halmaza.

$\mathcal{L}$ nyilván rendelkezik hatékony szeparációs algoritmussal, ha $\mathcal{L}$ méretét mátrixával definiáljuk.

Ekkor „van időnk” az összes definiáló egyenlőtlenség ellenőrzésére.

Példa

Az $\mathcal{MP}(G)$ párosítási politóp is rendelkezik hatékony szeparációs algoritmussal akkor is, ha kódolása a kiinduló gráffal történik.

Itt is ismert explicit leírás (Edmonds-féle politóp tétel).

A leíró egyenlőtlenségek ellenőrzése azonban exponenciális lehet a gráf méretében.

Az ellipszoidokon alapuló algoritmusok: Az input, példák

Példa

$\mathcal{L}$: egy explicit módon felírt LP feladat lehetséges megoldásainak halmaza.

$\mathcal{L}$ nyilván rendelkezik hatékony szeparációs algoritmussal, ha $\mathcal{L}$ méretét mátrixával definiáljuk.

Ekkor „van időnk” az összes definiáló egyenlőtlenség ellenőrzésére.

Példa

Az $\mathcal{MP}(G)$ párosítási politóp is rendelkezik hatékony szeparációs algoritmussal akkor is, ha kódolása a kiinduló gráffal történik.

Itt is ismert explicit leírás (Edmonds-féle politóp tétel).

A leíró egyenlőtlenségek ellenőrzése azonban exponenciális lehet a gráf méretében.

A polinomiális (a gráf méretében) szeparációs orákulumot a Gomory—Hu-fa konstrukciójából és annak következményei adják.

Az ellipszoidokon alapuló algoritmusok: Alapfeladatok I

Az ellipszoidokon alapuló algoritmusok: Alapfeladatok I

Optimalizálás

Minimalizáljuk	$c(x)$-t
Feltéve, hogy	$x \in \mathcal{L} \subset \mathbb{R}^n$,

azaz az alap konvex optimalizálási feladat, ahol $\mathcal{L}$-hez van hatékony szeparációs orákulum.

Az ellipszoidokon alapuló algoritmusok: Alapfeladatok I

Optimalizálás

Minimalizáljuk	$c(x)$-t
Feltéve, hogy	$x \in \mathcal{L} \subset \mathbb{R}^n$,

azaz az alap konvex optimalizálási feladat, ahol $\mathcal{L}$-hez van hatékony szeparációs orákulum.

Igazából megelégszünk egy $x \in \mathcal{L}$ vektorral, amely ϵ szuboptimális, azaz $c(x) \geq p^* - \epsilon$.

Az ellipszoidokon alapuló algoritmusok: Alapfeladatok II

Az ellipszoidokon alapuló algoritmusok: Alapfeladatok II

Kielégíthetőség

Minimalizáljuk	0-t
Feltéve, hogy	$x \in \mathcal{L} \subset \mathbb{R}^n$,

azaz $\mathcal{L}$ nem-ürességének/a feltételrendszer kielégíthetőségének tesztelésére, ahol $\mathcal{L}$ egy hatékony szeparációs orákulummal adott.

Az ellipszoidokon alapuló algoritmusok: Alapfeladatok II

Kielégíthetőség

Minimalizáljuk	0-t
Feltéve, hogy	$x \in \mathcal{L} \subset \mathbb{R}^n$,

azaz $\mathcal{L}$ nem-ürességének/a feltételrendszer kielégíthetőségének tesztelésére, ahol $\mathcal{L}$ egy hatékony szeparációs orákulummal adott.

Ha $\mathcal{L}$ nem-üres, akkor eljárásunk outputja $\mathcal{L}$ egy tetszőleges pontja halmazunknak.

A két feladat kapcsolata

A két feladat kapcsolata

A tesztelési feladatot $\mathcal{L}$-SAT-nak nevezzük.

A két feladat kapcsolata

A tesztelési feladatot $\mathcal{L}$-SAT-nak nevezzük.

Jelölés: LP-SAT

Amennyiben $\mathcal{L}$ véges sok lineáris egyenlőtlenséggel definiált halmaz a probléma neve LP-SAT.

A két feladat kapcsolata

A tesztelési feladatot $\mathcal{L}$-SAT-nak nevezzük.

Jelölés: LP-SAT

Amennyiben $\mathcal{L}$ véges sok lineáris egyenlőtlenséggel definiált halmaz a probléma neve LP-SAT.

- A tesztelési feladat speciális esete az elsőnek. Ennek ellenére egy ezt megoldó szubrutin alkalmazható az optimalizálási feladat megoldására is.

A két feladat kapcsolata

A tesztelési feladatot $\mathcal{L}$-SAT-nak nevezzük.

Jelölés: LP-SAT

Amennyiben $\mathcal{L}$ véges sok lineáris egyenlőtlenséggel definiált halmaz a probléma neve LP-SAT.

- A tesztelési feladat speciális esete az elsőnek. Ennek ellenére egy ezt megoldó szubrutin alkalmazható az optimalizálási feladat megoldására is.
- A két feladat között mindkét irányban van hatékony redukció.

A két feladat kapcsolata

A tesztelési feladatot $\mathcal{L}$-SAT-nak nevezzük.

Jelölés: LP-SAT

Amennyiben $\mathcal{L}$ véges sok lineáris egyenlőtlenséggel definiált halmaz a probléma neve LP-SAT.

- A tesztelési feladat speciális esete az elsőnek. Ennek ellenére egy ezt megoldó szubrutin alkalmazható az optimalizálási feladat megoldására is.
- A két feladat között mindkét irányban van hatékony redukció.
- Az ehhez szükséges ötlet ismét a bináris keresés.

A két feladat kapcsolata: bináris keresés

A két feladat kapcsolata: bináris keresés

- Tegyük fel, hogy ismert egy I_0 zárt intervallum, ami tartalmazza p^*-t.

A két feladat kapcsolata: bináris keresés

- Tegyük fel, hogy ismert egy I_0 zárt intervallum, ami tartalmazza p^*-t.
- Legyen t_0 az intervallum középpontja. Nézzük meg, hogy $\mathcal{L}^+ = \mathcal{L} \cap \{x : c(x) \leq t_0\}$ üres-e. (Ha a kiinduló feladat LP, akkor az LP-SAT probléma fenti módosítása egy új LP-SAT kérdés lesz.)

A két feladat kapcsolata: bináris keresés

- Tegyük fel, hogy ismert egy I_0 zárt intervallum, ami tartalmazza p^*-t.
- Legyen t_0 az intervallum középpontja. Nézzük meg, hogy $\mathcal{L}^+ = \mathcal{L} \cap \{x : c(x) \leq t_0\}$ üres-e. (Ha a kiinduló feladat LP, akkor az LP-SAT probléma fenti módosítása egy új LP-SAT kérdés lesz.)
- Ha igen, akkor I_1 legyen I_0 felső félintervalluma.

A két feladat kapcsolata: bináris keresés

- Tegyük fel, hogy ismert egy I_0 zárt intervallum, ami tartalmazza p^*-t.
- Legyen t_0 az intervallum középpontja. Nézzük meg, hogy $\mathcal{L}^+ = \mathcal{L} \cap \{x : c(x) \leq t_0\}$ üres-e. (Ha a kiinduló feladat LP, akkor az LP-SAT probléma fenti módosítása egy új LP-SAT kérdés lesz.)
- Ha igen, akkor I_1 legyen I_0 felső félintervalluma.
- Ha nem, akkor I_1 legyen I_0 alsó félintervalluma.

A két feladat kapcsolata: bináris keresés

- Tegyük fel, hogy ismert egy I_0 zárt intervallum, ami tartalmazza p^*-t.
- Legyen t_0 az intervallum középpontja. Nézzük meg, hogy $\mathcal{L}^+ = \mathcal{L} \cap \{x : c(x) \leq t_0\}$ üres-e. (Ha a kiinduló feladat LP, akkor az LP-SAT probléma fenti módosítása egy új LP-SAT kérdés lesz.)
- Ha igen, akkor I_1 legyen I_0 felső félintervalluma.
- Ha nem, akkor I_1 legyen I_0 alsó félintervalluma.
- Iterálva a felezést ismét megkapjuk p^* egy jó közelítését és ebből könnyen adódik egy $x^{\approx}$ szuboptimális hely, azaz olyan hely, ahol $c(x^{\approx}) \geq p^* - \epsilon$ egy előírt ϵ-ra.

A két feladat kapcsolata: bináris keresés

- Tegyük fel, hogy ismert egy I_0 zárt intervallum, ami tartalmazza p^*-t.
- Legyen t_0 az intervallum középpontja. Nézzük meg, hogy $\mathcal{L}^+ = \mathcal{L} \cap \{x : c(x) \leq t_0\}$ üres-e. (Ha a kiinduló feladat LP, akkor az LP-SAT probléma fenti módosítása egy új LP-SAT kérdés lesz.)
- Ha igen, akkor I_1 legyen I_0 felső félintervalluma.
- Ha nem, akkor I_1 legyen I_0 alsó félintervalluma.
- Iterálva a felezést ismét megkapjuk p^* egy jó közelítését és ebből könnyen adódik egy $x^{\approx}$ szuboptimális hely, azaz olyan hely, ahol $c(x^{\approx}) \geq p^* - \epsilon$ egy előírt ϵ-ra.
- A redukció esete ellenére mindkét változatra tárgyaljuk az ellipszoid módszert.

Ellipszoid módszer: optimalizálás

Ellipszoid módszer (optimalizálásra):

Feltételezések:

Egy $\mathcal{L}$ konvex halmaz hatékony szeparációs orákulummal adott

c egy differenciálható célfüggvény

Adott egy $\mathcal{E}_{\text{kül}}$ ellipszoid, amely tartalmazza x^* optimális értéket (esetleg a teljes $\mathcal{L}$-et tartalmazó gömb).

Adott egy $\mathcal{E}_{\text{bel}}$ ellipszoid, amely része az $\mathcal{L}$ halmaznak.

// Az algoritmus során egy $\mathcal{E}_i$ ellipszist változtatunk.
// $\mathcal{E}_i$ az aktuális ellipszoid.
// Kezdetben $\mathcal{E}_0 = \mathcal{E}_{\text{kül}}$.

Ellipszoid módszer: optimalizálás (folytatás)

Ellipszoid módszer optimalizálásra (folytatás):
FELEZÉSI LÉPÉS: Vizsgáljuk meg, hogy ellipszoidunk o_i középpontja kielégíti-e a feltételeket. Ha
• NEM, akkor vegyünk egy szeparáló félteret o_i-n keresztül. Tegyük fel, hogy $\mathcal{F} = \{x : g^T x \leq g^T o_i\}$ féltér tartalmazza $\mathcal{E}_i \cap \mathcal{F}$-t.
// Ekkor az $\mathcal{E}_i/g$ félellipszoid is tartalmazza $\mathcal{E}_i \cap \mathcal{L}$-t, speciálisan x^*-t is.
• IGEN, akkor nézzük meg lokálisan a $c(x)$ célfüggvényt o_i-ben. Legyen g a gradiense ($g = \nabla c(o_i)$).
// Ekkor $c(x) \geq c(o_i) + g^T(x - o_i)$ minden $x \in \mathcal{L}$ esetén.
Legyen $\mathcal{F} = \{x : g^T x \leq g^T o_i\}$.
// A g által definiált (o_i-n átmenő) hipersíknak a gradiens irányával
// ellentétes oldalát írtuk le. Ha az $\mathcal{E}_i/g$ félellipszoidra szorítkozunk,
// akkor nem veszítjük el a keresett optimális helyet.
• *KILÉPÉSI FELTÉTEL:* Ha $\sqrt{g^T P g} \leq \epsilon$, akkor leállunk: outputunk o_i.
Ha nem, akkor továbbhaladunk.

Ellipszoid módszer: optimalizálás (folytatás)

// Azt várnánk, hogy továbbhaladás esetén a szüküló halmazunk
// új értéke $\mathcal{E}_i/g$ lesz. Most jön egy merész lépés.
// A félellipszoidhoz hozáadunk olyan részeket, amelyekről
// tudjuk, hogy feleslegek. A haszon: ismét elliszoiddal
// becsüljük a keresett helyet.
UPDATE LÉPÉS: Tegyük fel, hogy $\mathcal{E}_i = \mathcal{E}(P, o)$. Legyen

$$\widetilde{g} = \frac{1}{\sqrt{g^T P g}} g,$$

$$o^+ = o - \frac{1}{n+1} P\widetilde{g}, \qquad P^+ = \frac{n^2}{n^2-1}\left(P - \frac{2}{n+1} P\widetilde{g}\widetilde{g}^T P\right).$$

Legyen $\mathcal{E}_{i+1} = \mathcal{E}(P^+, o^+)$, $i \leftarrow i+1$. Térjünk vissza a felezési lépéshez.

Kiterjesztések

Kiterjesztések

- A módszer nem differenciálható célfüggvények esetén is alkalmazható, ha minden $x_0 \in \mathcal{F}$ esetén tudunk egy szubgradienst, azaz egy olyan g_{x_0} vektort, hogy $c(x) \geq c(x_0) + g_{x_0}^T(x - x_0)$.

Kiterjesztések

- A módszer nem differenciálható célfüggvények esetén is alkalmazható, ha minden $x_0 \in \mathcal{F}$ esetén tudunk egy szubgradienst, azaz egy olyan g_{x_0} vektort, hogy $c(x) \geq c(x_0) + g_{x_0}^T(x - x_0)$.

- Ekkor ezzel a szubgradiens vektorral ugyanúgy dolgozhatunk mint, ahogy g-vel tettük a differenciálható esetben.

A kilépési feltétel jogossága (lineáris célfüggvény esetén)

A kilépési feltétel jogossága (lineáris célfüggvény esetén)

Tétel

Legyen $c(x) = c^T x$. Tegyük fel, hogy az i paraméter értéknél $\sqrt{g^T P g} \leq \epsilon$ (azaz $\sqrt{c^T P c} \leq \epsilon$, hiszen $g = \nabla c(o_i) = c$), akkor

$$c(o_k) - c(x^*) \leq \epsilon.$$

A kilépési feltétel jogossága (lineáris célfüggvény esetén)

Tétel

Legyen $c(x) = c^T x$. Tegyük fel, hogy az i paraméter értéknél $\sqrt{g^T P g} \leq \epsilon$ (azaz $\sqrt{c^T P c} \leq \epsilon$, hiszen $g = \nabla c(o_i) = c$), akkor

$$c(o_k) - c(x^*) \leq \epsilon.$$

$$c(x^*) = c^T x^* \geq \inf_{z \in \mathcal{E}(P, o_i)} c^T z =$$

$$c^T \left(o_i - \frac{Pc}{|P^{1/2} c|} \right) = c^T o_i - \frac{c^T P c}{\sqrt{(P^{1/2} c)^T P^{1/2} c}} =$$

$$= c(o_i) - \sqrt{c^T P c}.$$

A kilépési feltétel jogossága (lineáris célfüggvény esetén)

Tétel

Legyen $c(x) = c^T x$. Tegyük fel, hogy az i paraméter értéknél $\sqrt{g^T P g} \leq \epsilon$ (azaz $\sqrt{c^T P c} \leq \epsilon$, hiszen $g = \nabla c(o_i) = c$), akkor

$$c(o_k) - c(x^*) \leq \epsilon.$$

$$c(x^*) = c^T x^* \geq \inf_{z \in \mathcal{E}(P, o_i)} c^T z =$$

$$c^T \left(o_i - \frac{Pc}{|P^{1/2}c|} \right) = c^T o_i - \frac{c^T Pc}{\sqrt{(P^{1/2}c)^T P^{1/2} c}} =$$

$$= c(o_i) - \sqrt{c^T Pc}.$$

A bizonyitásban szereplő minimalizálási feladat optimumának értékét korábban meghatároztuk, az ott bizonyított képlet alapján haladtunk tovább.

Ellipszoid módszer: tesztelés

Ellipszoid módszer (Kielégíthetőség tesztelésre):
Feltételezések:

Adott egy $\mathcal{E}_0$ ellipszoid, amely tartalmazza $\mathcal{L}$-et.

$\mathcal{L}$-re van egy hatékony szeparációs orákulum.

Van egy pozitív alsó becslésünk vol $\mathcal{L}$-re, amennyiben nem üres.

// Az algoritmus során egy $\mathcal{E}_i$ ellipszist változtunk. $\mathcal{E}_i$ az aktuális ellipszoid.
// Kezdetben $i = 0$. $\mathcal{E}_{i+1}$ mindig metszi $\mathcal{F}$-et, ha az nem üres.
FELEZÉSI LÉPÉS: Vizsgáljuk meg, hogy ellipszoidunk o_i középpontja $\mathcal{L}$-hez tartozik-e. Ha

- IGEN, akkor leállunk, o_i lesz a kiszámolt eleme $\mathcal{L}$-nek.
- NEM, akkor vegyünk egy szeparáló félteret o_i-n keresztül.

Ellipszoid módszer: tesztelés (folytatás)

Tegyük fel, hogy
$\mathcal{F} = \{x : g^T x \leq g^T o_i\}$ féltér tartalmazza $\mathcal{E}_i \cap \mathcal{F}$-t.
// Ekkor az $\mathcal{E}_i/g$ félellipszoid is tartalmazza $\mathcal{E}_i \cap \mathcal{L}$-t.
• *KILÉPÉSI FELTÉTEL:* Ha i „elég" nagy, akkor leállunk: outputunk „$\mathcal{L}$ üres".

Ha nem, akkor továbbhaladunk.
// Azt várnánk, hogy továbbhaladás esetén a szükülő halmazunk
// új értéke $\mathcal{E}_i/g$ lesz. Most jön egy merész lépés.
// A félellipszoidhoz hozáadunk olyan részeket, amelyekről tudjuk,
// hogy feleslegek. A haszon: ismét elliszoiddal becsüljük a vizsgált
// halmazt.

Ellipszoid módszer: tesztelés (folytatás)

UPDATE LÉPÉS: Tegyük fel, hogy $\mathcal{E}_i = \mathcal{E}(P, o)$. Legyen

$$\widetilde{g} = \frac{1}{\sqrt{g^T P g}} g,$$

$$o^+ = o - \frac{1}{n+1} P\widetilde{g}, \qquad P^+ = \frac{n^2}{n^2-1}\left(P - \frac{2}{n+1} P\widetilde{g}\widetilde{g}^T P\right).$$

Legyen $\mathcal{E}_{i+1} = \mathcal{E}(P^+, o^+)$, $i \leftarrow i+1$. Térjünk vissza a felezési lépéshez.

Az ördög a részletekben rejlik

Az ördög a részletekben rejlik

- A fenti leírás kissé vázlatos. Matematikailag nem írjuk le pontosan mikor állunk le. A részletekhez sajnos nincs elég időnk.

Az ördög a részletekben rejlik

- A fenti leírás kissé vázlatos. Matematikailag nem írjuk le pontosan mikor állunk le. A részletekhez sajnos nincs elég időnk.
- A fenti leírás már bizonyos programozási nyelvekben minden probléma nélkül kódolható.

Az ördög a részletekben rejlik

- A fenti leírás kissé vázlatos. Matematikailag nem írjuk le pontosan mikor állunk le. A részletekhez sajnos nincs elég időnk.
- A fenti leírás már bizonyos programozási nyelvekben minden probléma nélkül kódolható.
- Elméletileg azonban több sebből vérzik.

Az ördög a részletekben rejlik

- A fenti leírás kissé vázlatos. Matematikailag nem írjuk le pontosan mikor állunk le. A részletekhez sajnos nincs elég időnk.
- A fenti leírás már bizonyos programozási nyelvekben minden probléma nélkül kódolható.
- Elméletileg azonban több sebből vérzik.
- Sajnos az update formula négyzetgyökvonást is tartalmaz. Kerekítenünk kell. Ez numerikus kérdéseket vet fel.

Az ördög a részletekben rejlik

- A fenti leírás kissé vázlatos. Matematikailag nem írjuk le pontosan mikor állunk le. A részletekhez sajnos nincs elég időnk.
- A fenti leírás már bizonyos programozási nyelvekben minden probléma nélkül kódolható.
- Elméletileg azonban több sebből vérzik.
- Sajnos az update formula négyzetgyökvonást is tartalmaz. Kerekítenünk kell. Ez numerikus kérdéseket vet fel.
- Extrém esetben az ellipszoidot leíró mátrix pozitív definitását is elveszíthetjük pontatlan aritmetikával.

Az ördög a részletekben rejlik

- A fenti leírás kissé vázlatos. Matematikailag nem írjuk le pontosan mikor állunk le. A részletekhez sajnos nincs elég időnk.
- A fenti leírás már bizonyos programozási nyelvekben minden probléma nélkül kódolható.
- Elméletileg azonban több sebből vérzik.
- Sajnos az update formula négyzetgyökvonást is tartalmaz. Kerekítenünk kell. Ez numerikus kérdéseket vet fel.
- Extrém esetben az ellipszoidot leíró mátrix pozitív definitását is elveszíthetjük pontatlan aritmetikával.
- A numerikus matematikai korrekciók, az ehhez tartozó analízis meghaladja az előadás kereteit.

Az ördög a részletekben rejlik

- A fenti leírás kissé vázlatos. Matematikailag nem írjuk le pontosan mikor állunk le. A részletekhez sajnos nincs elég időnk.
- A fenti leírás már bizonyos programozási nyelvekben minden probléma nélkül kódolható.
- Elméletileg azonban több sebből vérzik.
- Sajnos az update formula négyzetgyökvonást is tartalmaz. Kerekítenünk kell. Ez numerikus kérdéseket vet fel.
- Extrém esetben az ellipszoidot leíró mátrix pozitív definitását is elveszíthetjük pontatlan aritmetikával.
- A numerikus matematikai korrekciók, az ehhez tartozó analízis meghaladja az előadás kereteit.
- A továbbiakban csak a számunkra fontos elméleti eredményeket összegezzük.

Elméleti eredmények

Elméleti eredmények

Tétel

Az ellipszoid módszer bizonyítja, hogy a LP-SAT nyelv $\mathcal{P}$-beli.

Elméleti eredmények

Tétel

Az ellipszoid módszer bizonyítja, hogy a LP-SAT nyelv $\mathcal{P}$-beli.

- Láttuk, hogy az ellipszoidunk térfogata mértani sorozatnak megfelelően csökken (ez a kerekítések mellett is megmarad).

Elméleti eredmények

Tétel

Az ellipszoid módszer bizonyítja, hogy a LP-SAT nyelv $\mathcal{P}$-beli.

- Láttuk, hogy az ellipszoidunk térfogata mértani sorozatnak megfelelően csökken (ez a kerekítések mellett is megmarad).
- Az ellipszoidunk térfogata egy felső becslés $\mathcal{L}$ térfogatára.

Elméleti eredmények

Tétel

Az ellipszoid módszer bizonyítja, hogy a LP-SAT nyelv $\mathcal{P}$-beli.

- Láttuk, hogy az ellipszoidunk térfogata mértani sorozatnak megfelelően csökken (ez a kerekítések mellett is megmarad).
- Az ellipszoidunk térfogata egy felső becslés $\mathcal{L}$ térfogatára.
- Ha ez kiinduló alsó becslésünk alá ér biztosak lehetünk abban, hogy $\mathcal{L} = \emptyset$.

Elméleti eredmények

Tétel

Az ellipszoid módszer bizonyítja, hogy a LP-SAT nyelv $\mathcal{P}$-beli.

- Láttuk, hogy az ellipszoidunk térfogata mértani sorozatnak megfelelően csökken (ez a kerekítések mellett is megmarad).
- Az ellipszoidunk térfogata egy felső becslés $\mathcal{L}$ térfogatára.
- Ha ez kiinduló alsó becslésünk alá ér biztosak lehetünk abban, hogy $\mathcal{L} = \emptyset$.
- Ebből kiszámolható lesz a leállási feltételben szereplő „elég nagy” kitétel pontos matematikai leírása.

Elméleti eredmények

Tétel

Az ellipszoid módszer bizonyítja, hogy a LP-SAT nyelv $\mathcal{P}$-beli.

• Láttuk, hogy az ellipszoidunk térfogata mértani sorozatnak megfelelően csökken (ez a kerekítések mellett is megmarad).

• Az ellipszoidunk térfogata egy felső becslés $\mathcal{L}$ térfogatára.

• Ha ez kiinduló alsó becslésünk alá ér biztosak lehetünk abban, hogy $\mathcal{L} = \emptyset$.

• Ebből kiszámolható lesz a leállási feltételben szereplő „elég nagy" kitétel pontos matematikai leírása.

• A kapott algoritmus polinomiális lesz. A részletek kidolgozása messze meghaladja a kurzus kereteit.

Elméleti eredmények II

Elméleti eredmények II

Tétel

Az ellipszoid módszer bizonyítja, hogy a LP feladatra ϵ szuboptimális megoldást találhatunk az input méretében és $\log(1/\epsilon)$-ban.

Elméleti eredmények II

Tétel

Az ellipszoid módszer bizonyítja, hogy a LP feladatra ϵ szuboptimális megoldást találhatunk az input méretében és $\log(1/\epsilon)$-ban.

- Ez következik abból, hogy az optimalizálási feladatot visszavezettük a kielégíthetőségi problémára.

Elméleti eredmények II

Tétel

Az ellipszoid módszer bizonyítja, hogy a LP feladatra ϵ szuboptimális megoldást találhatunk az input méretében és $\log(1/\epsilon)$-ban.

- Ez következik abból, hogy az optimalizálási feladatot visszavezettük a kielégíthetőségi problémára.
- Ennél több is mondható.

Elméleti eredmények II

Tétel

Az ellipszoid módszer bizonyítja, hogy a LP feladatra ϵ szuboptimális megoldást találhatunk az input méretében és $\log(1/\epsilon)$-ban.

- Ez következik abból, hogy az optimalizálási feladatot visszavezettük a kielégíthetőségi problémára.
- Ennél több is mondható.

Tétel

Az ellipszoid módszer bizonyítja, hogy az LP feladat egy optimális megoldását megtaláljuk $\mathcal{P}$-ben.

Elméleti eredmények II

Tétel

Az ellipszoid módszer bizonyítja, hogy a LP feladatra ϵ szuboptimális megoldást találhatunk az input méretében és $\log(1/\epsilon)$-ban.

- Ez következik abból, hogy az optimalizálási feladatot visszavezettük a kielégíthetőségi problémára.
- Ennél több is mondható.

Tétel

Az ellipszoid módszer bizonyítja, hogy az LP feladat egy optimális megoldását megtaláljuk $\mathcal{P}$-ben.

- A bizonyítás ismét technikai. Időnk nem engedi meg a részletek kidolgozását.

Elméleti eredmények III

Elméleti eredmények III

Tétel

Egy adott gráfban maximális súlyozott párosítás keresése lineáris programozási eszközökkel hatékonyan (polinomiális lépésben) megoldható.

Elméleti eredmények III

Tétel

Egy adott gráfban maximális súlyozott párosítás keresése lineáris programozási eszközökkel hatékonyan (polinomiális lépésben) megoldható.

- Az ellipszoid módszerrel lineáris célfüggvény polinom időben optimalizálható a párosítási politópon.

Elméleti eredmények III

Tétel

Egy adott gráfban maximális súlyozott párosítás keresése lineáris programozási eszközökkel hatékonyan (polinomiális lépésben) megoldható.

- Az ellipszoid módszerrel lineáris célfüggvény polinom időben optimalizálható a párosítási politópon.
- A maximális súlyozott párosítási probléma megfogalmazható így, a tétel állítása adódik.

Elméleti eredmények III

Tétel

Egy adott gráfban maximális súlyozott párosítás keresése lineáris programozási eszközökkel hatékonyan (polinomiális lépésben) megoldható.

- Az ellipszoid módszerrel lineáris célfüggvény polinom időben optimalizálható a párosítási politópon.
- A maximális súlyozott párosítási probléma megfogalmazható így, a tétel állítása adódik.
- A fentinél kombinatorikusabb, elemibb algoritmusok is léteznek. Azért megjegyezzük, hogy ezek is gyakran rendelkeznek LP interpretációval.

Elméleti eredmények IV

Elméleti eredmények IV

- Az ellipszoid módszer alkalmazható az SDP probléma megoldására is.

Elméleti eredmények IV

- Az ellipszoid módszer alkalmazható az SDP probléma megoldására is.
- Az SDP feladat jóval bonyolultabban néz ki mint az általános LP probléma.

Elméleti eredmények IV

- Az ellipszoid módszer alkalmazható az SDP probléma megoldására is.

- Az SDP feladat jóval bonyolultabban néz ki mint az általános LP probléma.

- Eddig véges sok lineáris egyenlőtlenség írta le a feltételrendszert. Most egy pozitív szemidefinitás is megjelenik:

Minimalizáljuk	$c^T x$-t
Feltéve, hogy	$Ax = b$
	$x_1 C_1 + x_2 C_2 + \ldots + x_n C_n \succeq D,$

ahol $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{k \times n}$, a C_i, D mátrixok szimmetrikus ($\ell \times \ell$ méretű) mátrixok.

Elméleti eredmények V

Elméleti eredmények V

Tétel

Tegyük fel, hogy adott egy SDP feladat,a mely I inputjában racionális számok állnak. Legyen $\|I\|$ az input leírásához szükséges bitek száma. Adott továbbá egy tetszőleges $\epsilon > 0$ paraméter. Ekkor létezik olyan algoritmus, amely egy ϵ-szuboptimális megoldást talál és polinomiális $\|I\|$-ben és $\log(1/\epsilon)$-ban.

Elméleti eredmények V

Tétel

Tegyük fel, hogy adott egy SDP feladat,a mely I inputjában racionális számok állnak. Legyen $\|I\|$ az input leírásához szükséges bitek száma. Adott továbbá egy tetszőleges $\epsilon > 0$ paraméter. Ekkor létezik olyan algoritmus, amely egy ϵ-szuboptimális megoldást talál és polinomiális $\|I\|$-ben és $\log(1/\epsilon)$-ban.

- LP esetén ragaszkodhattunk az optimális hely és optimális érték kereséséhez. Itt az input racionalitását kódolási okok miatt célszerű feltenni. Az optimális hely/érték azonban lehet, hogy irracionális.

Elméleti eredmények V

Tétel

Tegyük fel, hogy adott egy SDP feladat,a mely I inputjában racionális számok állnak. Legyen $\|I\|$ az input leírásához szükséges bitek száma. Adott továbbá egy tetszőleges $\epsilon > 0$ paraméter. Ekkor létezik olyan algoritmus, amely egy ϵ-szuboptimális megoldást talál és polinomiális $\|I\|$-ben és $\log(1/\epsilon)$-ban.

- LP esetén ragaszkodhattunk az optimális hely és optimális érték kereséséhez. Itt az input racionalitását kódolási okok miatt célszerű feltenni. Az optimális hely/érték azonban lehet, hogy irracionális.

- A feltételrendszer egy konvex halmazt definiál. Az ehhez tartozást egy pozitív szemidefinitás teszt ellenőrzi.

Elméleti eredmények V

Tétel

Tegyük fel, hogy adott egy SDP feladat,a mely I inputjában racionális számok állnak. Legyen $\|I\|$ az input leírásához szükséges bitek száma. Adott továbbá egy tetszőleges $\epsilon > 0$ paraméter. Ekkor létezik olyan algoritmus, amely egy ϵ-szuboptimális megoldást talál és polinomiális $\|I\|$-ben és $\log(1/\epsilon)$-ban.

• LP esetén ragaszkodhattunk az optimális hely és optimális érték kereséséhez. Itt az input racionalitását kódolási okok miatt célszerű feltenni. Az optimális hely/érték azonban lehet, hogy irracionális.

• A feltételrendszer egy konvex halmazt definiál. Az ehhez tartozást egy pozitív szemidefinitás teszt ellenőrzi. Ilyen eljárás alap lineáris algebrai ismeretek alapján könnyen megadható.

Elméleti eredmények V

Tétel

Tegyük fel, hogy adott egy SDP feladat,a mely I inputjában racionális számok állnak. Legyen $\|I\|$ az input leírásához szükséges bitek száma. Adott továbbá egy tetszőleges $\epsilon > 0$ paraméter. Ekkor létezik olyan algoritmus, amely egy ϵ-szuboptimális megoldást talál és polinomiális $\|I\|$-ben és $\log(1/\epsilon)$-ban.

• LP esetén ragaszkodhattunk az optimális hely és optimális érték kereséséhez. Itt az input racionalitását kódolási okok miatt célszerű feltenni. Az optimális hely/érték azonban lehet, hogy irracionális.

• A feltételrendszer egy konvex halmazt definiál. Az ehhez tartozást egy pozitív szemidefinitás teszt ellenőrzi. Ilyen eljárás alap lineáris algebrai ismeretek alapján könnyen megadható. Ha ügyesek vagyunk a pozitív szemidefinitás sérülése esetén egy szeparáló félsíkot is kiolvashatunk tesztünkből.

Szünet

Numerikus módszerek

Numerikus módszerek

Numerikus módszerek vannak feltétel nélküli optimalizálásra (esetleg gyökkeresési nyelvezettel elmondva).

Numerikus módszerek

Numerikus módszerek vannak feltétel nélküli optimalizálásra (esetleg gyökkeresési nyelvezettel elmondva).

$$\boxed{\text{Minimalizáljuk} \qquad c(x)\text{-t}}$$

Numerikus módszerek

Numerikus módszerek vannak feltétel nélküli optimalizálásra (esetleg gyökkeresési nyelvezettel elmondva).

$$\boxed{\text{Minimalizáljuk} \qquad c(x)\text{-t}}$$

Például

Numerikus módszerek

Numerikus módszerek vannak feltétel nélküli optimalizálásra (esetleg gyökkeresési nyelvezettel elmondva).

$$\boxed{\text{Minimalizáljuk} \qquad c(x)\text{-t}}$$

Például

- Gradiens módszer,

Numerikus módszerek

Numerikus módszerek vannak feltétel nélküli optimalizálásra (esetleg gyökkeresési nyelvezettel elmondva).

$$\boxed{\text{Minimalizáljuk} \qquad c(x)\text{-t}}$$

Például

- Gradiens módszer,
- Newton módszer.

Numerikus módszerek

Numerikus módszerek vannak feltétel nélküli optimalizálásra (esetleg gyökkeresési nyelvezettel elmondva).

$$\boxed{\text{Minimalizáljuk} \qquad c(x)\text{-t}}$$

Például

- Gradiens módszer,
- Newton módszer.

A numerikus módszerek lineáris egyenlőség feltételek esetére kiterjeszthetők.

Numerikus módszerek

Numerikus módszerek vannak feltétel nélküli optimalizálásra (esetleg gyökkeresési nyelvezettel elmondva).

Minimalizáljuk	$c(x)$-t

Például

- Gradiens módszer,
- Newton módszer.

A numerikus módszerek lineáris egyenlőség feltételek esetére kiterjeszthetők.

Minimalizáljuk	$c(x)$-t
Feltéve, hogy	$Ax = b$

Numerikus módszerek

Numerikus módszerek vannak feltétel nélküli optimalizálásra (esetleg gyökkeresési nyelvezettel elmondva).

Minimalizáljuk	$c(x)$-t

Például

- Gradiens módszer,
- Newton módszer.

A numerikus módszerek lineáris egyenlőség feltételek esetére kiterjeszthetők.

Minimalizáljuk	$c(x)$-t
Feltéve, hogy	$Ax = b$

Az igazi probléma az egyenlőtlenség feltételek kezelése.

Belső pontos módszerek: az ötlet

Belső pontos módszerek: az ötlet

- A numerikus módszerek feltezik, hogy a célfüggvény teljes értelmezési tartományán optimalizáltunk.

Belső pontos módszerek: az ötlet

- A numerikus módszerek feltezik, hogy a célfüggvény teljes értelmezési tartományán optimalizáltunk.
- A módszerek alkalmazása feltételek melletti optimalizálásra nem nyilvánvaló.

Belső pontos módszerek: az ötlet

- A numerikus módszerek feltezik, hogy a célfüggvény teljes értelmezési tartományán optimalizáltunk.
- A módszerek alkalmazása feltételek melletti optimalizálásra nem nyilvánvaló.
- **Egy ötlet:** A c célfüggvényt lecseréljük egy (paraméteres) c_p segédfüggvényre, amelyben az egyenlőtlenség feltételek bizonyos súllyal szerepelnek.

Belső pontos módszerek: az ötlet

• A numerikus módszerek feltezik, hogy a célfüggvény teljes értelmezési tartományán optimalizáltunk.
• A módszerek alkalmazása feltételek melletti optimalizálásra nem nyilvánvaló.
• **Egy ötlet:** A c célfüggvényt lecseréljük egy (paraméteres) c_p segédfüggvényre, amelyben az egyenlőtlenség feltételek bizonyos súllyal szerepelnek.

(a) a feltételnek megfelelő tartomány határozott belsejében közel egyenlő a célfüggvénnyel,
(b) a tartomány határához közel nagyon nagy értékeket vesz fel,
(c) továbbra is konvex,
(d) többszörösen differenciálható (a tartomány belsején kívül nem is lesz értelmezett vagy végtelenként értelmezzük).

Belső pontos módszerek: az ötlet

- A numerikus módszerek feltezik, hogy a célfüggvény teljes értelmezési tartományán optimalizáltunk.
- A módszerek alkalmazása feltételek melletti optimalizálásra nem nyilvánvaló.
- **Egy ötlet:** A c célfüggvényt lecseréljük egy (paraméteres) c_p segédfüggvényre, amelyben az egyenlőtlenség feltételek bizonyos súllyal szerepelnek.

(a) a feltételnek megfelelő tartomány határozott belsejében közel egyenlő a célfüggvénnyel,

(b) a tartomány határához közel nagyon nagy értékeket vesz fel,

(c) továbbra is konvex,

(d) többszörösen differenciálható (a tartomány belsején kívül nem is lesz értelmezett vagy végtelenként értelmezzük).

- Az egyenlőtlenség feltételeket eldobjuk.

Belső pontos módszerek: az ötlet (folytatás)

Belső pontos módszerek: az ötlet (folytatás)

- A paraméter értékét növelve az a tartomány, ahol a közelítő segédfüggvény jól approximálja a célfüggvényt egyre jobban a feltételek által leírt tartományhoz „simul”.

Belső pontos módszerek: az ötlet (folytatás)

- A paraméter értékét növelve az a tartomány, ahol a közelítő segédfüggvény jól approximálja a célfüggvényt egyre jobban a feltételek által leírt tartományhoz „simul”.
- A segédfüggvényre alkalmazva a most leírt módszereket egy $a = a_p$ aktuális értéket kapunk.

Belső pontos módszerek: az ötlet (folytatás)

- A paraméter értékét növelve az a tartomány, ahol a közelítő segédfüggvény jól approximálja a célfüggvényt egyre jobban a feltételek által leírt tartományhoz „simul”.
- A segédfüggvényre alkalmazva a most leírt módszereket egy $a = a_p$ aktuális értéket kapunk.
- A paraméter növelésével kapott jobb segédfüggvényt véve $a = a_p$-ból megkapjuk az update-lt $a^+ = a_{p^+}$ pontot.

Belső pontos módszerek: az ötlet (folytatás)

- A paraméter értékét növelve az a tartomány, ahol a közelítő segédfüggvény jól approximálja a célfüggvényt egyre jobban a feltételek által leírt tartományhoz „simul”.
- A segédfüggvényre alkalmazva a most leírt módszereket egy $a = a_p$ aktuális értéket kapunk.
- A paraméter növelésével kapott jobb segédfüggvényt véve $a = a_p$-ból megkapjuk az update-lt $a^+ = a_{p^+}$ pontot.
- Az aktuális pontok végig a lehetséges megoldások halmazának belsejéből kerülnek ki.

Belső pontos módszerek: az ötlet (folytatás)

- A paraméter értékét növelve az a tartomány, ahol a közelítő segédfüggvény jól approximálja a célfüggvényt egyre jobban a feltételek által leírt tartományhoz „simul”.

- A segédfüggvényre alkalmazva a most leírt módszereket egy $a = a_p$ aktuális értéket kapunk.

- A paraméter növelésével kapott jobb segédfüggvényt véve $a = a_p$-ból megkapjuk az update-lt $a^+ = a_{p^+}$ pontot.

- Az aktuális pontok végig a lehetséges megoldások halmazának belsejéből kerülnek ki. Szemben például a szimplex módszerrel, amely csúcspontokon (a határ speciális pontjain) keresztül közelít az optimumhoz.

Belső pontos módszerek: az ötlet (folytatás)

• A paraméter értékét növelve az a tartomány, ahol a közelítő segédfüggvény jól approximálja a célfüggvényt egyre jobban a feltételek által leírt tartományhoz „simul”.

• A segédfüggvényre alkalmazva a most leírt módszereket egy $a = a_p$ aktuális értéket kapunk.

• A paraméter növelésével kapott jobb segédfüggvényt véve $a = a_p$-ból megkapjuk az update-lt $a^+ = a_{p^+}$ pontot.

• Az aktuális pontok végig a lehetséges megoldások halmazának belsejéből kerülnek ki. Szemben például a szimplex módszerrel, amely csúcspontokon (a határ speciális pontjain) keresztül közelít az optimumhoz. Az ilyen módszereket belső pontos módszernek nevezik.

A kiinduló feladat

A kiinduló feladat

Legyen $\mathcal{O}$ az alábbi optimalizálási feladat:

$$\mathcal{O}: \quad \boxed{\begin{array}{lll} \text{Minimalizáljuk} & c(x)\text{-et} & \\ \text{Feltéve, hogy} & Ax = b, & \\ & f_i(x) \leq 0, & i = 1, 2, \ldots, k, \end{array}}$$

ahol $A \in \mathbb{R}^{\ell \times n}$, $b \in \mathbb{R}^{\ell}$, $x \in \mathbb{R}^n$.

A kiinduló feladat

Legyen $\mathcal{O}$ az alábbi optimalizálási feladat:

$$\mathcal{O}: \quad \boxed{\begin{array}{lll} \text{Minimalizáljuk} & c(x)\text{-et} & \\ \text{Feltéve, hogy} & Ax = b, & \\ & f_i(x) \leq 0, & i = 1, 2, \ldots, k, \end{array}}$$

ahol $A \in \mathbb{R}^{\ell \times n}$, $b \in \mathbb{R}^{\ell}$, $x \in \mathbb{R}^{n}$.

Az eddigiektől új, nehezítő egyenlőtlenség feltételeket a célfüggvénybe olvasztjuk.

A feltételek beolvasztása a célfüggvénybe, a csalás

A feltételek beolvasztása a célfüggvénybe, a csalás

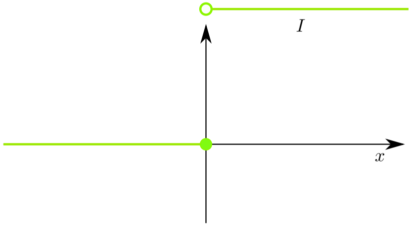
Definíció

Legyen $I(x) : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$, ami nem pozitív értékeken 0-t, pozitív értékeken ∞ értéket vesz fel.

A feltételek beolvasztása a célfüggvénybe, a csalás

Definíció

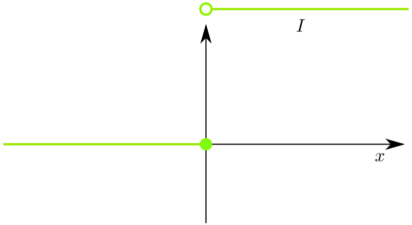
Legyen $I(x) : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$, ami nem pozitív értékeken 0-t, pozitív értékeken ∞ értéket vesz fel.



A feltételek beolvasztása a célfüggvénybe, a csalás

Definíció

Legyen $I(x) : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$, ami nem pozitív értékeken 0-t, pozitív értékeken ∞ értéket vesz fel.



Az eredeti $\mathcal{O}$ problémát ekvivalens módon megfogalmazhatjuk az I függvény segítségével:

Minimalizáljuk	$c(x) + \sum_{i=1}^{k} I(f_i(x))$-t
Feltéve, hogy	$Ax = b$.

A feltételek beolvasztása a célfüggvénybe, a csalás kiváltása

A feltételek beolvasztása a célfüggvénybe, a csalás kiváltása

- Természetesen a csalás problémája, hogy általában szép c, f_i függvényekkel találkozunk.

A feltételek beolvasztása a célfüggvénybe, a csalás kiváltása

- Természetesen a csalás problémája, hogy általában szép c, f_i függvényekkel találkozunk.
- A korábbi módszereink differenciálhatósági feltételek mellett működnek.

A feltételek beolvasztása a célfüggvénybe, a csalás kiváltása

- Természetesen a csalás problémája, hogy általában szép c, f_i függvényekkel találkozunk.
- A korábbi módszereink differenciálhatósági feltételek mellett működnek.
- A bevezetett függvény nem differenciálható. A kiút, hogy az I függvényt differenciálható függvénnyel közelítjük.

A feltételek beolvasztása a célfüggvénybe, a csalás kiváltása

- Természetesen a csalás problémája, hogy általában szép c, f_i függvényekkel találkozunk.
- A korábbi módszereink differenciálhatósági feltételek mellett működnek.
- A bevezetett függvény nem differenciálható. A kiút, hogy az I függvényt differenciálható függvénnyel közelítjük.
- Az I függvény olyan x-eket enged a „c-t minimalizáló versenybe”, amelyeknél az f_i-k előjele nem pozitív.

A feltételek beolvasztása a célfüggvénybe, a csalás kiváltása

- Természetesen a csalás problémája, hogy általában szép c, f_i függvényekkel találkozunk.
- A korábbi módszereink differenciálhatósági feltételek mellett működnek.
- A bevezetett függvény nem differenciálható. A kiút, hogy az I függvényt differenciálható függvénnyel közelítjük.
- Az I függvény olyan x-eket enged a „c-t minimalizáló versenybe", amelyeknél az f_i-k előjele nem pozitív.
- Az I függvény *gátat szab* a „versenyzőknek". Az ilyen függvényeket barrier/gátfüggvényeknek nevezzük.

A feltételek beolvasztása a célfüggvénybe, a csalás kiváltása

- Természetesen a csalás problémája, hogy általában szép c, f_i függvényekkel találkozunk.
- A korábbi módszereink differenciálhatósági feltételek mellett működnek.
- A bevezetett függvény nem differenciálható. A kiút, hogy az I függvényt differenciálható függvénnyel közelítjük.
- Az I függvény olyan x-eket enged a „c-t minimalizáló versenybe”, amelyeknél az f_i-k előjele nem pozitív.
- Az I függvény *gátat szab* a „versenyzőknek”. Az ilyen függvényeket barrier/gátfüggvényeknek nevezzük.
- I közelítése differenciálható függvénnyel, ami a gátfüggvényység tulajdoságot „szimulálja” sokféleképpen megoldható.

A feltételek beolvasztása a célfüggvénybe, a csalás kiváltása

- Természetesen a csalás problémája, hogy általában szép c, f_i függvényekkel találkozunk.
- A korábbi módszereink differenciálhatósági feltételek mellett működnek.
- A bevezetett függvény nem differenciálható. A kiút, hogy az I függvényt differenciálható függvénnyel közelítjük.
- Az I függvény olyan x-eket enged a „c-t minimalizáló versenybe”, amelyeknél az f_i-k előjele nem pozitív.
- Az I függvény *gátat szab* a „versenyzőknek”. Az ilyen függvényeket barrier/gátfüggvényeknek nevezzük.
- I közelítése differenciálható függvénnyel, ami a gátfüggvénység tulajdoságot „szimulálja” sokféleképpen megoldható. Mi egy lehetőségét emelünk ki.

A kiváltás: logaritmikus gátfüggvény, log barrier

A kiváltás: logaritmikus gátfüggvény, log barrier

Logaritmikus gátfüggvény

Legyen $\widetilde{I}_t(x) = -\frac{1}{t}\log(-x)$, ahol $t > 0$.

A kiváltás: logaritmikus gátfüggvény, log barrier

Logaritmikus gátfüggvény

Legyen $\widetilde{I}_t(x) = -\frac{1}{t}\log(-x)$, ahol $t > 0$.

- Minél nagyobb a t értéke, annál jobban közelíti a szükséges egyenlőtlenség feltételt szimuláló függvényt.

A kiváltás: logaritmikus gátfüggvény, log barrier

Logaritmikus gátfüggvény

Legyen $\widetilde{I}_t(x) = -\frac{1}{t}\log(-x)$, ahol $t > 0$.

• Minél nagyobb a t értéke, annál jobban közelíti a szükséges egyenlőtlenség feltételt szimuláló függvényt.
• $\widetilde{I}_t(x)$ logaritmikus gátfüggvények grafikonjai különböző t értékek esetén.

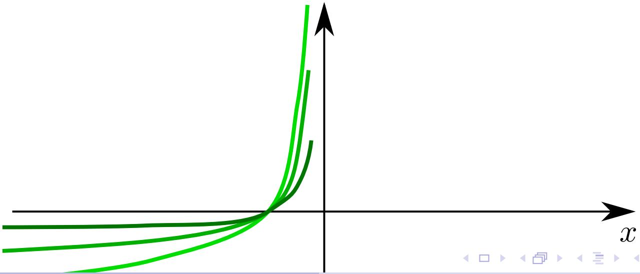
A kiváltás: logaritmikus gátfüggvény, log barrier

Logaritmikus gátfüggvény

Legyen $\widetilde{I}_t(x) = -\frac{1}{t}\log(-x)$, ahol $t > 0$.

• Minél nagyobb a t értéke, annál jobban közelíti a szükséges egyenlőtlenség feltételt szimuláló függvényt.

• $\widetilde{I}_t(x)$ logaritmikus gátfüggvények grafikonjai különböző t értékek esetén. A sötétebb grafikon nagyobb t értékhez tartozik, jobban közelíti a szükséges egyenlőtlenség feltételt szimuláló függvényt.



A kiváltás: az új célfüggvény

A kiváltás: az új célfüggvény

- Fixáljunk egy t értéket és a csaló célfüggvény helyett a logaritmikus gátfüggvénnyel küszöböljük ki az egyenlőtlenség feltételeket: $c_t(c) = c(x) + \sum_{i=1}^{k} \widetilde{I}_t(f_i(x))$.

A kiváltás: az új célfüggvény

- Fixáljunk egy t értéket és a csaló célfüggvény helyett a logaritmikus gátfüggvénnyel küszöböljük ki az egyenlőtlenség feltételeket: $c_t(c) = c(x) + \sum_{i=1}^{k} \widetilde{I}_t(f_i(x))$.
- Az optimum helye nem változik, ha a célfüggvényt a fix t-vel megszorozzuk:

$$\widetilde{c}_t(x) = tc(x) + \sum_{i=1}^{k} t\widetilde{I}_t(f(x)).$$

A kiváltás: az új célfüggvény

- Fixáljunk egy t értéket és a csaló célfüggvény helyett a logaritmikus gátfüggvénnyel küszöböljük ki az egyenlőtlenség feltételeket: $c_t(c) = c(x) + \sum_{i=1}^{k} \widetilde{I}_t(f_i(x))$.
- Az optimum helye nem változik, ha a célfüggvényt a fix t-vel megszorozzuk:

$$\widetilde{c}_t(x) = tc(x) + \sum_{i=1}^{k} t\widetilde{I}_t(f(x)).$$

Jelölés

$$\Phi_{\mathcal{F}}(x) = \Phi(x) = -\sum_{i=1}^{k} \log(-f_i(x)).$$

A kiváltás: az új célfüggvény

- Fixáljunk egy t értéket és a csaló célfüggvény helyett a logaritmikus gátfüggvénnyel küszöböljük ki az egyenlőtlenség feltételeket: $c_t(c) = c(x) + \sum_{i=1}^{k} \widetilde{I}_t(f_i(x))$.
- Az optimum helye nem változik, ha a célfüggvényt a fix t-vel megszorozzuk:

$$\widetilde{c}_t(x) = tc(x) + \sum_{i=1}^{k} t\widetilde{I}_t(f(x)).$$

Jelölés

$$\Phi_{\mathcal{F}}(x) = \Phi(x) = -\sum_{i=1}^{k} \log(-f_i(x)).$$

- A ϕ függvény a feltételrendszerünktől függ, igazából csak az egyenlőtlenségek rendszerétől.

1. eset: $\mathcal{L}$ poliéder, az új célfüggvény

1. eset: $\mathcal{L}$ poliéder, az új célfüggvény

- Legyen $\mathcal{F} : Ax \preceq b$ ($A \in \mathbb{R}^{k \times n}$, $b \in \mathbb{R}^k$).

1. eset: $\mathcal{L}$ poliéder, az új célfüggvény

- Legyen $\mathcal{F} : Ax \preceq b$ ($A \in \mathbb{R}^{k \times n}$, $b \in \mathbb{R}^k$).
- Ekkor

$$\Phi(x) = -\sum_{i=1}^{k} \log(b_i - a_i^T x),$$

ahol $b = (b_1, b_2, \ldots, b_k)^T$ és a_i^T az A mátrix i-edik sora.

1. eset: $\mathcal{L}$ poliéder, az új célfüggvény

- Legyen $\mathcal{F} : Ax \preceq b$ ($A \in \mathbb{R}^{k\times n}$, $b \in \mathbb{R}^k$).
- Ekkor

$$\Phi(x) = -\sum_{i=1}^{k} \log(b_i - a_i^T x),$$

ahol $b = (b_1, b_2, \ldots, b_k)^T$ és a_i^T az A mátrix i-edik sora.

- $\phi(x)$ „szép” függvény, könnyű vele számolni:

$$\nabla\Phi(x) = \sum_{i=1}^{k} \frac{1}{b_i - a_i^T x} \cdot a_i,$$

$$\nabla^2\Phi(x) = \sum_{i=1}^{k} \frac{1}{(b_i - a_i^T x)^2} a_i a_i^T.$$

Az új célfüggvény, a centrális út

Az új célfüggvény, a centrális út

Legyen $\widetilde{\mathcal{O}}_t$ a következő optimalizálási feladat, ahol $t > 0$ fix szám:

$$\widetilde{\mathcal{O}}_t : \quad \boxed{\begin{array}{ll} \text{Minimalizáljuk} & t \cdot c(x) + \Phi(x)\text{-et} \\ \text{Feltéve, hogy} & Ax = b. \end{array}}$$

Az új célfüggvény, a centrális út

Legyen $\widetilde{\mathcal{O}}_t$ a következő optimalizálási feladat, ahol $t > 0$ fix szám:

$$\widetilde{\mathcal{O}}_t : \quad \boxed{\begin{array}{ll} \text{Minimalizáljuk} & t \cdot c(x) + \Phi(x)\text{-et} \\ \text{Feltéve, hogy} & Ax = b. \end{array}}$$

Definíció

Legyen $x^*(t)$ az $\widetilde{\mathcal{O}}_t$ optimalizálási feladat optimális helye.

Az új célfüggvény, a centrális út

Legyen $\widetilde{\mathcal{O}}_t$ a következő optimalizálási feladat, ahol $t > 0$ fix szám:

$$\widetilde{\mathcal{O}}_t : \quad \boxed{\begin{array}{ll} \text{Minimalizáljuk} & t \cdot c(x) + \Phi(x)\text{-et} \\ \text{Feltéve, hogy} & Ax = b. \end{array}}$$

Definíció

Legyen $x^*(t)$ az $\widetilde{\mathcal{O}}_t$ optimalizálási feladat optimális helye.

Az $x^*(t)$ pontok az optimalizálási feladat *centrális pontjai*.

Az új célfüggvény, a centrális út

Legyen $\widetilde{\mathcal{O}}_t$ a következő optimalizálási feladat, ahol $t > 0$ fix szám:

$$\widetilde{\mathcal{O}}_t : \quad \boxed{\begin{array}{ll} \text{Minimalizáljuk} & t \cdot c(x) + \Phi(x)\text{-et} \\ \text{Feltéve, hogy} & Ax = b. \end{array}}$$

Definíció

Legyen $x^*(t)$ az $\widetilde{\mathcal{O}}_t$ optimalizálási feladat optimális helye.

Az $x^*(t)$ pontok az optimalizálási feladat *centrális pontjai*.

Ha t végig fut a $\mathbb{R}_{>0}$ halmazon, akkor az $x^*(t)$ helyek a *centrális utat* írják le.

Az új céfüggvény, a centrális út és KKT

Az új céfüggvény, a centrális út és KKT

KKT-tétel alapján könnyű karakterizálni $x^*(t)$ helyeket:

Az új céfüggvény, a centrális út és KKT

KKT-tétel alapján könnyű karakterizálni $x^*(t)$ helyeket:

Lemma

$x^*(t)$-t karakterizálják az alábbi feltételek:

(i) $Ax^*(t) = b$,

(ii) $f_i(x^*(t)) < 0$ minden $i = 1, 2, \ldots, k$ esetén,

(iii) alkalmas $\mu \in \mathbb{R}^\ell$ esetén

$$t\nabla c(x^*(t)) + \nabla\Phi(x^*(t)) + A^T\mu = 0.$$

Az új céfüggvény, a centrális út és KKT

KKT-tétel alapján könnyű karakterizálni $x^*(t)$ helyeket:

Lemma

$x^*(t)$-t karakterizálják az alábbi feltételek:

(i) $Ax^*(t) = b$,

(ii) $f_i(x^*(t)) < 0$ minden $i = 1, 2, \ldots, k$ esetén,

(iii) alkalmas $\mu \in \mathbb{R}^\ell$ esetén

$$t\nabla c(x^*(t)) + \nabla\Phi(x^*(t)) + A^T\mu = 0.$$

A lemma bizonyítása egyből adódik a KKT-tételből.

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény

- Ismét legyen $\mathcal{F} : Ax \preceq b$, azaz a feltételrendszert kielégítő x-ek halmaza egy $\mathcal{P}$ politóp.

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény

- Ismét legyen $\mathcal{F} : Ax \preceq b$, azaz a feltételrendszert kielégítő x-ek halmaza egy $\mathcal{P}$ politóp. Továbbá legyen $c(x) = c^T \cdot x$ egy lineáris célfüggvény.

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény

- Ismét legyen $\mathcal{F} : Ax \preceq b$, azaz a feltételrendszert kielégítő x-ek halmaza egy $\mathcal{P}$ politóp. Továbbá legyen $c(x) = c^T \cdot x$ egy lineáris célfüggvény.
- Vegyük a $\Phi(x)$ (lásd előző példa) logaritmikus gátfüggvény $S_{\leq\alpha} = \{x : \Phi(x) \leq \alpha\}$ szub-szinthalmazait.

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény

- Ismét legyen $\mathcal{F} : Ax \preceq b$, azaz a feltételrendszert kielégítő x-ek halmaza egy $\mathcal{P}$ politóp. Továbbá legyen $c(x) = c^T \cdot x$ egy lineáris célfüggvény.
- Vegyük a $\Phi(x)$ (lásd előző példa) logaritmikus gátfüggvény $S_{\leq\alpha} = \{x : \Phi(x) \leq \alpha\}$ szub-szinthalmazait.
- Ezek egy növekvő halmazrendszer ($\alpha < \beta$ esetén $S_\alpha \subset S_\beta$), uniójuk kiadja $\mathcal{P}$ politóp belsejét.

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény

- Ismét legyen $\mathcal{F} : Ax \preceq b$, azaz a feltételrendszert kielégítő x-ek halmaza egy $\mathcal{P}$ politóp. Továbbá legyen $c(x) = c^T \cdot x$ egy lineáris célfüggvény.
- Vegyük a $\Phi(x)$ (lásd előző példa) logaritmikus gátfüggvény $S_{\leq\alpha} = \{x : \Phi(x) \leq \alpha\}$ szub-szinthalmazait.
- Ezek egy növekvő halmazrendszer ($\alpha < \beta$ esetén $S_\alpha \subset S_\beta$), uniójuk kiadja $\mathcal{P}$ politóp belsejét.
- Ahogy α nő S_α hozzásímul a $\mathcal{P}$ politóphoz.

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény

- Ismét legyen $\mathcal{F} : Ax \preceq b$, azaz a feltételrendszert kielégítő x-ek halmaza egy $\mathcal{P}$ politóp. Továbbá legyen $c(x) = c^T \cdot x$ egy lineáris célfüggvény.
- Vegyük a $\Phi(x)$ (lásd előző példa) logaritmikus gátfüggvény $S_{\leq\alpha} = \{x : \Phi(x) \leq \alpha\}$ szub-szinthalmazait.
- Ezek egy növekvő halmazrendszer ($\alpha < \beta$ esetén $S_\alpha \subset S_\beta$), uniójuk kiadja $\mathcal{P}$ politóp belsejét.
- Ahogy α nő S_α hozzásímul a $\mathcal{P}$ politóphoz.
- Másképpen $S_{=\alpha} = \{x : \Phi(x) = \alpha\}$ szinthalmazok $\mathcal{P}$ határához símulnak.

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény

- Ismét legyen $\mathcal{F} : Ax \preceq b$, azaz a feltételrendszert kielégítő x-ek halmaza egy $\mathcal{P}$ politóp. Továbbá legyen $c(x) = c^T \cdot x$ egy lineáris célfüggvény.
- Vegyük a $\Phi(x)$ (lásd előző példa) logaritmikus gátfüggvény $S_{\leq\alpha} = \{x : \Phi(x) \leq \alpha\}$ szub-szinthalmazait.
- Ezek egy növekvő halmazrendszer ($\alpha < \beta$ esetén $S_\alpha \subset S_\beta$), uniójuk kiadja $\mathcal{P}$ politóp belsejét.
- Ahogy α nő S_α hozzásímul a $\mathcal{P}$ politóphoz.
- Másképpen $S_{=\alpha} = \{x : \Phi(x) = \alpha\}$ szinthalmazok $\mathcal{P}$ határához símulnak.
- Az (ii) feltétel azt monjda $x^*(t)$-ről, hogy Φ értelmezési tartományába esik.

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény

- Ismét legyen $\mathcal{F} : Ax \preceq b$, azaz a feltételrendszert kielégítő x-ek halmaza egy $\mathcal{P}$ politóp. Továbbá legyen $c(x) = c^T \cdot x$ egy lineáris célfüggvény.

- Vegyük a $\Phi(x)$ (lásd előző példa) logaritmikus gátfüggvény $S_{\leq\alpha} = \{x : \Phi(x) \leq \alpha\}$ szub-szinthalmazait.

- Ezek egy növekvő halmazrendszer ($\alpha < \beta$ esetén $S_\alpha \subset S_\beta$), uniójuk kiadja $\mathcal{P}$ politóp belsejét.

- Ahogy α nő S_α hozzásímul a $\mathcal{P}$ politóphoz.

- Másképpen $S_{=\alpha} = \{x : \Phi(x) = \alpha\}$ szinthalmazok $\mathcal{P}$ határához símulnak.

- Az (ii) feltétel azt monjda $x^*(t)$-ről, hogy Φ értelmezési tartományába esik. Azaz $\Phi(x^*(t)) = \alpha$ esetén (azaz alkalmas α esetén) $x^*(t)$ a $S_{=\alpha}$ szinthalmaz egy eleme.

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény

- Az (iii) feltétel egyszerűsödik, hiszen $\nabla c(x) = c \in \mathbb{R}^n$.

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény

- Az (iii) feltétel egyszerűsödik, hiszen $\nabla c(x) = c \in \mathbb{R}^n$.
- μ hiányzik, hiszen nincs egyenlőség feltételünk.

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény

- Az (iii) feltétel egyszerűsödik, hiszen $\nabla c(x) = c \in \mathbb{R}^n$.
- μ hiányzik, hiszen nincs egyenlőség feltételünk.
- $\nabla\Phi$-t az előző példában kiszámoltuk:

$$tc + \nabla\Phi(x^*(t)) =$$

$$tc + A^T \mathrm{diag}\left(\frac{1}{b_1 - a_1^T x^*(t)}, \frac{1}{b_2 - a_2^T x^*(t)}, \dots, \frac{1}{b_k - a_k^T x^*(t)}\right) = 0.$$

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény

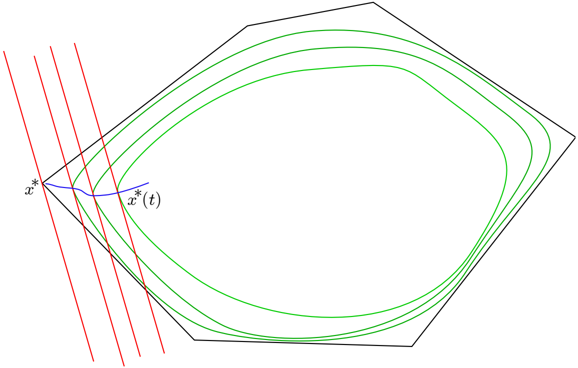


Figure: A zöld görbék a szinthalmazok, a kék görbe a centrális görbe, a fekete görbe a $\mathcal{P}$ politóp határa. A piros egyenesek c normálvektorú hipersíkok.

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény

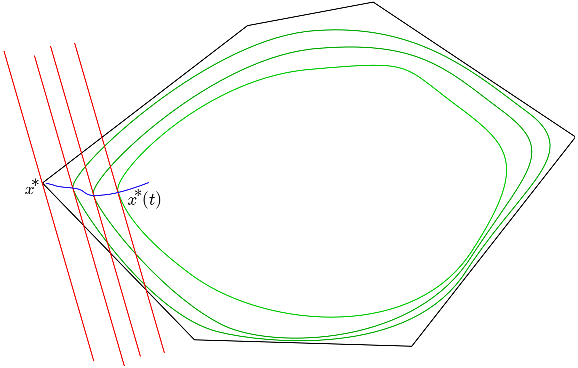


Figure: A zöld görbék a szinthalmazok, a kék görbe a centrális görbe, a fekete görbe a $\mathcal{P}$ politóp határa. A piros egyenesek c normálvektorú hipersíkok.

- Azaz $\nabla\Phi(x^*(t))$ párhuzamos c-vel. Azaz a $c^T x = c^T x^*(t)$ hipersík az $S_{=\alpha}$ szinhalmaz egy érintője a centrális út összes pontjára.

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény, a TÉTEL

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény, a TÉTEL

Tétel

$x^*(t)$ legyen egy centrális pont, azaz az $\widetilde{\mathcal{O}}_t$ feladat egy optimális megoldása (optimalizálás csak lineáris egyenlőség feltételekkel). A megoldáshoz vezető úton egy $w^*(t)$ duális optimumhelyet is megkapunk. Legyen

$$\lambda_i^*(t) = -\frac{1}{tf_i(x^*(t))}, \qquad \mu^*(t) = \frac{w^*(t)}{t}.$$

Ekkor

(i) $\lambda_i^*(t)$ és $\mu^*(t)$ duális megengedett megoldása az eredeti $\mathcal{O}$ feladatnak.

(ii) Továbbá a duális hézag $x^*(t)$ primál megengedett megoldás és $\lambda_i^*(t)$ és $\mu^*(t)$ duál megengedett megoldás között $\frac{\ell}{t}$.

1. eset: $\mathcal{L}$ poliéder, lineáris célfüggvény, a TÉTEL

Tétel

$x^*(t)$ legyen egy centrális pont, azaz az $\widetilde{\mathcal{O}}_t$ feladat egy optimális megoldása (optimalizálás csak lineáris egyenlőség feltételekkel). A megoldáshoz vezető úton egy $w^*(t)$ duális optimumhelyet is megkapunk. Legyen

$$\lambda_i^*(t) = -\frac{1}{tf_i(x^*(t))}, \qquad \mu^*(t) = \frac{w^*(t)}{t}.$$

Ekkor

(i) $\lambda_i^*(t)$ és $\mu^*(t)$ duális megengedett megoldása az eredeti $\mathcal{O}$ feladatnak.

(ii) Továbbá a duális hézag $x^*(t)$ primál megengedett megoldás és $\lambda_i^*(t)$ és $\mu^*(t)$ duál megengedett megoldás között $\frac{\ell}{t}$.

A tétel bizonyítása egyszerű számolás, az érdeklődő hallgatóra

LP és belsőpontos módszerek

LOGARITMIKUS GÁTFÜGGVÉNY MÓDSZER:
Kiinduló lépés: Legyen $x^{(0)}$ egy erősen megengedett megoldás
// minden egyenlőtlenség szigorúan teljesül.
Legyen $t = t^{(0)} = 1$.
// a kiinduló gátfüggvény paramétere.
$\mu (> 1)$.
// egy fix paraméter, a gátfüggvény paraméter növelési tényezője.
Centralizáló lépés: Számoljuk ki az $\widetilde{\mathcal{O}}_t$ optimalizálási feladat
$x^*(t)$ optimális értékét.
// A centrális út egy pontját számoljuk ki.
$x^+ = x^*(t)$
Kilépési kritérium: Ha $\frac{k}{t} < \epsilon$ akkor leállunk. Különben $t^+ = \mu \cdot t$
és visszatérünk a centralizáló lépésre.
// A centrális út egy későbbi (pontosabb) helyével próbálkozunk.

Elméleti következmények

Elméleti következmények

- A részletek kidolgozása, az analízis messze meghaladja az előadás keretei. A fentiek lényege csak az ötletek felvillantása volt.

Elméleti következmények

• A részletek kidolgozása, az analízis messze meghaladja az előadás kereteit. A fentiek lényege csak az ötletek felvillantása volt.

• A részletek kidolgozása, a numerikus problémák analízise nagyon sok optimalizálási feladat hatékony kezeléséhez vezet. Egy ízelítő:

Elméleti következmények

- A részletek kidolgozása, az analízis messze meghaladja az előadás kereteit. A fentiek lényege csak az ötletek felvillantása volt.
- A részletek kidolgozása, a numerikus problémák analízise nagyon sok optimalizálási feladat hatékony kezeléséhez vezet. Egy ízelítő:

Tétel

Az LP feladat polinomiális időben megoldható.

Elméleti következmények

- A részletek kidolgozása, az analízis messze meghaladja az előadás kereteit. A fentiek lényege csak az ötletek felvillantása volt.
- A részletek kidolgozása, a numerikus problémák analízise nagyon sok optimalizálási feladat hatékony kezeléséhez vezet. Egy ízelítő:

Tétel

Az LP feladat polinomiális időben megoldható.

- Ezt a tételt már láttuk az ellipszoid módszer tárgyalásakor. Ott megemlítettük, hogy a bizonyításként használt ellipszoid módszer a gyakorlatban nem versenyképes a szimplex módszerrel (amely elméleti szempontból nem kielégítő).

Elméleti következmények

• A részletek kidolgozása, az analízis messze meghaladja az előadás kereteit. A fentiek lényege csak az ötletek felvillantása volt.

• A részletek kidolgozása, a numerikus problémák analízise nagyon sok optimalizálási feladat hatékony kezeléséhez vezet. Egy ízelítő:

Tétel

Az LP feladat polinomiális időben megoldható.

• Ezt a tételt már láttuk az ellipszoid módszer tárgyalásakor. Ott megemlítettük, hogy a bizonyításként használt ellipszoid módszer a gyakorlatban nem versenyképes a szimplex módszerrel (amely elméleti szempontból nem kielégítő).

• Most azt sugalljuk, hogy a bizonyítás a gátfüggvény módszerrel is bizonyítható. Megemlítjük, hogy a fenti módszer kifinomult megvalósítása bizonyos paraméter értékek esetén versenyképes a szimplex módszerrel.

2. eset: SDP

2. eset: SDP

Minimalizáljuk	$c^T x$-t
Feltéve, hogy	$Ax = b$
	$x_1 C_1 + x_2 C_2 + \ldots + x_n C_n \succeq D$,

ahol $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{k \times n}$, a C_i, D mátrixok szimmetrikus ($\ell \times \ell$ méretű) mátrixok.

2. eset: SDP

Minimalizáljuk	$c^T x$-t
Feltéve, hogy	$Ax = b$
	$x_1 C_1 + x_2 C_2 + \ldots + x_n C_n \succeq D$,

ahol $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{k \times n}$, a C_i, D mátrixok szimmetrikus ($\ell \times \ell$ méretű) mátrixok.

- Az előjelfeltételek helyett pozitívszemidefinitség feltételünk van.

2. eset: SDP

Minimalizáljuk	$c^T x$-t
Feltéve, hogy	$Ax = b$
	$x_1 C_1 + x_2 C_2 + \ldots + x_n C_n \succeq D$,

ahol $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{k\times n}$, a C_i, D mátrixok szimmetrikus ($\ell \times \ell$ méretű) mátrixok.

- Az előjelfeltételek helyett pozitívszemidefinitség feltételünk van.
- Az egyváltozó logaritmus függvény helyett egy a szimmetrikus mátrixokon értelmezett függvényre van szükségünk.

2. eset: SDP, a gátfüggvény

2. eset: SDP, a gátfüggvény

Definíció: SDP gátfüggvény

Legyen $S \in \mathcal{S}^n$. Ekkor

$$\beta(S) = -\log\det(S).$$

2. eset: SDP, a gátfüggvény

Definíció: SDP gátfüggvény

Legyen $S \in \mathcal{S}^n$. Ekkor

$$\beta(S) = -\log\det(S).$$

- A korábbi gondolatmenetek (belsőpontos módszer) és a fenti gátfüggvény (technikai és összetett) összegzéséből adódik a következő tétel.

2. eset: SDP, a gátfüggvény

Definíció: SDP gátfüggvény

Legyen $S \in \mathcal{S}^n$. Ekkor

$$\beta(S) = -\log\det(S).$$

- A korábbi gondolatmenetek (belsőpontos módszer) és a fenti gátfüggvény (technikai és összetett) összegzéséből adódik a következő tétel. A részletek kidolgozása meghaladja előadásunk kereteit.

2. eset: SDP, a gátfüggvény

Definíció: SDP gátfüggvény

Legyen $S \in \mathcal{S}^n$. Ekkor

$$\beta(S) = -\log\det(S).$$

• A korábbi gondolatmenetek (belsőpontos módszer) és a fenti gátfüggvény (technikai és összetett) összegzéséből adódik a következő tétel. A részletek kidolgozása meghaladja előadásunk kereteit.

Tétel

Tegyük fel, hogy adott egy SDP feladat,a mely I inputjában racionális számok állnak. Legyen $\|I\|$ az input leírásához szükséges bitek száma. Adott továbbá egy tetszőleges $\epsilon > 0$ paraméter. Ekkor létezik olyan algoritmus, amely egy ϵ-szuboptimális megoldást talál és polinomiális $\|I\|$-ben és $\log(1/\epsilon)$-ban.

Vége van!

Köszönöm a figyelmet!