

多項式計画に対する半正定値緩和の不思議な性質

村松正和

電気通信大学情報理工学研究科

2011/10/14 at 九州大学

Outline

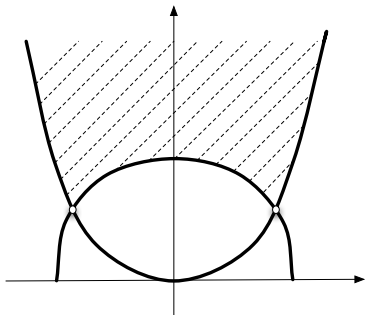
- ① 主双対内点法の奇妙な振る舞い
- ② 奇妙な振る舞いとその理由（制約なしの POP の場合）
- ③ 奇妙な振る舞いとその理由（制約のある POP の場合）

Outline

- ① 主双対内点法の奇妙な振る舞い
- ② 奇妙な振る舞いとその理由（制約なしの POP の場合）
- ③ 奇妙な振る舞いとその理由（制約のある POP の場合）

ある POP の例

$$\begin{cases} \min & y \\ \text{s.t.} & x^4 + y^4 - 1 \geq 0 \\ & -x^2 + y \geq 0 \end{cases}$$



最適値: $\sqrt{(\sqrt{5} - 1)/2} = 0.78615\dots$

SDP 緩和の数値実験

- 緩和次数 $r \leftrightarrow$ SDP 緩和問題 (昨日の話を思い出そう！)

SDP 緩和の数値実験

- 緩和次数 $r \leftrightarrow$ SDP 緩和問題 (昨日の話を出そう！)
- ζ_r : 緩和次数 r の SDP 緩和問題の最適値
 $\Rightarrow \zeta_r \leq$ POP の最適値 = 0.78615...

SDP 緩和の数値実験

- 緩和次数 $r \leftrightarrow$ SDP 緩和問題 (昨日の話を思い出そう！)
- ζ_r : 緩和次数 r の SDP 緩和問題の最適値
 $\Rightarrow \zeta_r \leq$ POP の最適値 = 0.78615...
- $\bar{\zeta}_r$: 緩和次数 r の SDP 緩和問題に対し、主双対内点法 (SeDuMi) が (最適値として) 返した値

SDP 緩和の数値実験

- 緩和次数 $r \leftrightarrow$ SDP 緩和問題 (昨日の話を思い出そう!)
- ζ_r : 緩和次数 r の SDP 緩和問題の最適値
 $\Rightarrow \zeta_r \leq$ POP の最適値 = 0.78615...
- $\bar{\zeta}_r$: 緩和次数 r の SDP 緩和問題に対し、主双対内点法 (SeDuMi) が (最適値として) 返した値

r	2	3	4	5	6	7
$\bar{\zeta}_r$	0.0000	0.0000	0.0024	0.0761	0.6813	0.7862

SDP 緩和の数値実験

- 緩和次数 $r \leftrightarrow$ SDP 緩和問題 (昨日の話を思い出そう!)
- ζ_r : 緩和次数 r の SDP 緩和問題の最適値
 $\Rightarrow \zeta_r \leq$ POP の最適値 = 0.78615...
- $\bar{\zeta}_r$: 緩和次数 r の SDP 緩和問題に対し、主双対内点法 (SeDuMi) が (最適値として) 返した値

r	2	3	4	5	6	7
$\bar{\zeta}_r$	0.0000	0.0000	0.0024	0.0761	0.6813	0.7862

- 緩和次数 $r = 7$ で最適値に到達

証明できる事実

定理 : $\zeta_r = 0$ ($r \geq 2$)

NOTE: 許容領域がコンパクトでない

⇒ Lasserre の定理が成り立つための条件は整っていない

証明できる事実

定理 : $\zeta_r = 0 (r \geq 2)$

NOTE: 許容領域がコンパクトでない

⇒ Lasserre の定理が成り立つための条件は整っていない

起こっていること

- ① SDP ソルバーは SDP の最適値として 正しくない値 を返している
- ② その「正しくない値」は POP の最適値 に収束している

知っておいてほしいこと

- ① SDP ソルバーが SDP の最適値として正しくない値を返す
 - 数値的困難を理由に解けない場合も多い

知っておいてほしいこと

- ① SDP ソルバーが SDP の最適値として正しくない値を返す
 - 数値的困難を理由に解けない場合も多い
 - POP の SDP 緩和に限らない.

知っておいてほしいこと

- ① SDP ソルバーが SDP の最適値として正しくない値を返す
 - 数値的困難を理由に解けない場合も多い
 - POP の SDP 緩和に限らない.
 - 「内点許容解」が存在しない SDP はたちが悪い
⇒ 内点許容解の回復 (Facial Reduction Algorithm)

知っておいてほしいこと

- ① SDP ソルバーが SDP の最適値として正しくない値を返す
 - 数値的困難を理由に解けない場合も多い
 - POP の SDP 緩和に限らない.
 - 「内点許容解」が存在しない SDP はたちが悪い
 - ⇒ 内点許容解の回復 (Facial Reduction Algorithm)
 - SDPA-GMP の使用
 - ⇒ 任意精度の計算ができる
 - ⇒ 非常に遅い

知っておいてほしいこと

- ① SDP ソルバーが SDP の最適値として正しくない値を返す
 - 数値的困難を理由に解けない場合も多い
 - POP の SDP 緩和に限らない.
 - 「内点許容解」が存在しない SDP はたちが悪い
 - ⇒ 内点許容解の回復 (Facial Reduction Algorithm)
 - SDPA-GMP の使用
 - ⇒ 任意精度の計算ができる
 - ⇒ 非常に遅い
- ② POP の SDP 緩和の場合、SDP ソルバーが返す正しくない値は、計算したい値である
 - スーパーラッキー？

知っておいてほしいこと

- ① SDP ソルバーが SDP の最適値として正しくない値を返す
 - 数値的困難を理由に解けない場合も多い
 - POP の SDP 緩和に限らない.
 - 「内点許容解」が存在しない SDP はたちが悪い
 - ⇒ 内点許容解の回復 (Facial Reduction Algorithm)
 - SDPA-GMP の使用
 - ⇒ 任意精度の計算ができる
 - ⇒ 非常に遅い
- ② POP の SDP 緩和の場合、SDP ソルバーが返す正しくない値は、計算したい値である
 - スーパーラッキー？
 - **必然性あり** ⇒ 今回のお話

(寄り道)SDP の双対定理

$$\theta_P = \inf \{ C \bullet X \mid A_i \bullet X = b_i (i = 1, \dots, m), X \succeq O \}$$

$$\theta_D = \sup \left\{ \sum_{i=1}^m b_i y_i \mid S = C - \sum_{i=1}^m y_i A_i, S \succeq O \right\}$$

弱双対定理: $\theta_P \geq \theta_D$

(寄り道)SDP の双対定理

$$\theta_P = \inf \{ C \bullet X \mid A_i \bullet X = b_i (i = 1, \dots, m), X \succeq O \}$$

$$\theta_D = \sup \left\{ \sum_{i=1}^m b_i y_i \mid S = C - \sum_{i=1}^m y_i A_i, S \succeq O \right\}$$

弱双対定理: $\theta_P \geq \theta_D$

双対定理 1: もし θ_P に $\bar{X} \succ O$ なる許容解が存在すれば、 $\theta_P = \theta_D$ である。このとき、もし $\theta_D > -\infty$ ならば θ_D には最適解が存在する。

(寄り道)SDP の双対定理

$$\theta_P = \inf \{ C \bullet X \mid A_i \bullet X = b_i (i = 1, \dots, m), X \succeq O \}$$
$$\theta_D = \sup \left\{ \sum_{i=1}^m b_i y_i \mid S = C - \sum_{i=1}^m y_i A_i, S \succeq O \right\}$$

弱双対定理: $\theta_P \geq \theta_D$

双対定理 1: もし θ_P に $\bar{X} \succ O$ なる許容解が存在すれば、 $\theta_P = \theta_D$ である。このとき、もし $\theta_D > -\infty$ ならば θ_D には最適解が存在する。

双対定理 2: もし θ_D に $\bar{S} \succ O$ なる許容解が存在すれば、 $\theta_P = \theta_D$ である。このとき、もし $\theta_P < \infty$ ならば θ_P には最適解が存在する。

(寄り道)SDP の双対定理

$$\theta_P = \inf \{ C \bullet X \mid A_i \bullet X = b_i (i = 1, \dots, m), X \succeq O \}$$
$$\theta_D = \sup \left\{ \sum_{i=1}^m b_i y_i \mid S = C - \sum_{i=1}^m y_i A_i, S \succeq O \right\}$$

弱双対定理: $\theta_P \geq \theta_D$

双対定理 1: もし θ_P に $\bar{X} \succ O$ なる許容解が存在すれば、 $\theta_P = \theta_D$ である。このとき、もし $\theta_D > -\infty$ ならば θ_D には最適解が存在する。

双対定理 2: もし θ_D に $\bar{S} \succ O$ なる許容解が存在すれば、 $\theta_P = \theta_D$ である。このとき、もし $\theta_P < \infty$ ならば θ_P には最適解が存在する。

NOTE: 上記は SDP より一般的な錐線形計画の枠組みで成立する

内点許容解の無い例

$$\theta_P = \inf \{ C \bullet X \mid A_1 \bullet X = b_1, X \succeq 0 \}$$

$$C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, b_1 = 2$$

$$\Downarrow X = \begin{pmatrix} x_1 & 1 \\ 1 & x_2 \end{pmatrix}$$

$$\theta_P = \inf \{ x_2 \mid x_1 x_2 \geq 1, x_1 \geq 0, x_2 \geq 0 \}$$

内点許容解の無い例

$$\theta_P = \inf \{ C \bullet X \mid A_1 \bullet X = b_1, X \succeq 0 \}$$

$$C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, b_1 = 2$$

$$\Downarrow X = \begin{pmatrix} x_1 & 1 \\ 1 & x_2 \end{pmatrix}$$

$$\theta_P = \inf \{ x_2 \mid x_1 x_2 \geq 1, x_1 \geq 0, x_2 \geq 0 \}$$

- 内点許容解が存在 ($x_1 x_2 > 1$)
- $\theta_P = 0$ だが最適解は存在しない

例の双対問題

$$\theta_D = \sup \{ b_1 y \mid S = C - yA_1, S \succeq O \}$$

$$C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, b_1 = 2$$

$$\Downarrow S = \begin{pmatrix} 0 & -y \\ -y & 1 \end{pmatrix}$$

$$\theta_D = \sup \{ 2y \mid y = 0 \} = 0 = \theta_P$$

- θ_D には最適解が存在
- θ_D には内点許容解がない
- C の (2, 2) 成分が少しでも正であれば、最適解は存在 (寄り道終)

Outline

- ① 主双対内点法の奇妙な振る舞い
- ② 奇妙な振る舞いとその理由（制約なしの POP の場合）
- ③ 奇妙な振る舞いとその理由（制約のある POP の場合）

非負多項式と SOS

非負多項式：

$$f(x) \geq 0 \quad (\forall x \in \mathbb{R}^n)$$

SOS(Sum Of Squares)：

$$f(x) = \sum_{i=1}^q g_i(x)^2 \quad (g_i(x) \text{ は多項式})$$

非負多項式と SOS

非負多項式 :

$$f(x) \geq 0 \quad (\forall x \in \mathbb{R}^n)$$

SOS(Sum Of Squares) :

$$f(x) = \sum_{i=1}^q g_i(x)^2 \quad (g_i(x) \text{ は多項式})$$

- SOS は非負多項式である

非負多項式と SOS

非負多項式 :

$$f(x) \geq 0 \quad (\forall x \in \mathbb{R}^n)$$

SOS(Sum Of Squares) :

$$f(x) = \sum_{i=1}^q g_i(x)^2 \quad (g_i(x) \text{ は多項式})$$

- SOS は非負多項式である
- 非負多項式は SOS とは限らない

例 : Motzkin Polynomial : $x^4y^2 + x^2y^4 - 3x^2y^2 + 1$

SOS と半正定値行列の関係

$2r$ 次多項式 $f(x)$ が SOS



$\exists G \succeq 0$, such that $f(x) = u_r(x)^T G u_r(x)$

$u_r(x)$: x の r 次以下の単項式を全て並べた縦ベクトル

例: $x \in \mathbb{R}^2$, $r = 2$ の場合

$$u_2(x) = (1, x_1, x_2, x_1^2, x_1x_2, x_2^2)^T$$

制約なし多項式計画の SOS 緩和

$$\begin{aligned} & \inf \{ f(x) \mid x \in \mathbb{R}^n \} \\ = & \sup \{ \zeta \mid f(x) \geq \zeta \quad (\forall x \in \mathbb{R}^n) \} \\ = & \sup \{ \zeta \mid f - \zeta \text{ が非負多項式} \} \end{aligned}$$

制約なし多項式計画の SOS 緩和

$$\begin{aligned} & \inf \{ f(x) \mid x \in \mathbb{R}^n \} \\ &= \sup \{ \zeta \mid f(x) \geq \zeta \quad (\forall x \in \mathbb{R}^n) \} \\ &= \sup \{ \zeta \mid f - \zeta \text{ が非負多項式} \} \\ &\geq \sup \{ \zeta \mid f - \zeta \text{ が SOS} \} \end{aligned}$$

SOS 緩和

$$\begin{cases} \text{maximize} & \zeta \\ \text{subject to} & u_r(x)^T G u_r(x) = f(x) - \zeta \quad \dots \quad (*) \\ & G \succeq 0 \end{cases}$$

- (*) を「両辺の単項式にかかる係数が全て等しい」と読む
⇒ G の成分に関する等式制約

Motzkin 多項式に対する SOS 緩和

$$f(x, y) = x^4y^2 + x^2y^4 - 3x^2y^2 + 1$$

- $\min f(x, y) = 0$
- SOS 緩和問題は許容解が無い
(Motzkin 多項式は SOS ではない非負多項式)

Motzkin 多項式に対する SOS 緩和

$$f(x, y) = x^4 y^2 + x^2 y^4 - 3x^2 y^2 + 1$$

- $\min f(x, y) = 0$
- SOS 緩和問題は許容解が無い
(Motzkin 多項式は SOS ではない非負多項式)

数値計算 (SparsePOP 2.97 + SeDuMi 1.3)

r	3	4	5	6	7	8
$\bar{\zeta}_r$	-1867	-42.74	-4.071	-0.471	1.820e-08	1.256e-08

- どの問題も許容解が存在すると判定
- $r = 7$ でほぼゼロが得られる

定理 (by Lasserre)

$f(x)$ を n 変数非負多項式とするとき、
任意の $\epsilon > 0$ に対しある r が存在して

$$f_{\epsilon,r}(x) = f(x) + \epsilon \sum_{k=0}^r \sum_{j=1}^n \frac{x_j^{2k}}{k!}$$

が SOS となる

定理 (by Lasserre)

$f(x)$ を n 変数非負多項式とすると、
任意の $\epsilon > 0$ に対しある r が存在して

$$f_{\epsilon,r}(x) = f(x) + \epsilon \sum_{k=0}^r \sum_{j=1}^n \frac{x_j^{2k}}{k!}$$

が SOS となる

NOTE: f と $f_{\epsilon,r}$ において、同じ単項式に対する係数の差は ϵ 以下

SOS 緩和がうまくいく仕組み

- 多項式の係数の差 \Leftrightarrow SOS 緩和問題における等式制約の差
- Motzkin 多項式に対する SOS 緩和問題において、等式制約に少しの誤差を許せば、許容解が存在するようになる。
(SOS 緩和問題は weakly infeasible)
- この誤差はいくら小さくても良い
⇒ 有限桁の計算により、自然に導入されてしまう
⇒ ただし、小さいとそれだけ緩和次数を上げる必要あり

ここまでのまとめ

- Motzkin 多項式に対する SOS 緩和問題は、
 - ① 許容解を持たないはず
 - ② SDP ソルバーは許容解を持つと判定
 - ③ SDP ソルバーの返す値は Motzkin 多項式の最小値に収束
- 非負多項式を少し「ずらす」と SOS になるため、このような現象が起こる
 - ① 「ずれる量」はいくら小さくても良い
⇒ 有限桁計算により自然に混入
 - ② 「ずれる量」が小さいと SOS になるための緩和次数を大きくとる必要あり

Outline

- ① 主双対内点法の奇妙な振る舞い
- ② 奇妙な振る舞いとその理由（制約なしの POP の場合）
- ③ 奇妙な振る舞いとその理由（制約のある POP の場合）

制約のある POP の SDP 緩和

$$\min. f_0(x) \text{ s.t. } x \in K$$

where

$$K = \{ x \mid f_i(x) \geq 0 \ (i = 1, \dots, m) \}$$

制約のある POP の SDP 緩和

$$\min. f_0(x) \text{ s.t. } x \in K$$

where

$$K = \{x \mid f_i(x) \geq 0 \ (i = 1, \dots, m)\}$$

(制約なしの場合の)SOS 緩和の拡張

$$f_0(x) - \zeta = \sigma_0(x) + \sum_{i=1}^m f_i(x)\sigma_i(x)$$

- $\sigma_i(x)$: SOS. ($\sigma_i(x)$ の次数はそれぞれの項における最大次数が緩和次数 r に対して $2r$ か $2r - 1$ となるようにとる)
- 制約なし \Leftrightarrow 第 1 項のみ

制約のある POP の SDP 緩和

$$\min. f_0(x) \text{ s.t. } x \in K$$

where

$$K = \{ x \mid f_i(x) \geq 0 \ (i = 1, \dots, m) \}$$

(制約なしの場合の) SOS 緩和の拡張

$$f_0(x) - \zeta = \sigma_0(x) + \sum_{i=1}^m f_i(x) \sigma_i(x)$$

- $\sigma_i(x)$: SOS. ($\sigma_i(x)$ の次数はそれぞれの項における最大次数が緩和次数 r に対して $2r$ か $2r - 1$ となるようにとる)
- 制約なし \Leftrightarrow 第 1 項のみ
- もし等式が成り立てば、 $f_0(x) - \zeta \geq 0$ for all $x \in K$.

制約のある POP の SDP 緩和

$$\min. f_0(x) \text{ s.t. } x \in K$$

where

$$K = \{ x \mid f_i(x) \geq 0 \ (i = 1, \dots, m) \}$$

(制約なしの場合の) SOS 緩和の拡張

$$f_0(x) - \zeta = \sigma_0(x) + \sum_{i=1}^m f_i(x) \sigma_i(x)$$

- $\sigma_i(x)$: SOS. ($\sigma_i(x)$ の次数はそれぞれの項における最大次数が緩和次数 r に対して $2r$ か $2r - 1$ となるようにとる)
- 制約なし \Leftrightarrow 第 1 項のみ
- もし等式が成り立てば、 $f_0(x) - \zeta \geq 0$ for all $x \in K$.
- それぞれの SOS は半正定値行列としてモデル化

$$\sigma_i(x) = u_{r_i}(x)^T G_i u_{r_i}(x) \quad (i = 0, \dots, m)$$

Lasserre の SDP 緩和

$$\min. f_0(x) \text{ s.t. } x \in K$$

↓

$$\begin{aligned} \max. \quad & \zeta \\ \text{s.t.} \quad & f_0(x) - \zeta \\ & = u_{r_0}(x)^T G_0 u_{r_0}(x) + \sum_{i=1}^m f_i(x) u_{r_i}(x)^T G_i u_{r_i}(x) \\ & G_i \succeq O \quad (i = 0, \dots, m) \end{aligned}$$

- Lasserre の SDP 緩和の最適値は $r_0 \rightarrow \infty$ で POP の最適値へ収束
- 必要な仮定: K がコンパクト + もう一つ

Lasserre の SDP 緩和

$$\min. f_0(x) \text{ s.t. } x \in K$$

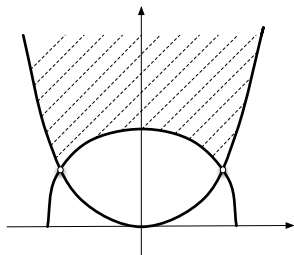
↓

$$\begin{aligned} \max. \quad & \zeta \\ \text{s.t.} \quad & f_0(x) - \zeta \\ & = u_{r_0}(x)^T G_0 u_{r_0}(x) + \sum_{i=1}^m f_i(x) u_{r_i}(x)^T G_i u_{r_i}(x) \\ & G_i \succeq O \quad (i = 0, \dots, m) \end{aligned}$$

- Lasserre の SDP 緩和の最適値は $r_0 \rightarrow \infty$ で POP の最適値へ収束
- 必要な仮定： K がコンパクト + もう一つ
NOTE: 最初の問題は K がコンパクトでない

POP の例 (再掲)

$$\begin{cases} \min & x_2 \\ \text{s.t.} & x_1^4 + x_2^4 - 1 \geq 0 \\ & -x_1^2 + x_2 \geq 0 \end{cases}$$



- Lasserre の定理の仮定を満たさない
- SDP の最適値は POP の最適値に収束しているように見える
- 本当の SDP の最適値は 0

摂動定理の記述に必要な仮定

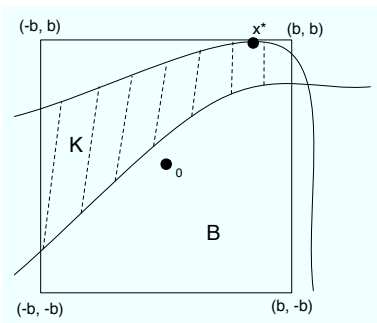
- K は非空 (コンパクト性は仮定しない)
- 最適解 x^* が存在

$$b = \max(1, \max\{|x_j^*| \mid j = 1, \dots, n\})$$

$$B = [-b, b]^n$$

$$\bar{K} = K \cap B$$

$$R_i = \max\{|f_i(x)| \mid x \in B\}$$



摂動定理

- 仮定 : $f(x) - \zeta > 0$ for every $x \in \bar{K}$
($\Leftrightarrow \zeta$ は POP の最適値の下界)
- 結論 : 任意の $\epsilon > 0$ に対し次を満たす正整数 \hat{r} と \tilde{r} が存在 :
「 $r \geq \hat{r}$ なる任意の r に対し、

$$f(x) - \zeta + \epsilon \Theta_r(x/b) + \psi_{\tilde{r}}(x)$$

が SOS」ただし、

$$\Theta_r(x) = 1 + \sum_{j=1}^n x_j^{2r}$$
$$\psi_r(x) = - \sum_{i=1}^m f_i(x) \left(1 - \frac{f_i(x)}{R_i} \right)^{2r}$$

摂動定理とSDP緩和の関係

$$f(x) - \zeta + \epsilon \Theta_r(x/b) + \psi_r(x) = \sigma_0(x)$$

$$\psi_r(x) = - \sum_{i=1}^m f_i(x) \left(1 - \frac{f_i(x)}{R_i}\right)^{2r}$$

↓

$$f(x) - \zeta + \epsilon \Theta_r(x/b) = \sigma_0(x) + \sum_{i=1}^m f_i(x) \left(1 - \frac{f_i(x)}{R_i}\right)^{2r}$$

摂動定理とSDP緩和の関係

$$f(x) - \zeta + \epsilon \Theta_r(x/b) + \psi_r(x) = \sigma_0(x)$$

$$\psi_r(x) = - \sum_{i=1}^m f_i(x) \left(1 - \frac{f_i(x)}{R_i}\right)^{2r}$$

↓

$$f(x) - \zeta + \epsilon \Theta_r(x/b) = \sigma_0(x) + \sum_{i=1}^m f_i(x) \left(1 - \frac{f_i(x)}{R_i}\right)^{2r}$$

cf. Lasserre のSDP緩和

$$f(x) - \zeta = \sigma_0(x) + \sum_{i=1}^m f_i(x) \sigma_i(x)$$

摂動定理の意味すること

- SOS 緩和問題に少しの誤差を許せば、許容解が存在 (NOTE: $\Theta_r(x)$ の係数の最大値は 1)
- この誤差はいくら小さくても良い
 - ⇒ 有限桁の計算により、自然に誤差が導入される
 - ⇒ ただし、誤差が小さいと緩和次数を大きくする必要有り
- 以上のことが一般的な場合に言えていることが面白い

まとめ

POP の SOS 緩和においては、

- SDP ソルバーが最適値でない値を返すことがある。
- その値は POP の最適値へ「収束」する
- この差は、有限桁計算により誤差が混入することで起こる
⇒ 誤差のおかげで最適値が計算できる
- 上記の現象を説明するのは
 - ① 制約なしの場合：Lasserre の定理
 - ② 制約有りの場合：今回の摂動定理