不確実性下での最適化手法:
ロバスト最適化法の紹介

数理最適化分野でのロバストネス

東京大学・理研　　武田朗子

概要

- ロバスト最適化法
  - ✓ 概要 （考え方、歴史的経緯）
  - ✓ ロバスト最適化問題の解き方
  - ✓ 機械学習分野の判別モデルにおけるロバスト性

不確実性を考慮した最適化法

最適化問題:

$$\min_{\boldsymbol{x}} \quad f(\boldsymbol{x})$$
$$\text{s.t.} \quad g_i(\boldsymbol{x}) \leq 0, \qquad i = 1, \ldots, m$$

●確率計画法

- ✓ 古くから（1950年代より）研究あり
- ✓ 確率分布を仮定して，リスク尺度を用いたモデル化

●ロバスト最適化法

- ✓ 1998年にBen-Tal & Nemirovskiによって提案
- ✓ 不確実なデータの取りうる範囲を仮定して，最悪状況で最適な意思決定をするようなモデル化

ロバスト最適化法の基礎と使い方について紹介

ロバスト最適化法とは？

Ben-Tal & Nemirovski ['98]

不確実性を含んだ最適化問題のモデル化とその解法

目的:　どんな状況が起こっても大きな損害を被らないよう、意思決定をしたい

●電力供給コスト最小にするための生産量は？

$$\min_{x \geq 0} \sum_{i=1}^{3} u_i x_i$$

$$\text{s.t.} \quad \sum_{i=1}^{3} x_i \geq d$$

3 基の発電機
(燃料: 石油,石炭,ガス)

x_i : 生産量

u_i : 単位コスト, d : 需要量

不確実性大

ロバスト最適化問題

不確実なデータの範囲をあらかじめ設定した上で、
最悪事態の発生を想定して最適化を行なう問題

不確実なデータの範囲を設定：$\boldsymbol{u} \in \mathcal{U}, \ \ d \in [d_\ell, d_u]$

$$\min_{\boldsymbol{x} \geq \boldsymbol{0}} \ \sum_{i=1}^{3} u_i x_i \quad \Rightarrow \quad \min_{\boldsymbol{x}} \max_{\boldsymbol{u} \in \mathcal{U}} \sum_{i=1}^{3} u_i x_i$$

$$\text{s.t.} \ \sum_{i=1}^{3} x_i \geq d \quad \Rightarrow \quad \sum x_i \geq d, \ \ \forall d \in [d_\ell, d_u]$$

$$\Leftrightarrow \quad \sum_{i=1}^{3} x_i \geq \max_{d \in [d_\ell, d_u]} d$$

データの不確実性を考えることの大切さ

Ben-Tal & Nemirovski ['00]

PILOT4 (NETLIB library)

1000変数、410制約式、最適解を $\boldsymbol{x}^*$ とする

$$\boldsymbol{a}^\top \boldsymbol{x} \equiv$$

$$-15.79081x_{826} - 8.598819x_{827} - 1.88789x_{828} - 1.362417 \cdots$$
$$-0.031883x_{849} - 28.725555x_{850} - 10.792065x_{851} - 0.190 \cdots$$
$$-12.290832x_{854} + 717.562256x_{855} - 0.057865x_{856} - 3.785$$
$$-122.163055x_{859} - 6.46609x_{860} - 0.48371x_{861} - 0.615264 \cdots$$
$$-84.644257x_{864} - 122.459045x_{865} - 43.15593x_{866} - 1.712 \cdots$$
$$+x_{880} - 0.946049x_{898} - 0.946049x_{916} \geq \boxed{23.387405} \equiv b$$

実数値係数を0.1%ずらす $\rightarrow \overline{\boldsymbol{a}}$

たとえば、$15.79081 \times 0.001 = 0.0157908$

$\boldsymbol{a}^\top \boldsymbol{x}^* - b \geq 0$ を満たす解 $\boldsymbol{x}^*$で，摂動を加えた式を評価：

$\overline{\boldsymbol{a}}^\top \boldsymbol{x}^* - b < -104.9$ （大きく制約式を破る）

ロバスト最適化法の適用例 -1-

ロバスト性を考慮した建築構造設計: 寒野 ['08]

建築構造物は外力に対し、不安定現象を生じないように設計の必要あり

→構造物に作用する外力を，正確に予測することは難しい

→構造物の解析・設計において，外力のばらつきを考慮

ロバスト化学プロセス設計 Lin, Janak & Floudas ['04]

化学プロセスは反応器、蒸留塔などの単位操作の組み合わせで、原材料から目的物質を製造

→環境影響(外因)による安全性への影響を考慮して、経済性最大化を達成する必要あり

ロバスト最適化法の適用例 -2-

土谷 & 笹川 ['05]

超電導磁気浮上式列車の磁気シールド最適設計:
車体を強磁性体のシールドで覆い、車内を強力な磁場から遮蔽する必要あり
→車内へ磁場が漏洩しない条件下でシールド重量最小化
→“近似計算”による外部磁場の不確実性を考慮

Yu & Yang ['98]

ロバスト最短路問題(離散最適化):
渋滞等による所要時間の不確実性を考慮した最短路問題

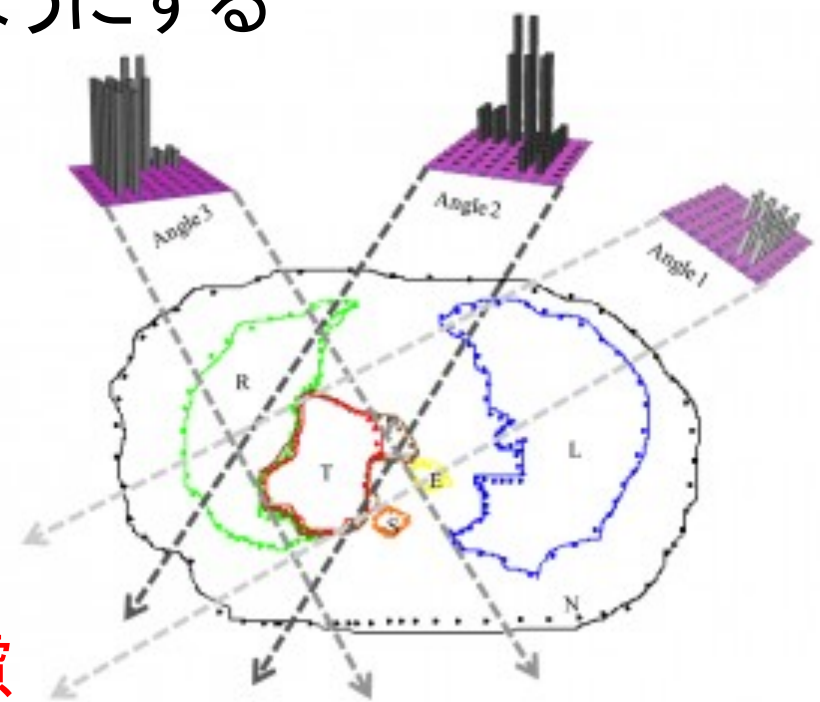
クラス分類モデル (機械学習):
観測値に観測誤差が含まれることを想定した手法

ポートフォリオ選択 (金融工学):
株価の不確実性を考慮して、最適資産配分を決定

放射線治療への適用

癌患者のための放射線治療　Chan, Bortfeld & Tsitsiklis ['06]

放射線をいろいろな角度から照射する際に，
＊腫瘍に重なるようにする
＊周辺の重要な臓器を避けるようにする



→ 放射線ビームの角度と
強度などを最適化法で決定

→ 腫瘍の位置の不確実性を考慮
（例えば，治療中に患者が呼吸することで，肺腫瘍の位置が動く）

他手法との関係 (vs. 確率計画)

不確実な最適化問題:

$$\min_{\boldsymbol{x} \in X} f(\boldsymbol{x}, \boldsymbol{u}_0) \text{ s.t. } g(\boldsymbol{x}, \boldsymbol{u}_1) \leq 0$$

$\boldsymbol{u}_0, \boldsymbol{u}_1$
不確実なデータ

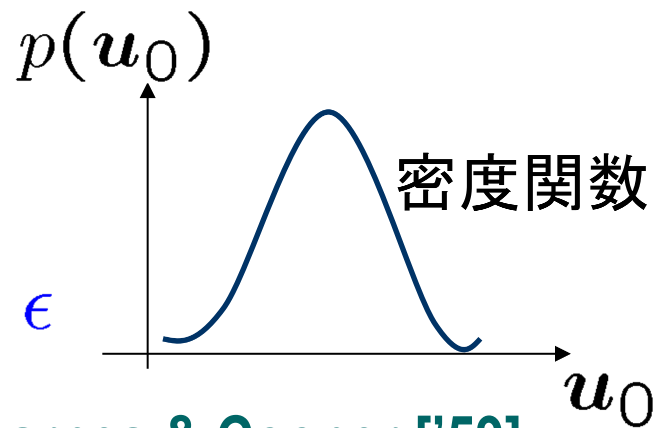
●確率計画法

Dantzig, Beale ['55]

仮定: $\boldsymbol{u}_0, \boldsymbol{u}_1$ の確率分布が既知

$$\min_{\boldsymbol{x} \in X} E_{\boldsymbol{u}_0}[f(\boldsymbol{x}, \boldsymbol{u}_0)]$$

$$\text{s.t. } \Pr_{\boldsymbol{u}_1}\{g(\boldsymbol{x}, \boldsymbol{u}_1) \leq 0\} \geq 1 - \epsilon$$



機会制約(確率制約) Charnes & Cooper ['59]

Cf.) ロバスト最適化法

仮定: $\boldsymbol{u}$ の属する領域 $\mathcal{U}$ が既知

$$\min_{\boldsymbol{x} \in X} \max_{\boldsymbol{u} \in \mathcal{U}} f(\boldsymbol{x}, \boldsymbol{u})$$

最適化分野でのロバスト最適化法の歴史

ロバスト最適化法

$$\min_{\boldsymbol{x}\in X}\max_{\boldsymbol{u}\in\mathcal{U}} f(\boldsymbol{x},\boldsymbol{u})$$

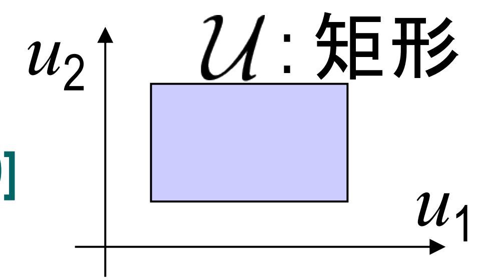
$\boldsymbol{u}$ の属する領域 $\mathcal{U}$ が既知

■1973年 A.L.Soyster が 矩形の $\mathcal{U}$

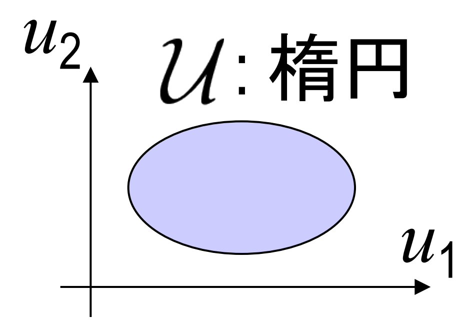
(手法はinexact LPと呼ばれ，解は “保守的すぎる”)

この間、進展なし (論文2本†)

†) **reported by Ben-Tal, El Ghaoui & Nemirovski ['09]**



■1998年 Ben-Tal & Nemirovskiが楕円形
El Ghaoui & Lebret, Kouvelis & Yuも同時期に同様の定式化



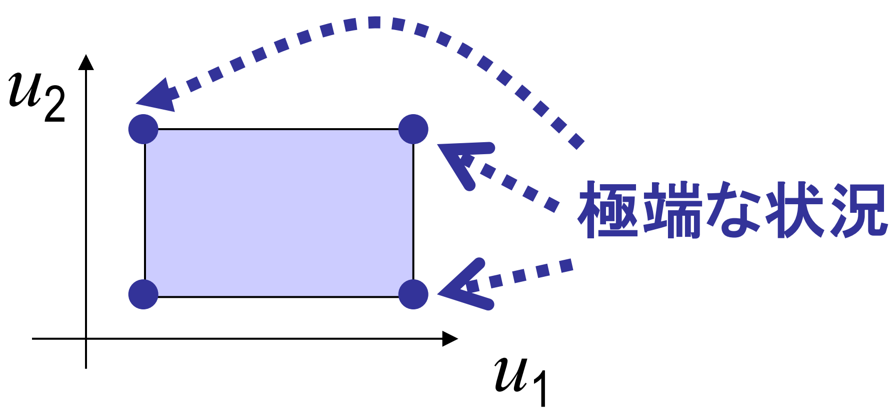
■続々と、研究つづく

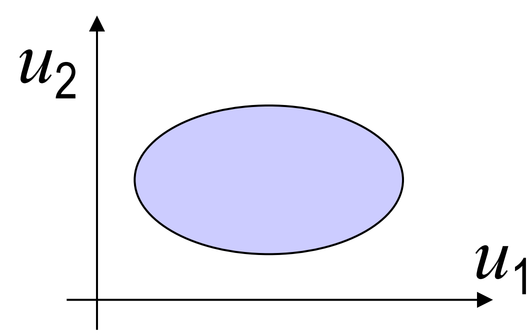
ロバスト最適化がうけた理由

① 矩形だと、極端な状況を想定するだけ.
　この欠点が解消された.

② 二次錐計画や半正定値計画に帰着
　（ ➡ 内点法で解ける）.

inexact LP Soyster ['73]		ロバストLP Ben-Tal & Nemirovski ['98]
$\mathcal{U}$ が矩形	$\Longrightarrow$	$\mathcal{U}$ が楕円 etc..





ロバスト最適化研究の様々な流れ

ロバスト最適化問題の原型
Soyster ['73]

ロバスト最適化問題の
基本的な定式化
Ben-Tal & Nemirovski ['98,'99]

確率的アプローチ
Calafiore & Campi ['05,'06]

金融工学,
機械学習の判別分析
etc. への応用

多期間モデルへの拡張
Ben-Tal, et.al. ['04]

解きやすいロバスト
最適化問題の特徴づけ
Goldfarb & Iyengar ['03]

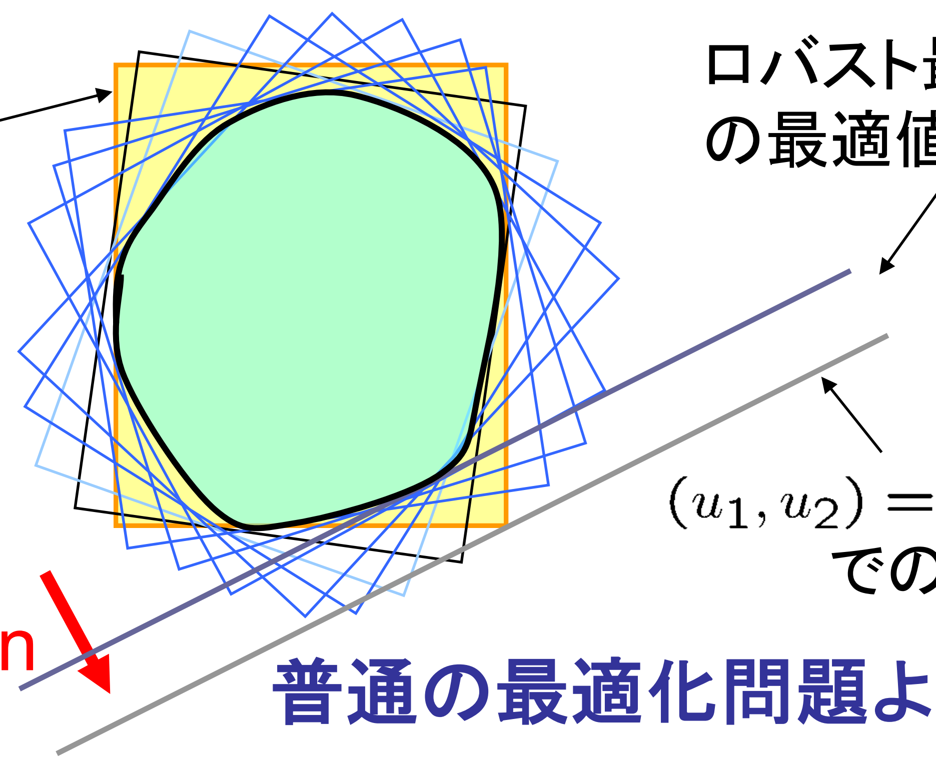
概要

- ロバスト最適化法
  - ✓ 概要（考え方、歴史的経緯）
  - ✓ ロバスト最適化問題の解き方
  - ✓ 機械学習分野の判別モデルにおけるロバスト性

ロバスト最適化問題のイメージ

制約式は無限本

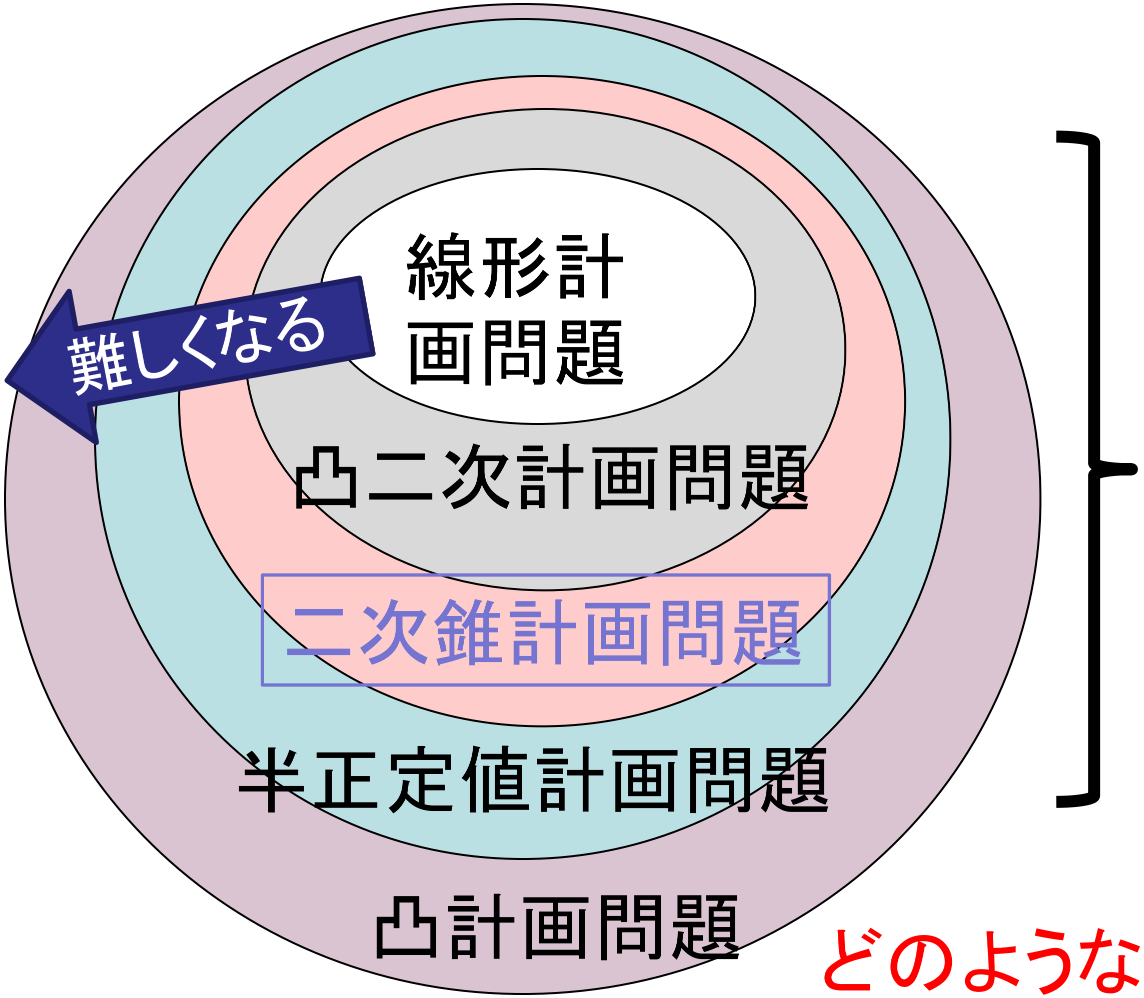
$$
\begin{aligned}
\min_{\boldsymbol{x}} \quad & -x_1 + x_2 \\
\text{s.t.} \quad & -1 \le \quad u_1 x_1 + u_2 x_2 \le 1 \\
& -1 \le -u_2 x_1 + u_1 x_2 \le 1 \\
& \forall \boldsymbol{u} \in \mathcal{U} = \left\{ (u_1, u_2) \mid u_1^2 + u_2^2 = 1 \right\}
\end{aligned}
$$



普通の最適化問題より難しい！

代表的な最適化問題

連続最適化



効率的に解くための
アルゴリズムやその
ソフトウェアが揃っている

どのようなロバスト最適化
問題が，これらに帰着できるか？

ロバスト最適化問題の標準形

$$\min_{\boldsymbol{x} \in X} \ \boldsymbol{c}^\top \boldsymbol{x} \quad \text{s.t.} \quad f_i(\boldsymbol{x}, \boldsymbol{u}_i) \le 0, \quad \forall \boldsymbol{u}_i \in \mathcal{U}_i, \\ i = 1, \ldots, m$$

制約毎の不確実性 （constraint-wise uncertainty）

$f_i(\boldsymbol{x}, \boldsymbol{u}_i)$: $\boldsymbol{u}_i$ について凸　$(\forall \boldsymbol{u}_i \in \mathcal{U}_i)$

X: 閉凸集合,　$\mathcal{U}_i$: 有界閉集合

- 目的関数が不確実なとき

$$\min_{\boldsymbol{x} \in X} \max_{\boldsymbol{u}_0 \in \mathcal{U}_0} f_0(\boldsymbol{x}, \boldsymbol{u}_0)$$

$$\Longrightarrow \quad \min_{\boldsymbol{x} \in X, t} \ t \quad \text{s.t.} \quad f_0(\boldsymbol{x}, \boldsymbol{u}_0) \le t, \ \forall \boldsymbol{u}_0 \in \mathcal{U}_0$$

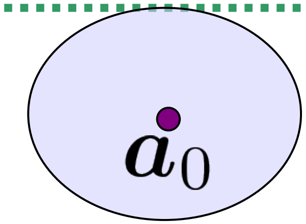
ロバストLP （楕円型不確実性集合）

Ben-Tal & Nemirovski ['99]

$$\min_{\boldsymbol{x}} \ \boldsymbol{c}^\top \boldsymbol{x} \quad \text{s.t.} \quad \boldsymbol{a}^\top \boldsymbol{x} \le b, \quad \forall \boldsymbol{a} \in \mathcal{U}$$

楕円型:

$$\mathcal{U} = \{\boldsymbol{a}_0 + \boldsymbol{A}\boldsymbol{u} : \|\boldsymbol{u}\|_2 \le 1\}$$



$$\min_{\boldsymbol{x}} \ \boldsymbol{c}^\top \boldsymbol{x} \quad \text{s.t.} \quad \boldsymbol{a}_0^\top \boldsymbol{x} + \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{u} \le b, \quad \|\boldsymbol{u}\|_2 \le 1,$$

$$\boldsymbol{a}_0^\top \boldsymbol{x} + \left(\max_{\boldsymbol{u}:\|\boldsymbol{u}\|_2 \le 1} \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{u} \right) \le b$$

$$\boxed{\boldsymbol{u}^* = \frac{\boldsymbol{A}^\top \boldsymbol{x}}{\|\boldsymbol{A}^\top \boldsymbol{x}\|_2}}$$

$$\min_{\boldsymbol{x}} \ \boldsymbol{c}^\top \boldsymbol{x} \quad \text{s.t.} \quad \boldsymbol{a}_0^\top \boldsymbol{x} + \|\boldsymbol{A}^\top \boldsymbol{x}\|_2 \le b$$

二次錐計画問題

ロバスト最適化問題は解ける！？

$$\min_{\boldsymbol{x}\in X} \quad \boldsymbol{c}^\top \boldsymbol{x} \quad \text{s.t.} \quad f_i(\boldsymbol{x}, \boldsymbol{u}_i) \le 0, \quad \forall \boldsymbol{u}_i \in \mathcal{U}_i,$$
$$i = 1, \ldots, m$$

➡ 解きやすい凸計画問題に変形するためには .....

Ben-Tal & Nemirovski ['98], Goldfarb & Iyengar ['03]

3つの仮定

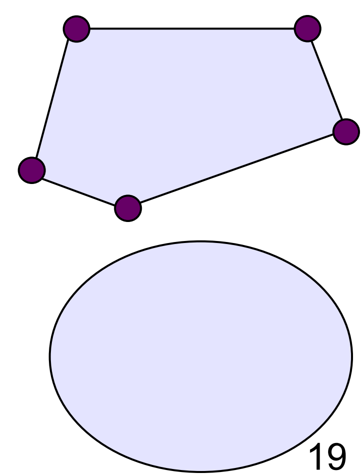
(1) $f(\boldsymbol{x}, \boldsymbol{u})$ は $\boldsymbol{x}$ について凸二次

$$f(\boldsymbol{x}, \boldsymbol{u}) = \boldsymbol{x}^\top \boxed{\boldsymbol{Q}(\boldsymbol{u})} \boldsymbol{x} + \boxed{\boldsymbol{q}(\boldsymbol{u})}^\top \boldsymbol{x} + \boxed{\gamma(\boldsymbol{u})}$$

(2) 不確実データは $\boldsymbol{u}$ について線形

$$\boldsymbol{Q}(\boldsymbol{u}) = \boldsymbol{Q}_0 + \textstyle\sum_i \boldsymbol{Q}_i u_i$$
$$\boldsymbol{q}(\boldsymbol{u}) = \boldsymbol{q}_0 + \textstyle\sum_i \boldsymbol{q}_i u_i$$

(3) $\mathcal{U}_i$ は有限集合、その凸包、もしくは楕円



ロバスト最適化の定式化のコツ

ロバスト最適化では.....
✓不確実データの表現方法がポイント！
✓その表現方法に制限が大きい
- 不確実データは$\boldsymbol{u}$について線形、
- $\boldsymbol{u}$は楕円内を動く, etc.

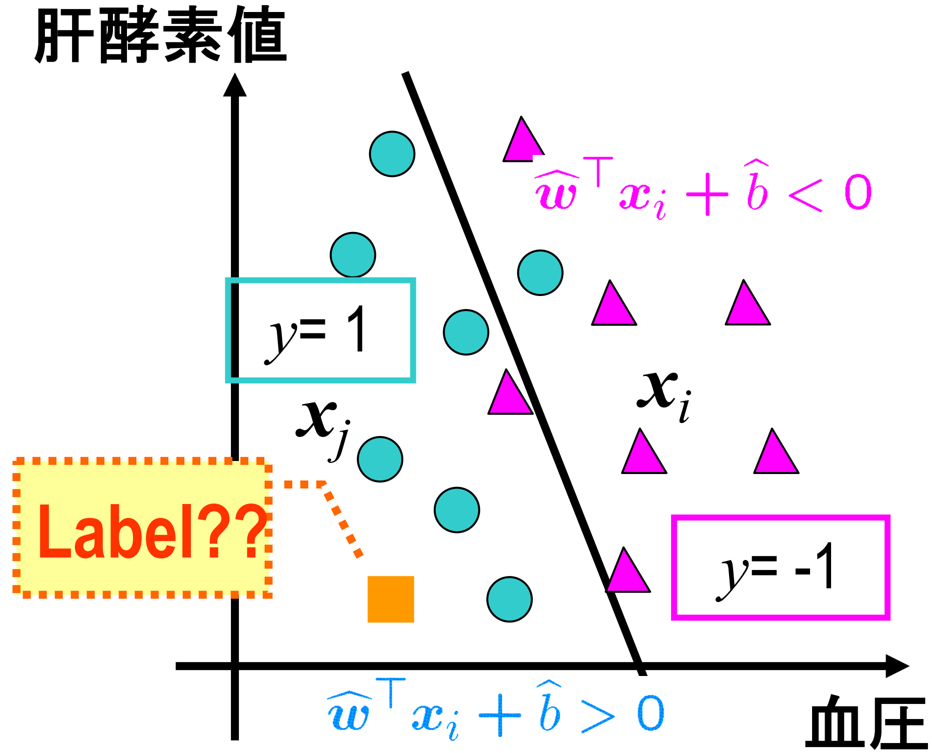
✓その制限を満たせば、扱いやすい問題に変換できる

- どういう問題構造をしていれば扱いやすい問題に変換できる？
- 条件を満たさない場合は、無限本制約をすべて考慮するのでなく、うまくサンプリングすることによる確率的解法

概要

- ロバスト最適化法
  - ✓ 概要（考え方、歴史的経緯）
  - ✓ ロバスト最適化問題の解き方
  - ✓ 機械学習分野の判別モデルにおけるロバスト性

肝臓疾患の判別



特徴ベクトル

$\boldsymbol{x}_i$= (赤血球の濃さ, ……, 肝酵素値, 一日の飲酒量)

$$h(\boldsymbol{x}) = \mathrm{sign}(\boldsymbol{w}^\top \boldsymbol{x} + b)$$

y_i= { 1 健康 ; -1 肝臓疾患

線形判別関数を扱うがカーネル法により非線形関数に拡張可能

i =1 ,…, m

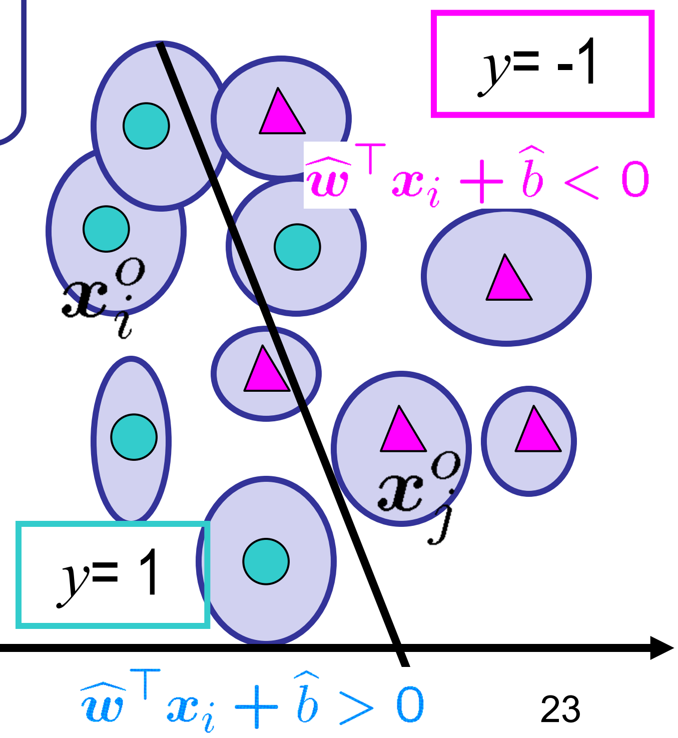
学習に用いなかった未知のデータ xのラベル y をうまく予測したい

SVM（サポートベクトルマシーン）における ロバスト性

所与のデータに過適合（オーバーフィット）しない判別関数を求めるために**正則化**

定式化は同値
Xu, Caramanis & Mannor ['09]

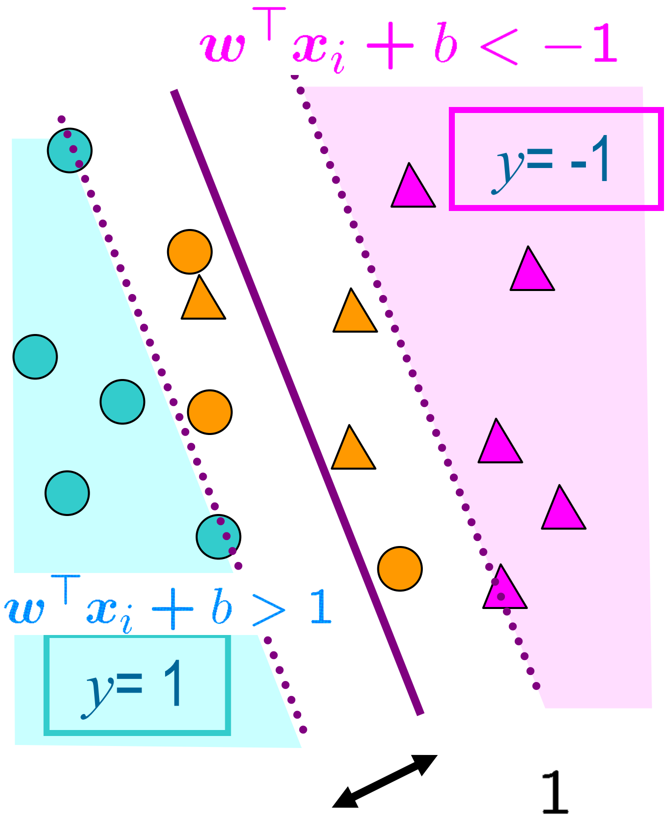
所与のデータの微小変動に対して頑健な解を求めるために**ロバスト化**



判別分析における「ロバスト性」は「正則化」によって考慮されている

ソフトマージンSVM（線形分離不可能な場合）

Cortes & Vapnik ['95]



マージン= $\dfrac{1}{\|\boldsymbol{w}\|}$

$$
\begin{aligned}
\min_{\boldsymbol{w},b,\boldsymbol{z}} \quad & \delta\|\boldsymbol{w}\| + \sum_{i=1}^{m} z_i \\
\text{s.t.} \quad & y_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b) \geq 1 - z_i, \\
& z_i \geq 0, \quad i = 1, \ldots, m
\end{aligned}
$$

δ は2つの目的をコントロール

$$
\left\{
\begin{array}{ll}
\text{●損失最小} & \left(\min \sum_{i=1}^{m} z_i \right) \\
\text{●マージン最大化} & \left(\max \dfrac{1}{\|\boldsymbol{w}\|} \right)
\end{array}
\right.
$$

この目的を加えることを正則化
（過適合を防ぐ役割）

正則化＝ロバスト化

正則化

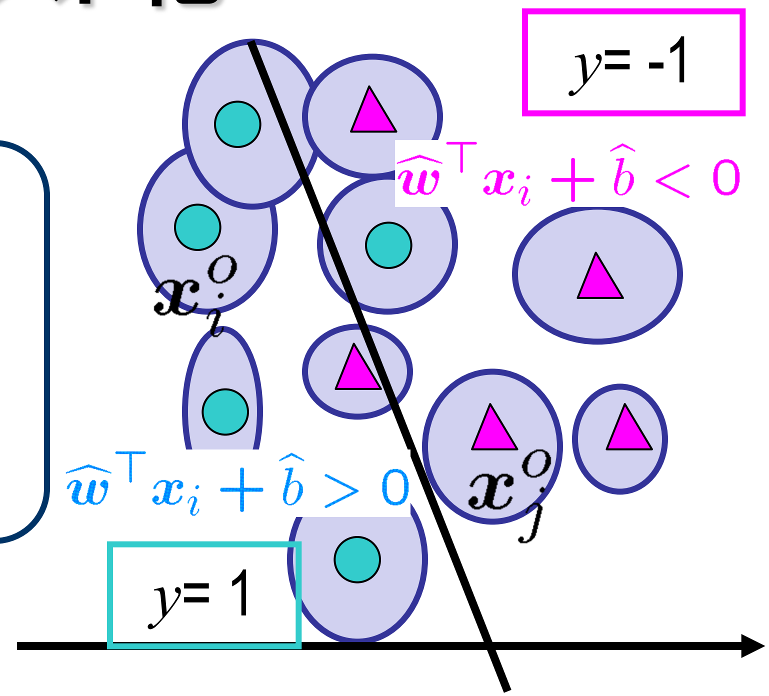
$$\min_{\boldsymbol{w},b,\boldsymbol{z}} \quad \delta\|\boldsymbol{w}\| + \sum_{i=1}^{m} z_i$$
$$\text{s.t.} \quad y_i(\boldsymbol{w}^\top \boldsymbol{x}_i^o + b) \geq 1 - z_i,$$
$$z_i \geq 0, \quad i = 1, \ldots, m$$

同値

正則化項を除く

$$\min_{\boldsymbol{w},b} \quad \sum_{i=1}^{m} [1 - y_i(\boldsymbol{w}^\top \boldsymbol{x}_i^o + b)]^+$$

$$[a]^+ := \max\{a, 0\}$$



ロバスト化

Xu, Caramanis & Mannor ['09]

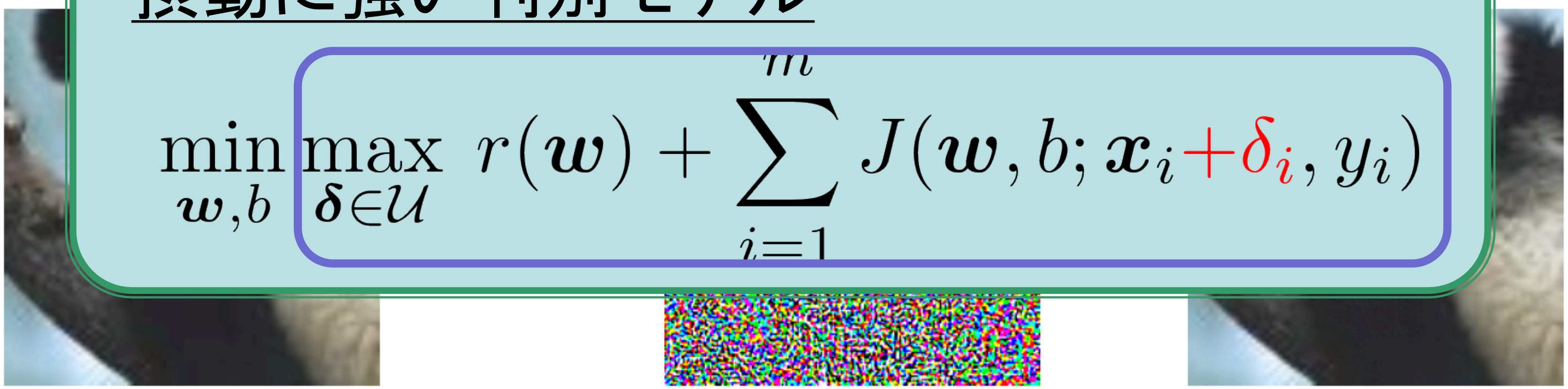
$$\min_{\boldsymbol{w},b} \max_{(\Delta\boldsymbol{x}_1,\ldots,\Delta\boldsymbol{x}_m)\in\mathcal{U}} \sum_{i=1}^{m} [1 - y_i\{\boldsymbol{w}^\top(\boldsymbol{x}_i^o + \Delta\boldsymbol{x}_i) + b\}]^+$$

$$\mathcal{U} = \{(\Delta\boldsymbol{x}_1, \ldots, \Delta\boldsymbol{x}_m) : \sum_{i=1}^{m} \|\Delta\boldsymbol{x}_i\| \leq \delta\}$$

敵対的摂動 (adversarial example)

画像に摂動を与えることにより誤分類してしまう

り抜粋



<u>摂動に強い判別モデル</u>

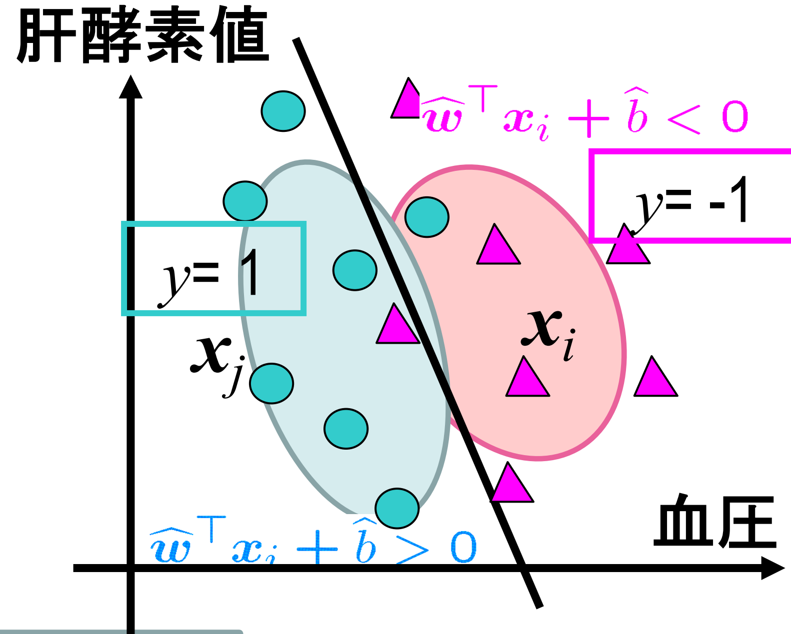
$$\min_{\boldsymbol{w}, b} \boxed{\max_{\boldsymbol{\delta} \in \mathcal{U}} \; r(\boldsymbol{w}) + \sum_{i=1}^{m} J(\boldsymbol{w}, b; \boldsymbol{x}_i + \delta_i, y_i)}$$

目的関数値が増大する勾配

判別モデル: $$\min_{\boldsymbol{w}, b} \; r(\boldsymbol{w}) + \sum_{i=1}^{m} J(\boldsymbol{w}, b; \boldsymbol{x}_i, y_i)$$

ロバスト最適化による判別モデルの解釈

Takeda, Mitsugi & Kanamori
(*ICML* ['12], *Neural Computation* ['13])



ロバスト最適化法による,
観測データに頑健な判別モデル

$$\max_{\|\boldsymbol{w}\|=1} \min_{\boldsymbol{x}_+ \in \mathcal{U}_+, \boldsymbol{x}_- \in \mathcal{U}_-} (\boldsymbol{x}_+ - \boldsymbol{x}_-)^\top \boldsymbol{w}$$

$\Rightarrow$ $f(\boldsymbol{x}) = \boldsymbol{w}^{*\top} \boldsymbol{x} + b^*$

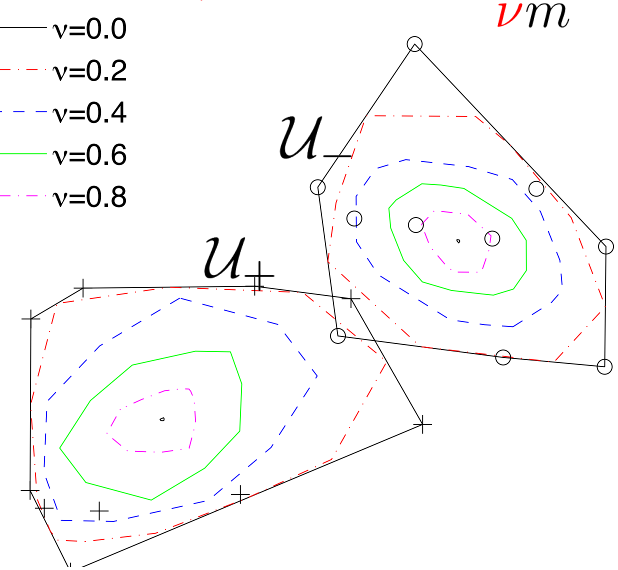
- ✓ $\boldsymbol{x}_+, \boldsymbol{x}_-$: 各クラスの平均ベクトル（真の分布が未知なので不確定）
- ✓ $\mathcal{U}_+, \mathcal{U}_-$: 有界閉凸な不確実性集合 ($\boldsymbol{x}_+, \boldsymbol{x}_-$の動きうる範囲）
- ✓ b は $\boldsymbol{x}_+^*$ と $\boldsymbol{x}_-^*$ の真ん中を通るように計算される

不確実性集合の例

$\mathcal{U}_+$ と $\mathcal{U}_-$ は各クラスのデータ集合より構築される

⇦ パラメータκによって大きさコントロール

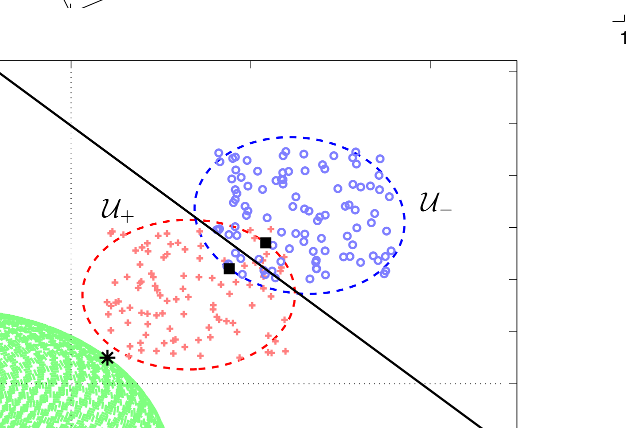
パラメータ：$\kappa = \dfrac{2}{\nu m}$



縮退凸包 (reduced convex hull)

Bennett & Bredensteiner ['00], Crisp & Burges ['00]

$$\mathcal{U}_+ = \left\{ \sum_{i \in M_+} \lambda_i \boldsymbol{x}_i \,:\, \boldsymbol{e}^\top \boldsymbol{\lambda} = 1,\ \boldsymbol{0} \le \boldsymbol{\lambda} \le \kappa \boldsymbol{e} \right\}$$



楕円2：

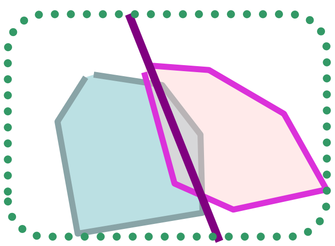
$$\mathcal{U}_+ = \left\{ \bar{\boldsymbol{x}}_+ + \Sigma_+^{1/2} \boldsymbol{u} \,:\, \|\boldsymbol{u}\| \le \kappa \right\}$$

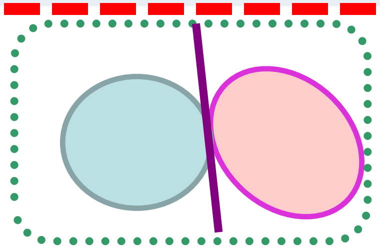
訓練データの平均: $\bar{\boldsymbol{x}}_+$, $\bar{\boldsymbol{x}}_-$

分散共分散行列： Σ_+, Σ_-

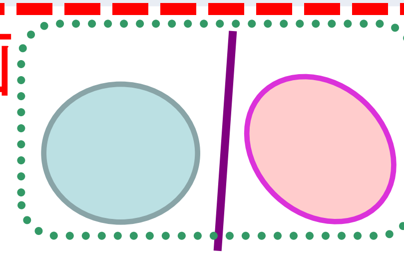
既存モデルとの関係

u_+, u_-	交わりあり	接する	交わりなし
楕円1：	なし	**Fisher Discriminant Analysis (FDA)** Fukunaga ('90)	**Sparse Feature Selection** Bhattacharyya ('04)
楕円2：	なし	**Minimax Probability Machine (MPM)** Lanckriet et al. ('02)	**Minimum Margin-MPM** Nath & Bhattacharyya ('07)
縮退凸包:	**Eν-SVM** Perez-Cruz et al. ('03)	**ν-SVM** (= Soft Margin SVM) Scholkopf et al. ('00)	
凸包:		------	**Hard Margin SVM** Boser et al. ('92)





凸計画



他に，L2-SVM, ロジスティック回帰なども…

既存モデルとの関係

$\mathcal{U}_+, \mathcal{U}_-$	交わりあり	接する	交わりなし
楕円1：	モデル提案 (非凸2次最適化)	**Fisher Discriminant Analysis (FDA)** Fukunaga ('90)	**Sparse Feature Selection** Bhattacharyya ('04)
楕円2：	モデル提案 (非凸2次最適化)	**Minimax Probability Machine (MPM)** Lanckriet et al. ('02)	**Minimum Margin-MPM** Nath & Bhattacharyya ('07)
縮退凸包：	**Eν-SVM** Perez-Cruz et al. ('03)	**ν-SVM** (= Soft Margin SVM) Scholkopf et al. ('00)	
凸包：		------	**Hard Margin SVM** Boser et al. ('92)

凸計画

一般的な非凸2次最適化問題の解法提案
(AISTATS'14, SIOPT'15,'16)

一次最適化法による統一的な解法の提案
(JMLR'17)

まとめ

- 現代社会は変化のスピードが速く，不確実性への対応は必須
- min-maxの定式化は機械学習分野で活発な研究対象 (e.g., adversarial attack)
- ロバスト最適化は比較的新しい分野
  Ben-Tal, El Ghaoui & Nemirovski ['09]
- 分野を超えてロバスト最適化研究が広がっている
  - ✓上記の本の冒頭で
    ロバスト制御 (H_∞制御)、ロバスト統計、機械学習（SVM）
    などについて言及あり
- 機械学習分野でもロバスト最適化が使われている