ロバストネスを探る
What is Robustness?

京都大学　　永原　正章
Masaaki Nagahara
Kyoto University

Abstract　Robustness is a key issue in engineering. This article is intended as an investigation of common mathematical structures in robustness in various fields. In particular, we point out a common structure between robust control, regularization, and Bayes estimation.

1　はじめに

ロバストネスという用語は，数理工学のさまざまな研究分野で用いられている．大規模化し複雑化した現代のシステムを予測したり制御するときに，ロバストネスはもっとも重要なキーワードの一つである．たとえば制御工学では，複雑なシステムにおけるモデル化できないダイナミクスを不確かさととらえ，その不確かさをシステムの集合として定式化するロバスト制御が盛んに議論されている [9]．また，統計的学習においては，教師なし学習におけるロバスト統計や教師あり学習における汎化能力の問題はロバストネスと関連がある [8]．他にもディジタル通信における誤り訂正符号の議論 [4] やシステムバイオロジー [6]，複雑ネットワーク [2] などにおけるロバストネス研究など幅広い分野で多くの研究が行われている．

しかし，それらのロバストネスの数理的な共通点はあまり詳しく調べられていない．そこで本稿では，ロバストネスという概念の裏に潜む数理的構造を解明するための足がかりとして，制御工学の立場から統計的学習の問題を眺め，ロバスト制御と正則化法，そして Bayes 推定の問題の定式化における数理的な共通点を明らかにする．なお，統計的学習の立場からのロバストネスの考察，およびロバスト制御との関連については [8] が参考になる．

本稿の構成は以下のとおりである．第 2 章で，ロバスト制御の基礎的な事項について復習し，ノミナル性能の向上とロバスト安定化を達成するロバスト制御の問題（$\mathcal{H}^\infty$ 混合感度問題）が 2 ブロックの $\mathcal{H}^\infty$ 最適制御問題で記述されることを述べる．第 3 章では，統計的学習における正則化と Bayes 推定の問題の定式化について述べ，それらが 2 ブロックのベクトルノルムの最適化問題で記述されることを述べる．第 4 章でそれらの関連について述べる．

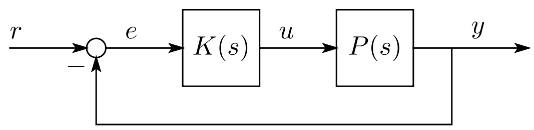

図 1: 単一フィードバック系：$P(s)$, $K(s)$ はそれぞれ制御対象および制御器の伝達関数を表す.

2　ロバスト制御

2.1　感度最適化問題

制御対象として，次の線形システムを考える．

$$\begin{cases} \dot{\boldsymbol{x}}(t) = A\boldsymbol{x}(t) + \boldsymbol{b}u(t), \\ y(t) = \boldsymbol{c}^\top \boldsymbol{x}(t), \quad t \in [0,\infty). \end{cases} \tag{1}$$

ここで，$\boldsymbol{x}(t) \in \mathbb{R}^\nu$ をシステムの状態，$u(t) \in \mathbb{R}$ をシステムへの入力，$y(t) \in \mathbb{R}$ をシステムの出力とし，$A \in \mathbb{R}^{\nu\times\nu}$, $\boldsymbol{b} \in \mathbb{R}^\nu$, $\boldsymbol{c} \in \mathbb{R}^\nu$ は与えられているとする．このシステムの入出力伝達関数を $P(s)$ とおく．すなわち，$P(s) = \boldsymbol{c}^\top(sI-A)^{-1}\boldsymbol{b}$ とする．このシステムに対して，図 1 の単一フィードバック系を考え，参照信号 r に出力 y を追従させる制御を考える．すなわち，追従誤差 $e = r - y$ をなるべく小さくする制御を考えたい．これは，参照信号 r から追従誤差 e への伝達関数（感度関数と呼ばれる）

$$S(s) = \frac{1}{1+P(s)K(s)}$$

のゲイン $|S(\mathrm{j}\omega)|$ を小さくすることにより達成される．単純に考えれば，制御器 $K(s)$ のゲインを大きくすれば，感度関数 $S(s)$ のゲインはいくらでも小さくできる．しかし，後で述べるロバスト性との関連から，感度関数 $S(s)$ のゲインは全ての周波数で小さくするのではなく，安定な伝達関数（重み関数） $W_1(s)$ のゲイン $|W_1(\mathrm{j}\omega)|$ で定められる周波数帯で感度関数を小さくすることにし

たい[1]．この問題を $\mathcal{H}^\infty$ ノルムにより定式化すると，

$$\|W_1 S\|_\infty = \sup_{\omega\in\mathbb{R}} |W_1(\mathrm{j}\omega)S(\mathrm{j}\omega)|$$

を最小化する安定化制御器 K を求める問題となる．これを感度最適化問題と呼ぶ．ここで簡単のため $P(s)$ は安定と仮定すると，図 1 に対する安定化制御器全体は次式で表される [3, Chap. 5]．

$$\mathcal{K} := \left\{ K(s) = \frac{Q(s)}{1-P(s)Q(s)} : \ Q \in \mathcal{H}^\infty \right\}. \tag{2}$$

また，$K(s) \in \mathcal{K}$ とすれば，簡単な計算により $S(s) = 1 - P(s)Q(s)$ が示される．したがって，感度最適化問題は次の評価関数

$$J_1(Q) = \|W_1(1-PQ)\|_\infty \tag{3}$$

を最小化する $Q \in \mathcal{H}^\infty$ を求める問題となる．

2.2 不確かさの表現

上記の感度最適化問題では，モデル (1) によって現実の制御対象が寸分の狂いもなく表現されていると仮定する．しかし，現実のシステムには例えば以下のような不確かさが存在する．

- パラメータ公称値のばらつき（たとえば，抵抗値など．パーツの経年変化も含む）
- 高周波歪み（たとえば，高次共振モードなど）
- 未知外乱の存在（たとえば，電磁波など）
- 非線形性

これらのばらつきやダイナミクスは，モデリングの段階で同定できなかったり，もしくは (1) のようには簡単に表現できなかったり，表現できても非常に高次になったりするので，非常に扱いにくい．

ロバスト制御では，これらの不確かさを，伝達関数の集合として記述する [9]．具体的には，

1. ノミナルモデル
2. 不確かさを表現する摂動の構造
3. 摂動の最大値（上限値）

はあらかじめわかっているものとして，不確かさを記述する [5, 2.1 節]．たとえば，ノミナルモデルとして A, $\boldsymbol{b}$,

[1] 応用上は追従性能は低周波でより良くさせたいことから，$W_1(s)$ はローパスにとることが多い．

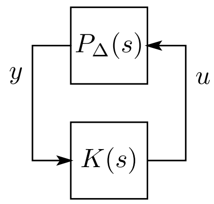

図 2: フィードバック系: $P_\Delta(s) \in \mathcal{P}$, $K(s)$ はそれぞれ制御対象および制御器の伝達関数を表す．

$\boldsymbol{c}$ を固定した (1) の伝達関数 $P(s) = \boldsymbol{c}^\top (sI - A)^{-1}\boldsymbol{b}$ を与え，摂動の構造として次の乗法的摂動を仮定する．

$$P_\Delta(s) = (1 + W_2(s)\Delta(s))P(s).$$

ここで，$P_\Delta(s)$ は $\Delta(s) \in \mathcal{H}^\infty$ によって摂動を受けた制御対象であり，$W_2(s)$ は安定な伝達関数（重み関数）である[2]．また $\Delta(s)$ は，その最大値が $\mathcal{H}^\infty$ ノルムを用いて

$$\|\Delta\|_\infty < \delta^{-1}$$

で制約されているとする．これらを用いて，伝達関数の集合を

$$\mathcal{P} := \left\{ P_\Delta(s) = (1 + W_2(s)\Delta(s))P(s) : \|\Delta\|_\infty < \delta^{-1} \right\} \tag{4}$$

と表現する．これがロバスト制御における不確かさの一つのモデルである．

2.3 ロバスト安定化

評価関数の集合 (4) でモデル化された不確かさに対して，ロバスト制御ではまず安定化が重要である．すなわち，(4) で表される集合の要素すべてに対して，図 2 のフィードバック系が安定となる制御器 $K(s)$ を特徴づける．これをロバスト安定化と呼ぶ．安定化制御器のパラメトリゼーション (2) およびスモールゲイン定理 [9, Chap. 9] を用いれば，(4) の不確かさに対して図 2 のフィードバック系がロバスト安定であるための必要十分条件は，

$$K(s) \in \mathcal{K} := \left\{ K(s) = \frac{Q(s)}{1-P(s)Q(s)} : \ Q \in \mathcal{H}^\infty,\ \ \|W_2 PQ\|_\infty \le \delta \right\} \tag{5}$$

であることがわかる．ここで，

$$J_2(Q) = \|W_2 PQ\|_\infty \tag{6}$$

[2] 通常，高周波で不確かさは大きくなるので，$W_2(s)$ はハイパスにすることが多い．

を小さくすればするほど，δ は小さくとれるので，不確かさの集合 $\mathcal{P}$ は大きくなる．すなわち，上記の評価関数 $J_2(Q)$ は小さければ小さいほど，大きな変動に対しても安定となり，ロバストネスは高まる．ここで，$J_2(Q)$ を最小化する Q は $Q(s)=0$ であり，これはすなわち $K(s)=0$ であるので，明らかに意味のない解であることに注意する．

2.4　$\mathcal{H}^\infty$ 混合感度問題

感度最適化問題では，評価関数 (3) を小さくすることで（ノミナルの）追従性能を良くすることが目的であった．いっぽう，ロバスト安定化では，評価関数 (6) を小さくすることで (4) で定義される摂動プラントの集合が大きくなり，ロバストネスが向上する．しかし，両者を同時に小さくするような $Q\in\mathcal{H}^\infty$ は存在しない．たとえば，$Q(s)$ をノミナル制御対象 $P(s)$ の逆システムに近づけることにより，(3) の $\mathcal{H}^\infty$ ノルムはいくらでも小さくなるが，同時に (6) で定義される $J_2(Q)$ は $\|W_2\|_\infty$ に近づくだけで小さくはならない．また，$Q(s)$ を 0 に近づけることにより $J_2(Q)$ はいくらでも小さくなるものの，$J_1(Q)$ は $\|W_1\|_\infty$ に近づくだけで，小さくはできない．すなわち，ノミナルの性能とロバストネスはトレードオフの関係にあることがわかる．このトレードオフを調整するのが，重み関数 $W_1(s)$ と $W_2(s)$ である．前で述べたように，$W_1(s)$ はローパスに，いっぽう $W_2(s)$ はハイパスにとることが多い．ノミナル性能とロバストネスのトレードオフを考えると，この重み関数の選び方は妥当である．この重み関数のもとで

$$J(Q)=\left\|\begin{bmatrix}W_1(1-PQ)\\ W_2PQ\end{bmatrix}\right\|_\infty \qquad (7)$$

を最小化すれば，感度関数 $1-P(s)Q(s)$ は低周波で小さくなり，相補感度関数 $P(s)Q(s)$ は高周波で小さくなることが期待できる．これを $\mathcal{H}^\infty$ 混合感度問題と呼ぶ．数値計算により，評価関数 (7) の最小値（下限値）に任意に近い $Q(s)$ を求めることができる [9]．なお，(7) を最小化することは，ロバスト性能問題（ロバスト安定化し，かつ集合 $\mathcal{P}$ の伝達関数のうち $W_1(s)(1-P_\Delta(s)Q(s))$ の $\mathcal{H}^\infty$ ノルムを最大化するものに関して感度を最適化する問題）における評価関数のある上界を最小化することが知られており [9]，$\mathcal{H}^\infty$ 混合感度問題によって，制御対象に変動があっても追従性能に変動があまりない制御系が得られる．

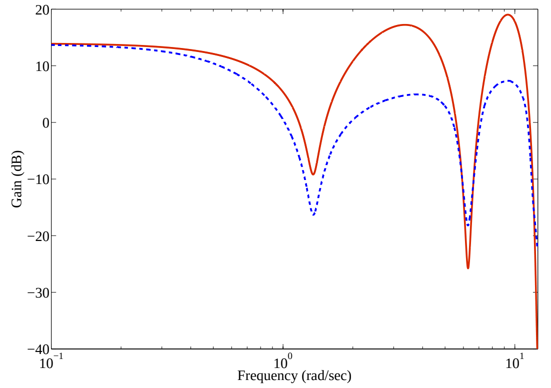

図 3: ディジタル繰り返し制御器，ノミナルモデルに対する制御器（実線）とロバスト制御器（破線）．

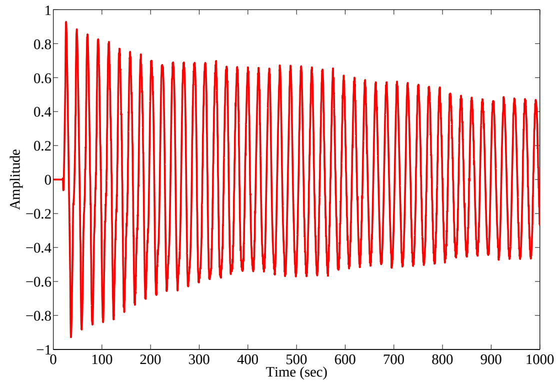

図 4: ノミナルモデルに対する制御器による追従誤差

2.5　ロバスト制御の例題

ここでは，筆者らが最近提案しているディジタル繰り返し制御系におけるロバスト制御設計 [7] を例題に，$\mathcal{H}^\infty$ 混合感度問題 (7) の解の性質を確認する．この設計問題において，感度最適化問題，すなわち (3) の評価関数 $J_1(s)$ だけを最適化する $Q(s)$ と $\mathcal{H}^\infty$ 混合感度問題，すなわち (7) の評価関数 $J(s)$ を最適化する $Q(s)$ のゲインの周波数応答を図 3 に示す．この図から，ロバスト制御器は，単に感度関数 $J_1(Q)$ を最適化した場合よりもゲインが低くなっていることがわかる．特にこのロバスト制御器は，高周波で不確かさをより大きく見積ったため，高周波でゲインが落ちていることに注意する．このロバスト制御器の効果は，実際に制御対象に変動があった場合にあらわれる． 図 4 はノミナルモデルに対する制御器で，制御対象に変動があったときの追従誤差を示す．この制御器は，制御対象の変動を想定していないの

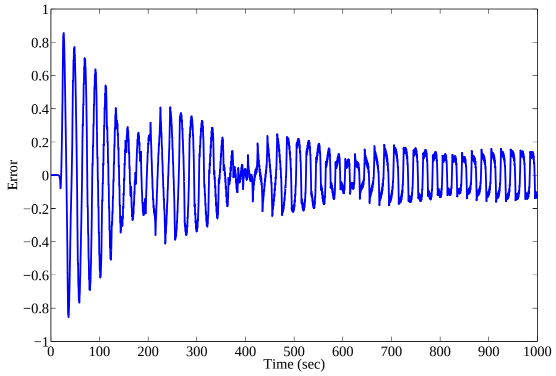

図 5: ロバスト制御器による追従誤差

で，実際に変動が起こると誤差が大きくなる．いっぽう，図5はロバスト設計をした場合の制御系の追従誤差である．図4に比べて誤差が小さくなっており，ロバスト設計の効果があらわれている．もちろん，制御対象に変動がない場合は，ノミナルモデルに対する制御器のほうがロバスト制御器よりも良い性能を示す．すなわち，ロバスト制御器はノミナル性能を犠牲にしてロバストネスを獲得しており，ノミナル性能とロバストネスのトレードオフを示している．

以上のような性質は，最小二乗推定問題における正則化法および機械学習における Bayes 推定でもあらわれる．これを次節で示す．

3 正則化法と Bayes 推定

次の最小二乗推定問題を考える．N 個の訓練データの組 $\{(x_1,y_1),(x_2,y_2),\ldots,(x_N,y_N)\}$ が与えられたとき，次の二乗誤差

$$J(\boldsymbol{w})=\sum_{n=0}^{N}\left|y_n-\boldsymbol{w}^\top\boldsymbol{\phi}(x_n)\right|^2 \tag{8}$$

を最小化するパラメータ $\boldsymbol{w}=[w_0,w_1,\ldots,w_M]^\top$ を求める．ただし，$\boldsymbol{\phi}(x)=[1,x,x^2,\ldots,x^M]^\top$ である．ここで，$\boldsymbol{y}=[y_1,y_2,\ldots,y_N]^\top$ とおき，

$$\Phi=\begin{bmatrix}1 & x_1 & \ldots & x_1^M\\ 1 & x_2 & \ldots & x_2^M\\ \vdots & \vdots & \ddots & \vdots\\ 1 & x_N & \ldots & x_N^M\end{bmatrix}$$

とおくと，(8) の二乗誤差 $J(\boldsymbol{w})$ は次のように表すことができる．

$$J(\boldsymbol{w})=\|\boldsymbol{y}-\Phi\boldsymbol{w}\|_2^2.$$

この評価関数を最小化する $\boldsymbol{w}$ は次で与えられる [1]．

$$\boldsymbol{w}_{\rm opt}=(\Phi^\top\Phi)^{-1}\Phi^\top\boldsymbol{y}.$$

この最適パラメータによる関数 $\boldsymbol{w}_{\rm opt}^\top\boldsymbol{\phi}(x)$ は，データ数 N が小さく，パラメータ数 M が大きいとき，データの間で振動する傾向があることが知られている．これを過学習と呼ぶ [1]．これはデータ $\boldsymbol{y}$ の微小な変化に対し，係数 $\boldsymbol{w}_{\rm opt}$ が大きく変化することが原因であり，これを抑えるために (8) の評価関数の代わりに次の評価関数を最小化することを考える．

$$\begin{aligned}J_{\rm reg}(\boldsymbol{w})&=\sum_{n=0}^{N}\left|y_n-\boldsymbol{w}^\top\boldsymbol{\phi}(x_n)\right|^2+\lambda^2\boldsymbol{w}^\top\boldsymbol{w}\\&=\left\|\begin{bmatrix}\boldsymbol{y}-\Phi\boldsymbol{w}\\ \lambda\boldsymbol{w}\end{bmatrix}\right\|_2^2\end{aligned} \tag{9}$$

これを正則化最小二乗法（またはリッジ回帰）と呼ぶ．この評価関数において，λ は正則化係数と呼ばれ，データのフィッティング性能（第一項の抑制）とパラメータ $\boldsymbol{w}$ の抑制のトレードオフを制御する係数である．この評価関数を最小化する $\boldsymbol{w}$ は次式で与えられる [1]．

$$\boldsymbol{w}_{\rm reg}=(\lambda I+\Phi^\top\Phi)^{-1}\Phi^\top\boldsymbol{y}.$$

この式より，λ を大きくすれば $\boldsymbol{w}_{\rm reg}$ の 2 ノルムが減少することが容易にわかる．

この正則化最小二乗法は Bayes 推定と関連が深いことが知られている．実際，データ $\{(x_1,y_1),\ldots,(x_N,y_N)\}$ は，関数 $y=\boldsymbol{w}^\top\boldsymbol{\phi}(x)$ に平均 0，分散 σ^2 の Gauss 雑音が加わったもの，すなわち，

$$p(y|x,\boldsymbol{w})=\mathbf{N}(y|\boldsymbol{w}^\top\boldsymbol{\phi}(x),\sigma^2)$$

と仮定し，パラメータ $\boldsymbol{w}$ の事前分布として，同じく Gauss 分布

$$p(\boldsymbol{w}|\tau)=\mathbf{N}(\boldsymbol{w}|0,\tau^2)$$

を仮定すると，Bayes の定理より事後分布は

$$\begin{aligned}p(\boldsymbol{w}|\boldsymbol{y},\tau,\sigma)&\propto p(\boldsymbol{y}|w,\sigma)p(\boldsymbol{w}|\tau)\\&=\prod_{n=1}^{N}\mathbf{N}(y_n|\boldsymbol{w}^\top\boldsymbol{\phi}(x_n),\sigma^2)\mathbf{N}(\boldsymbol{w}|\tau)\end{aligned}$$

となる．この事後分布を最大化する $\boldsymbol{w}$ を最大事後確率推定，または MAP(maximum posterior) 推定とよぶ．Gauss 分布の性質から，この $\boldsymbol{w}$ は評価関数

$$\begin{aligned}J_{\rm MAP}(\boldsymbol{w})&=\sum_{n=1}^{N}\left|y_n-\boldsymbol{w}^\top\boldsymbol{\phi}(x_n)\right|^2+\frac{\sigma^2}{\tau^2}\boldsymbol{w}^\top\boldsymbol{w}\\&=\left\|\begin{bmatrix}\boldsymbol{y}-\Phi\boldsymbol{w}\\ \frac{\sigma}{\tau}\boldsymbol{w}\end{bmatrix}\right\|_2^2\end{aligned} \tag{10}$$

を最小化する $\boldsymbol{w}$（これを $\boldsymbol{w}_{\mathrm{MAP}}$ とおく）と一致する [1]．上記の評価関数と (9) を比べると，MAP 推定によるパラメータ $\boldsymbol{w}_{\mathrm{MAP}}$ は，$\lambda = \sigma/\tau$ とおいたときの正則化最小二乗法の解と一致することがわかる．

4 ロバスト制御と正則化法・Bayes 推定の共通点

$\mathcal{H}^\infty$ 混合感度問題の評価関数 (7) と正則化法および Bayes 推定における評価関数 (9), (10) は，形式だけを見れば 2 ブロックの伝達関数またはベクトルのノルムの最適化を行っており，非常に良く似ている．しかし，形式だけでなく，これらの評価関数はノミナル性能（感度最適化，または訓練データへのフィッティング性能）とロバストネス（混合感度最適化，または汎化能力）とのトレードオフを考慮した最適化となっており，概念的にも共通点があることがわかる．

5 おわりに

本稿では，ロバスト制御と正則化法，Bayes 推定の定式化に共通点があることを指摘した．これを足がかりとして，それぞれの分野におけるロバストネスの数理的な構造の解明やそれぞれの分野における数理的手法（例えば $\mathcal{H}^\infty$ 最適化における数値最適化手法や統計的学習における確率的アルゴリズムなど）を取り入れた新しい設計法の確立は今後の課題である．

参考文献

[1] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer (2006)

[2] J. C. Doyle et al., The “robust yet fragile” nature of the Internet, *Proc. of Natl. Acad. Sci. USA*, vol. 102, no. 41, pp. 14497–14502 (2005)

[3] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, *Feedback Control Theory*, Macmillan (1992)

[4] G. A. Jones and J. M. Jones, *Information and Coding Theory*, Springer (2000)

[5] 木村，藤井，森，ロバスト制御，コロナ社 (1994)

[6] H. Kitano, Towards a theory of biological robustness, *Molecular Systems Biology*, 3:137 (2007)

[7] M. Nagahara and Y. Yamamoto, Robust repetitive control by sampled-data H^∞ filters, *Proc. of IEEE Conf. on Decision and Control* (2009)

[8] 田中，統計的学習における「ロバストネス」，第 52 回自動制御連合講演会 (2009)

[9] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*, Prentice Hall (1996)