様　式　Ｃ－１９、Ｆ－１９、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

平成 ２８ 年　 ５ 月 ２７ 日現在

機関番号：８２６７０
研究種目：若手研究(B)
研究期間：2014～2015
課題番号：２６８２０１７４

研究課題名（和文）逐次型LASSOとその設計手法の開発

研究課題名（英文）Development of Recursive LASSO and Its Design Method

研究代表者
金田　泰昌（Kaneda, Yasuaki）

地方独立行政法人東京都立産業技術研究センター・開発本部開発第一部情報技術グループ・主任研究員

研究者番号：２０４６３０１０

交付決定額（研究期間全体）：（直接経費）　　1,400,000円

研究成果の概要（和文）：L1正則化付き線形回帰問題に対して，その正則化パラメータが観測ノイズの二次モーメントと関連していることを示した．また，その関係性から正則化パラメータのシステマチックな設計手法を提案した．さらに，L1正則化付き線形回帰問題の近似解を導出し，その解が解析的かつ逐次型に計算できることを示した．数値シミュレーションにて提案手法の有効性を示した．

研究成果の概要（英文）：For L1 regularized linear regression problem, we show that a regularization parameter of the problem is related to a second moment of measurement noise and propose a systematic design method of the parameter using the relation. Moreover, we derive an approximated solution of the problem, which can be solved analytically and recursively. Numerical simulations demonstrate effectiveness of the proposed methods.

研究分野：制御工学

キーワード：スパース推定　L1正則化　正則化パラメータ　逐次アルゴリズム

１．研究開始当初の背景

近年，様々な分野においてスパース推定が注目を集めている．例えば，ネットワークにおけるリンクの推定や，パラメータ推定の際のパラメータ選択においてスパース推定が適用されている．制御工学の分野も例外ではなく，しばしばスパース推定が使われることがある．一例として，モデル予測制御における入力の削減手法としてもスパース推定が用いられている．さらには，我々はスパース推定をカルマンフィルタに応用することで，外れ値ノイズにロバストなロバストカルマンフィルタを提案している．

スパース推定手法の中でも，LASSO（Least Absolute Shrinkage and Selection Operator）と呼ばれる手法が注目を集めている．この手法は最適化問題に対して L1 正則化を適用することでスパースな解が得られる手法である．特に，評価関数が凸関数の場合，凸最適化問題になるため，取り扱いの容易さから様々な応用がなされている．しかしながら，以下の 2 つの問題が未だ存在し，特に制御工学の分野のようにオンラインでデータを処理する際に，LASSO の適用を妨げている．

一つ目の問題として，正則化パラメータの決定が試行錯誤的であり，設計の妥当性が保証できないという問題が挙げられる．そこで，正則化パラメータの設計に妥当性を与えるために，これまでに一般的なモデル選択指標を用いてクロスバリデーションで正則化パラメータを決定する方法や，L1 正則化に特化したモデル選択指標などが提案されている．しかしながら，これらは正則化パラメータの候補を複数用意しておき，全ての候補に対してある指標を計算する．そして，その指標を基に候補の中からパラメータを選択する．そのため，予め用意した候補の範囲内でしか妥当性は得られないことになる．また，正則化パラメータの候補数に比例して計算回数が増加するため，パラメータの候補が多い場合は計算時間が長くなり，通常オフラインで事前に計算する．しかしながら，前述の通り制御工学の分野ではオンラインでデータを処理することが多く，短いサンプリング周期内で計算する必要がある．そのような分野での応用を考えた場合，正則化パラメータの候補を複数計算することは非効率的で，事実上困難となる．

二つ目の問題として，データ数が多くなると大量のデータ処理や大規模な逆行列演算が必要となり，計算コストが高くなることが挙げられる．前述の通りスパース推定を制御分野に応用しようとした際，短いサンプリング周期内で計算できるアルゴリズムが求められる．そのため，一度に大量のデータ処理が必要なアルゴリズムは計算コストの点で実用的ではない．また，近年ではビックデータの活用が注目されているが，観測データの肥大化に伴い，メモリの観点から観測データを一括処理するのではなく，逐次的に処理する方法が望まれている．

２．研究の目的

１．で示した背景を受け，本研究では，正則化パラメータの試行錯誤設計を必要とせず，かつ逐次的に計算できる LASSO のアルゴリズムを開発することを目的とする．具体的には以下の 2 つを目的とする．

(1) LASSO における正則化パラメータのシステマチックな設計手法を明らかにする．

(2) LASSO の解を 1 ステップ前の解から逐次的に計算可能な手法を明らかにする．

３．研究の方法

２．で示した 2 つの目的に対して，それぞれ数学的にアルゴリズムを導出し，導出したアルゴリズムの有効性を数値シミュレーションで評価する．具体的には以下の(1)および(2)の通りである．

(1) 正則化パラメータの設計アルゴリズムの導出方法を①に，その評価方法を②にそれぞれ示す．

①正則化パラメータの設計アルゴリズムを数学的に導出するために，まず本研究で対象とする LASSO を定式化する．今，N 番目までの観測値 $\boldsymbol{y}_k \in \mathbb{R}^m$ (k = 1, …, N) が与えられたとする．さらに，観測値 $\boldsymbol{y}_k$ に関して次式が成り立つとする．

$$\boldsymbol{y}_k = \boldsymbol{\phi}_k^T \boldsymbol{\theta} + \boldsymbol{v}_k$$

ここで，$\boldsymbol{\phi}_k \in \mathbb{R}^{n\times m}$ はレグレッサであり，$\boldsymbol{\theta} \in \mathbb{R}^n$ は推定したい未知パラメータベクトルである．また，$\boldsymbol{v}_k \in \mathbb{R}^m$ は平均値ゼロ，共分散行列 $\boldsymbol{R} \in \mathbb{R}^{m\times m}$ の観測ガウスノイズである．このとき，LASSO は次式として定式化される．

$$\min_{\boldsymbol{\theta}} J(\boldsymbol{\theta})$$

$$J(\boldsymbol{\theta}) = \frac{1}{N}\sum_{k=1}^{N} \left|\left|\boldsymbol{y}_k - \boldsymbol{\phi}_k^T\boldsymbol{\theta}\right|\right|^2_{\boldsymbol{R}^{-1}} + \lambda||\boldsymbol{\theta}||_1$$

ここで，λ は正則化パラメータである．本研究では，既知のパラメータ（$\boldsymbol{\phi}_k$ や $\boldsymbol{R}$）から正則化パラメータ λ を決定するアルゴリズムを数学的に導出する．そのために，まずは既知のパラメータと正則化パラメータとの関係性を数学的に解析する．そして，その解析結果を基に，正則化パラメータを自動的に計算するアルゴリズムを導出する．

②数値シミュレーションを用いて従来手法と提案手法を比較する．従来手法では，推定誤差が最も小さくなるように試行錯誤的に正則化パラメータを設計する．そして，両手法により得られる推定精度およびスパース度（ベクトル内に含まれるゼロの要素の割

合）を比較する．評価に用いるシステムはランダムに生成する．具体的には，システムのパラメータはガウス分布で生成し，パラメータの真値は非ゼロの値が全体の約 10%（スパース度が約 90%）となるようにベルヌーイ分布により生成する．

(2) 逐次的に解が計算可能なアルゴリズムの導出方法を①に，その評価方法を②にそれぞれ示す．

① 3．(1) ①で定式化された最適化問題を直接解くのではなく，近似的に解けるように問題を緩和する．そして導出された緩和問題から逐次的に計算できるアルゴリズムを導出する．

② 3．(1) ②と同様に，ランダムに生成されたシステムに対し，従来手法と提案手法を数値シミュレーションにて比較する．そして，それらの収束性を評価する．なお，本研究では，従来手法として近接勾配法を用いた手法である RDA (Regularized Dual Averaging). および AdaGrad (Adaptive Subgradient Method)を採用する．また，近接勾配法のアイディアを参考に，近接点法と準ニュートン法を組み合わせた手法（Prox + Newton）も比較対象とする．

４．研究成果

3．で示した方法で得られた成果は以下の(1)および(2)の通りである．

(1)正則化パラメータの設計アルゴリズムの導出に関する成果を①に，その評価結果を②に示す．

①既知パラメータと正則化パラメータの関係性を数学的に解析した結果，正則化パラメータと既知パラメータとの間に次式の不等式が成り立つことを示した．

$$|\nu_i| \leq \lambda$$

ただし，

$$\boldsymbol{\nu} = -\frac{2}{N}\boldsymbol{P}_N^{-1}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_N^{\mathrm{LS}}\right)$$

$$\boldsymbol{\theta}_N^{\mathrm{LS}} = \boldsymbol{P}_N \sum_{k=1}^{N} \boldsymbol{\phi}_k \boldsymbol{R}^{-1} \boldsymbol{y}_k$$

$$\boldsymbol{P}_N^{-1} = \sum_{k=1}^{N} \boldsymbol{\phi}_k \boldsymbol{R}^{-1} \boldsymbol{\phi}_k^T$$

であり，また$\hat{\boldsymbol{\theta}}$ は LASSO による解である．この不等式は，最小二乗解$\boldsymbol{\theta}_N^{\mathrm{LS}}$ と LASSO による解$\hat{\boldsymbol{\theta}}$ とのある種の差分$\boldsymbol{\nu}$を考えた時，その（絶対値の）上限がλで与えられることを意味する．言い換えると，LASSO は与えられた上限λの範囲内に差分$\boldsymbol{\nu}$が収まるように計算する手法であることを上記不等式は示している．そして，上限λを小さくすることで$\hat{\boldsymbol{\theta}}$ は$\boldsymbol{\theta}_N^{\mathrm{LS}}$に近づき，逆に上限$\lambda$を大きくすることで$\hat{\boldsymbol{\theta}}$ は$\boldsymbol{\theta}_N^{\mathrm{LS}}$から遠ざかり，よりスパースになっていく．これは$\lambda$を通常の正則化パラメータとしてとらえた場合と当然ながら一致する．ここで，上記解析結果を逆の視点で捉えると，差分$\boldsymbol{\nu}$の範囲が予め見積もることができれば，正則化パラメータが設計できることになる．そこで本研究では，最小二乗解の誤差の範囲内に真値があると考えることで差分$\boldsymbol{\nu}$の範囲を見積もり，最終的に正則化パラメータを次式で与えることとした．

$$\lambda = \frac{2}{N\sqrt{\eta^{\min}}}$$

ただし，$\eta^{\min}$は$\boldsymbol{P}_N$の最小固有値である．

②数値シミュレーションを用いた設計アルゴリズムの評価結果として，パラメータの推定誤差（RMSE）を図 1 に，パラメータのスパース度を図 2 にそれぞれ示す．なお，青△は提案手法を用いた場合の結果を，赤＋は試行錯誤設計をした際の結果をそれぞれ示す．

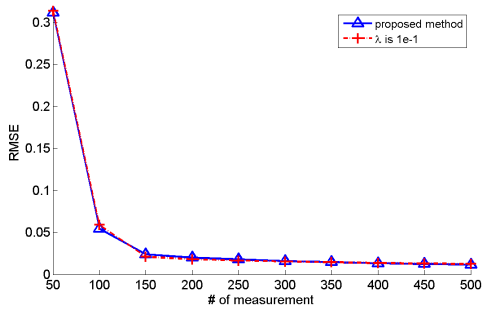

図 1　推定誤差（RMSE）

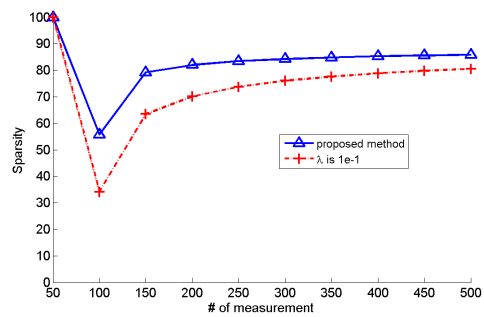

図 2　スパース度

試行錯誤設計では，複数の正則化パラメータの候補を試し，本結果では推定誤差が小さく，かつスパース度が高かった正則化パラメータとして 10^{-3} を採用している．これらの結果から分かるように，提案設計手法を用いることで，正則化パラメータを探索することなく，推定誤差の小さい結果が得られていることが分かる．また，提案手法を用いることで適切なスパース度が試行錯誤無く得られていることが分かる．

(2)逐次計算アルゴリズムの導出に関する成果を①に，その評価結果を②に示す．

①逐次計算アルゴリズムを導出するために，評価関数の上下限を導出し，緩和問題としてその上下限の最小化問題を採用した．そして，緩和問題の解は近接点法と最小二乗法を用いて解析的に計算できることを示した．具体的には，緩和問題の解は次式で与えられる．

$$\theta_i^{\min} = \mathrm{ST}_{\frac{N\lambda\eta^{\max}}{2}}(\theta_i^{\mathrm{LS}})$$

$$\theta_i^{\max} = \mathrm{ST}_{\frac{N\lambda\eta^{\min}}{2}}(\theta_i^{\mathrm{LS}})$$

ただし，$\theta_i^{\min}$および$\theta_i^{\max}$は評価関数の下限および上限の最小解を表す．また，$\eta^{\max}$は$\boldsymbol{P}_N$の最大固有値を表し，$\mathrm{ST}_\alpha(x)$はSoft-Thresholding関数と呼ばれ次式で与えられる．

$$\mathrm{ST}_\alpha(x) = \begin{cases} x-\alpha & x > \alpha \\ 0 & -\alpha \le x \le \alpha \\ x+\alpha & x < -\alpha \end{cases}$$

上記緩和問題では，解の候補が二つ存在する．一方，評価関数の近似誤差は，$\boldsymbol{P}_N - \eta^{\min}$および$\eta^{\max} - \boldsymbol{P}_N$で表される．そこで，本研究では，これらのノルムを評価し，ノルムが小さいときの解を最終的な近似解として採用する．ノルムの計算方法は種々存在するが，本研究では計算量の観点からフロベニウスノルムを採用する．また，上記緩和問題の解は最小二乗解$\boldsymbol{\theta}_N^{\mathrm{LS}}$を用いている．そこで，本研究ではこれを逐次最小二乗解に置き換えることで逐次的に緩和問題を解くアルゴリズムを導出した．

②数値シミュレーションを用いた逐次計算方法の評価結果として，$\lambda=10^{-3}$の場合のパラメータの推定誤差（RMSE）を図3に，パラメータのスパース度を図4にそれぞれ示す．なお，青△はRDAの結果を，緑◯はAdaGradの結果を，赤＋はProx + Newtonの結果を，薄青□は提案手法の結果をそれぞれ示す．これらの結果より，提案手法の推定誤差の収束性が最も良いことが分かる．また，提案手法により，適切なスパース度が得られていることが分かる．

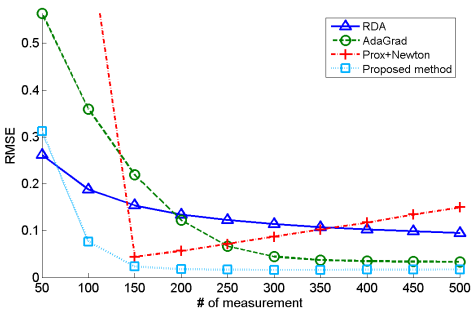

図3　$\lambda=10^{-3}$の場合の推定誤差（RMSE）

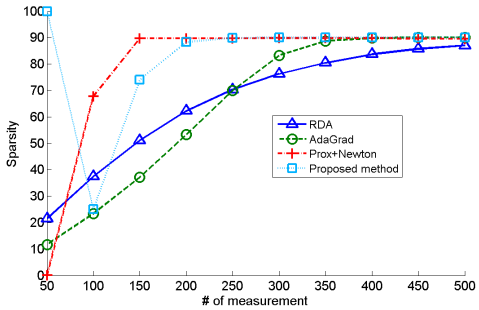

図4　$\lambda=10^{-3}$の場合のスパース度

5．主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計２件）

① 金田泰昌，入月康晴，「LASSOの逐次型アルゴリズム」，電気学会論文誌C編，Vol. 136，No. 7（2016）（to appear）

② 金田泰昌，入月康晴，「統計量に基づくL1最小化問題のパラメータ設計手法」，電気学会論文誌C編，Vol. 135，No. 11，pp. 1419-1426（2015）
DOI: 10.1541/ieejeiss.135.1419

〔学会発表〕（計４件）

① 金田泰昌，入月康晴，「逐次型LASSOとその設計手法」，第60回システム制御情報学会研究発表講演会（2016）

② 金田泰昌，入月康晴，「L1正則化を用いたロバストカルマンフィルタとその設計手法」，第58回自動制御連合講演会（2015）

③ 金田泰昌，入月康晴，「L1正則化付き線形回帰の逐次アルゴリズム」，平成27年度電気学会電子・情報・システム部門大会（2015）

④ 金田泰昌，入月康晴，「統計量に基づくLASSOの正則化パラメータの設計手法」，電子情報通信学会総合大会（2015）

6．研究組織

(1) 研究代表者

金田　泰昌（KANEDA, Yasuaki）
東京都立産業技術研究センター・
開発本部開発第一部情報技術グループ・
主任研究員
研究者番号：20463010