

[招待講演] 線形逆問題に対する凸最適化問題の特性解析とその応用

早川 諒†

† 大阪大学大学院基礎工学研究科, 〒560-8531 大阪府豊中市待兼山町 1-3

E-mail: †rhayakawa@sys.es.osaka-u.ac.jp

あらまし 未知のベクトルをその線形観測から推定する問題は線形逆問題と呼ばれ、信号処理における様々な応用で現れる。線形逆問題のための有力なアプローチの1つである凸最適化では、適切な正則化項を目的関数に含めることで、未知ベクトルに関する事前情報を活用した推定が可能である。凸最適化に基づくアプローチに関しては様々な理論的解析がなされており、中でもCGMT (Convex Gaussian Min-max Theorem) と呼ばれる定理を用いると、様々な凸最適化問題によって得られる推定値の漸近的な特性を解析できる。本稿では、 ℓ_1 正則化を用いた最適化問題を題材として、CGMTを用いた解析方法の概要や解析結果について説明する。また、CGMTに基づく理論解析の応用例として、未知の雑音分散を推定する手法や、線形逆問題のためのADMM (Alternating Direction Method of Multipliers) に関する特性解析について述べる。

キーワード 線形逆問題, 圧縮センシング, 凸最適化, 理論解析

[Invited Talk] Performance Analysis of Convex Optimization Problem for Linear Inverse Problems and Its Applications

Ryo HAYAKAWA†

† Graduate School of Engineering Science, Osaka University, 1-3, Machikaneyama, Toyonaka, Osaka, 560-8531 Japan

E-mail: †rhayakawa@sys.es.osaka-u.ac.jp

Abstract Linear inverse problems, i.e., the estimation of an unknown vector from its linear measurements, often appear in the field of signal processing. Convex optimization is a powerful approach for linear inverse problems because we can utilize the prior knowledge of the unknown vector by using a proper regularization in the objective function. There are various analyses for such convex optimization-based approaches, and we can obtain the asymptotic performance of the estimate by using convex Gaussian min-max theorem (CGMT). In this paper, we briefly explain the analysis via CGMT and the result for the optimization problem with the ℓ_1 regularizer. As applications of CGMT-based analysis, we also introduce a method for the estimation of the unknown noise variance and asymptotic analysis of alternating direction method of multipliers (ADMM) for linear inverse problems.

Key words Linear inverse problems, compressed sensing, convex optimization, theoretical analysis

1. はじめに

未知のベクトルをその線形観測から推定する問題は線形逆問題と呼ばれる。特に、未知ベクトルのスパース性を活用して少ない観測から推定を行う圧縮センシング [1]~[4] は、無線信号処理 [5], [6] や画像処理 [7], [8] などの信号処理分野で注目を集めている。圧縮センシングのアイデアは、低ランク性 [9], [10] や離散性 [11], [12] をもつ信号の推定にも応用されている。

圧縮センシングのための有力なアプローチの1つに、凸最適化を用いる方法がある。凸最適化に基づくアプローチでは、スパース性を活用するための正則化を目的関数に含んだ最適化問

題を解くことで未知ベクトルを推定する。有名な正則化の1つは ℓ_1 ノルムに基づく正則化 (ℓ_1 正則化) であり、ISTA (Iterative Soft Thresholding Algorithm) [13]~[15] や FISTA (Fast ISTA) [16] によって ℓ_1 正則化を用いた最適化問題を解くことができる。また、ADMM (Alternating Direction Method of Multipliers) [17]~[20] は ISTA や FISTA よりも広いクラスの最適化問題に適用可能であり、 ℓ_1 正則化を用いた最適化問題にも適用できる。

圧縮センシングの数理的な側面に関しては、さまざまな方法で理論的な解析がなされている [21]~[23]。特に、近年開発されたCGMT (Convex Gaussian Min-max Theorem) [24], [25] を用いると、様々な最適化問題によって得られる推定値の漸近的

な特性を解析できる。CGMT を用いた特性解析では、解析の対象となる最適化問題に対応する別の最適化問題の最適解を解析し、その結果が元の最適化問題の最適解に対しても成り立つことを CGMT によって示す。CGMT を用いた解析手法は様々な最適化問題に応用できる強力なものであるが、CGMT 自体もその解析の流れも比較的複雑であり、それらをすぐに使いこなすのは難しいように思われる。

そこで本稿では、 ℓ_1 正則化を用いた最適化問題を題材として、CGMT を用いた解析方法の概要や解析結果について説明する。 ℓ_1 正則化を用いた最適化問題によって得られる推定値の漸近的な MSE (Mean Squared Error, 平均二乗誤差) に関する結果を述べ、計算機シミュレーションによって推定値の MSE と理論的な結果を比較する。次に、CGMT に基づく理論解析の応用例として、雑音分散が未知の場合にそれを推定する手法 [26] を紹介する。この手法では、雑音分散の情報を使わずに計算できる推定値の残差と、雑音分散の値を使って計算される漸近的な残差の値を比較することで雑音分散を推定する。また、CGMT の別の応用例として、圧縮センシングのための ADMM に関する特性解析 [27] についても述べる。ADMM の更新式で現れる部分問題に対して CGMT を用いた解析を適用し、ADMM の各反復における推定値の漸近的な MSE を評価した結果を示す。

本稿では、以下の記法を用いる。 \mathbb{R} は実数全体の集合を表す。転置を $(\cdot)^T$ 、単位行列を \mathbf{I} で表す。ベクトル $\mathbf{u} = [u_1 \cdots u_N]^T \in \mathbb{R}^N$ に対し、 ℓ_1 ノルムを $\|\mathbf{u}\|_1 = \sum_{n=1}^N |u_n|$ 、 ℓ_2 ノルムを $\|\mathbf{u}\|_2 = \sqrt{\sum_{n=1}^N u_n^2}$ と書く。 $\|\mathbf{u}\|_0$ は \mathbf{u} の非零成分の個数を表す。下半連続関数 $\phi: \mathbb{R}^N \rightarrow \mathbb{R} \cup \{\infty\}$ に対して、その近接写像と Moreau envelope をそれぞれ $\text{prox}_\phi(\mathbf{u}) = \arg \min_{\mathbf{s} \in \mathbb{R}^N} \left\{ \phi(\mathbf{s}) + \frac{1}{2} \|\mathbf{s} - \mathbf{u}\|_2^2 \right\}$ および $\text{env}_\phi(\mathbf{u}) = \min_{\mathbf{s} \in \mathbb{R}^N} \left\{ \phi(\mathbf{s}) + \frac{1}{2} \|\mathbf{s} - \mathbf{u}\|_2^2 \right\}$ と定義する。確率変数の列 $\{\Theta_n\}$ ($n = 1, 2, \dots$) が Θ に確率収束するとき、 $\Theta_n \xrightarrow{P} \Theta$ ($n \rightarrow \infty$) と書く。

2. 準備

2.1 線形逆問題に対する凸最適化

未知の N 次元ベクトル $\mathbf{x} = [x_1 \cdots x_N]^T \in \mathbb{R}^N$ をその線形観測

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v} \in \mathbb{R}^M \quad (1)$$

から推定する問題は線形逆問題と呼ばれる。ここで、 $\mathbf{A} \in \mathbb{R}^{M \times N}$ は既知の観測行列であり、本稿では各成分が独立に平均 $0 \cdot$ 分散 $1/N$ のガウス分布に従うとする。また、 $\mathbf{v} \in \mathbb{R}^M$ は雑音ベクトルであり、本稿では各成分が独立に平均 $0 \cdot$ 分散 σ_v^2 のガウス分布に従うとする。観測率を $\Delta := M/N$ で表す。

本稿では、線形逆問題の例として、スパース (ほとんどの成分が 0) な未知ベクトル \mathbf{x} を推定する問題を考える。具体的には、 \mathbf{x} の成分が独立に確率密度関数

$$p_X(x) = p_0 \delta_0(x) + (1 - p_0) p_G(x) \quad (2)$$

に従うものとする。ここで、 $p_0 \in [0, 1]$ であり、 $\delta_0(\cdot)$ と $p_G(\cdot)$ はそれぞれディラックのデルタ関数と標準ガウス分布の確率密

度関数を表す。 p_0 が大きいほど、未知ベクトル \mathbf{x} はスパースになる。

スパースなベクトルを推定するための基本的な凸最適化問題として、 ℓ_1 正則化 $f(\mathbf{s}) = \|\mathbf{s}\|_1$ を用いた最適化問題

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{s} \in \mathbb{R}^N} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{s}\|_2^2 + \lambda f(\mathbf{s}) \right\} \quad (3)$$

がある。 $\lambda (> 0)$ は、第一項 $\frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{s}\|_2^2$ と第二項 $\lambda f(\mathbf{s})$ のバランスをとるための正則化パラメータである。

式 (3) の最適化問題を解くことで、未知ベクトル \mathbf{x} の推定値 $\hat{\mathbf{x}}$ を得られる。例えば ADMM を式 (3) に適用すると、 $k = 0, 1, 2, \dots$ として以下の計算を反復することで推定値 $\mathbf{s}^{(k)}$ を更新するアルゴリズムを導出できる。

$$\mathbf{s}^{(k+1)} = \arg \min_{\mathbf{s} \in \mathbb{R}^N} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{s}\|_2^2 + \frac{\rho}{2} \left\| \mathbf{s} - \mathbf{z}^{(k)} + \mathbf{w}^{(k)} \right\|_2^2 \right\} \quad (4)$$

$$= \left(\mathbf{A}^T \mathbf{A} + \rho \mathbf{I} \right)^{-1} \left(\mathbf{A}^T \mathbf{y} + \rho \left(\mathbf{z}^{(k)} - \mathbf{w}^{(k)} \right) \right), \quad (5)$$

$$\mathbf{z}^{(k+1)} = \arg \min_{\mathbf{z} \in \mathbb{R}^N} \left\{ \lambda f(\mathbf{z}) + \frac{\rho}{2} \left\| \mathbf{s}^{(k+1)} - \mathbf{z} + \mathbf{w}^{(k)} \right\|_2^2 \right\} \quad (6)$$

$$= \text{prox}_{\frac{\lambda}{\rho} f} \left(\mathbf{s}^{(k+1)} + \mathbf{w}^{(k)} \right), \quad (7)$$

$$\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} + \mathbf{s}^{(k+1)} - \mathbf{z}^{(k+1)}. \quad (8)$$

ここで、 $\rho (> 0)$ はパラメータである。

2.2 CGMT

CGMT [24], [25] は式 (3) のような正則化付き最適化問題の漸近的な特性を解析するために開発された定理である。CGMT では、以下の主問題 (Primary Optimization, PO) と補助問題 (Auxiliary Optimization, AO)

$$(\text{PO}): \Phi(\mathbf{G}) = \min_{\mathbf{e} \in \mathcal{S}_e} \max_{\mathbf{u} \in \mathcal{S}_u} \left\{ \mathbf{u}^T \mathbf{G} \mathbf{e} + \xi(\mathbf{e}, \mathbf{u}) \right\}, \quad (9)$$

$$(\text{AO}): \phi(\mathbf{g}, \mathbf{h}) = \min_{\mathbf{e} \in \mathcal{S}_e} \max_{\mathbf{u} \in \mathcal{S}_u} \left\{ \|\mathbf{e}\|_2 \mathbf{g}^T \mathbf{u} - \|\mathbf{u}\|_2 \mathbf{h}^T \mathbf{e} + \xi(\mathbf{e}, \mathbf{u}) \right\} \quad (10)$$

を考える。ここで、行列 $\mathbf{G} \in \mathbb{R}^{M \times N}$ とベクトル $\mathbf{g} \in \mathbb{R}^M, \mathbf{h} \in \mathbb{R}^N$ の要素はそれぞれ独立であり、すべて標準ガウス分布に従うものとする。集合 $\mathcal{S}_e \subset \mathbb{R}^N$ と $\mathcal{S}_u \subset \mathbb{R}^M$ はコンパクトであるとする。関数 $\xi(\mathbf{e}, \mathbf{u}): \mathbb{R}^N \times \mathbb{R}^M \rightarrow \mathbb{R}$ は連続であり、 \mathbf{e} に関しては凸関数で \mathbf{u} に関しては凹関数であると仮定する。このとき、 $M, N \rightarrow \infty$ ($M/N = \Delta$) の極限 (以後単に $M, N \rightarrow \infty$ と書く) を考えると、以下の定理が成り立つ。

[定理 1] (Asymptotic CGMT [25]) 集合 \mathcal{S} を \mathcal{S}_e 内の開集合とし、その補集合を $\mathcal{S}^c = \mathcal{S}_e \setminus \mathcal{S}$ とする。また、主問題 (PO) の \mathbf{e} の最適解を $\hat{\mathbf{e}}_\Phi(\mathbf{G})$ とおき、 $\mathbf{e} \in \mathcal{S}^c$ の制約のもとでの補助問題 (AO) の最適値を $\phi_{\mathcal{S}^c}(\mathbf{g}, \mathbf{h})$ とおく。もしある定数 $\bar{\phi}, \bar{\phi}_{\mathcal{S}^c}$ が存在して $\bar{\phi} < \bar{\phi}_{\mathcal{S}^c}$ かつ $\phi(\mathbf{g}, \mathbf{h}) \xrightarrow{P} \bar{\phi}, \phi_{\mathcal{S}^c}(\mathbf{g}, \mathbf{h}) \xrightarrow{P} \bar{\phi}_{\mathcal{S}^c}$ ($M, N \rightarrow \infty$) となるならば、 $\lim_{M, N \rightarrow \infty} \Pr(\hat{\mathbf{e}}_\Phi(\mathbf{G}) \in \mathcal{S}) = 1$ が成り立つ。

定理 1 は、直感的には「補助問題 (AO) の最適解が高い確率で集合 \mathcal{S} に属するならば、主問題 (PO) の最適解も高い確

率で集合 S に属する」ということを述べている。したがって、解析の対象となる最適化問題を主問題 (PO) の形で表現できる場合、主問題 (PO) の代わりに補助問題 (AO) を解析し、その結果を主問題 (PO) に適用することが可能となる。補助問題 (AO) は主問題 (PO) の第一項 $\mathbf{u}^T \mathbf{G} \mathbf{e}$ を $\|\mathbf{e}\|_2 \mathbf{g}^T \mathbf{u} - \|\mathbf{u}\|_2 \mathbf{h}^T \mathbf{e}$ に「分解」したものであり、主問題 (PO) よりも解析しやすい場合が多い。

3. CGMT を用いた特性解析

CGMT を用いることで、線形逆問題に対する様々な凸最適化問題の漸近的な特性を得られる。本章では、式 (3) の最適化問題を題材として、CGMT を用いた特性解析方法の大まかな流れを述べ、推定値 $\hat{\mathbf{x}}$ の MSE $\|\hat{\mathbf{x}} - \mathbf{x}\|_2^2 / N$ に関する解析結果を紹介する。

CGMT を用いた特性解析では、まず誤差ベクトルを $\mathbf{e} := \mathbf{s} - \mathbf{x}$ と定義し、式 (3) の最適化問題を \mathbf{e} に関する最適化問題

$$\min_{\mathbf{e} \in \mathbb{R}^N} \frac{1}{N} \left\{ \frac{1}{2} \|\mathbf{A} \mathbf{e} - \mathbf{v}\|_2^2 + \lambda f(\mathbf{x} + \mathbf{e}) \right\} \quad (11)$$

に書き換える。ただし、 $M, N \rightarrow \infty$ の極限で目的関数の値が発散しないようにするため、目的関数全体を $1/N$ 倍している。ここで、関数 $\psi(\mathbf{r}) := \|\mathbf{r}\|_2^2 / 2$ ($\mathbf{r} \in \mathbb{R}^M$) の凸共役が

$$\psi^*(\mathbf{z}^*) = \max_{\mathbf{u} \in \mathbb{R}^M} \left\{ \mathbf{u}^T \mathbf{z}^* - \frac{1}{2} \|\mathbf{u}\|_2^2 \right\} \quad (12)$$

$$= \frac{1}{2} \|\mathbf{z}^*\|_2^2 \quad (13)$$

で与えられることを利用すると、

$$\frac{1}{2} \|\mathbf{A} \mathbf{e} - \mathbf{v}\|_2^2 = \max_{\mathbf{u} \in \mathbb{R}^M} \left\{ \sqrt{N} \mathbf{u}^T (\mathbf{A} \mathbf{e} - \mathbf{v}) - \frac{N}{2} \|\mathbf{u}\|_2^2 \right\} \quad (14)$$

が成り立つ。式 (14) を用いて式 (11) を変形して整理すると、

$$\min_{\mathbf{e} \in \mathbb{R}^N} \max_{\mathbf{u} \in \mathbb{R}^M} \left\{ \frac{1}{N} \mathbf{u}^T (\sqrt{N} \mathbf{A}) \mathbf{e} - \frac{1}{\sqrt{N}} \mathbf{v}^T \mathbf{u} - \frac{1}{2} \|\mathbf{u}\|_2^2 + \frac{\lambda}{N} f(\mathbf{x} + \mathbf{e}) \right\} \quad (15)$$

と書ける。式 (15) は主問題 (PO) の形 (の定数倍) になっている¹ ため、対応する補助問題 (AO)

$$\min_{\mathbf{e} \in \mathbb{R}^N} \max_{\mathbf{u} \in \mathbb{R}^M} \left\{ \frac{1}{N} \left(\|\mathbf{e}\|_2 \mathbf{g}^T \mathbf{u} - \|\mathbf{u}\|_2 \mathbf{h}^T \mathbf{e} \right) - \frac{1}{\sqrt{N}} \mathbf{v}^T \mathbf{u} - \frac{1}{2} \|\mathbf{u}\|_2^2 + \frac{\lambda}{N} f(\mathbf{x} + \mathbf{e}) \right\} \quad (16)$$

の最適解や最適値について調べ、CGMT を適用することで以下の結果が得られる (解析方法の詳細は [25], [28], [29] を参照されたい)。

[定理 2] 観測行列 $\mathbf{A} \in \mathbb{R}^{M \times N}$ の各成分は独立に平均 $0 \cdot$ 分散

(注1): 式 (15), (16) では $\mathbf{e} \in \mathbb{R}^N, \mathbf{u} \in \mathbb{R}^M$ としているため厳密にはそのまま CGMT を適用することはできないが、最適解があるコンパクト集合に属することを示すことで CGMT を適用可能になる。

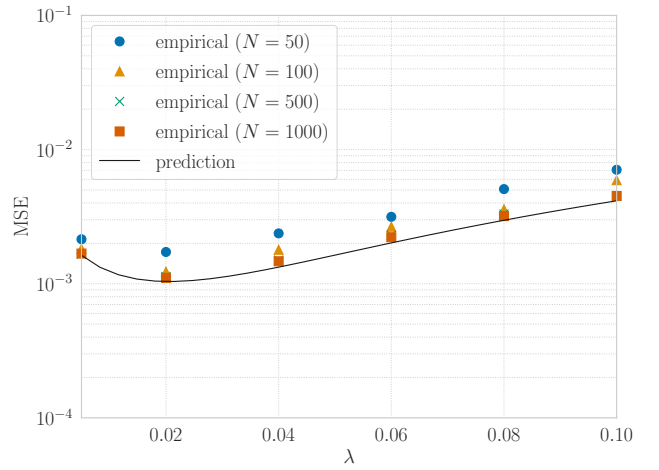


図1 最適化問題 (3) の MSE 特性とその理論解析結果 ($\Delta = 0.6, p_0 = 0.9, \sigma_v^2 = 0.001$)

$1/N$ のガウス分布に従うとする。雑音ベクトル $\mathbf{v} \in \mathbb{R}^M$ の各成分は独立に平均 $0 \cdot$ 分散 σ_v^2 のガウス分布に従うとする。また、

$$F(\alpha, \beta) = \frac{\alpha \beta \sqrt{\Delta}}{2} + \frac{\sigma_v^2 \beta \sqrt{\Delta}}{2\alpha} - \frac{\beta^2}{2} - \frac{\alpha \beta}{2\sqrt{\Delta}} + \frac{\beta \sqrt{\Delta}}{\alpha} \mathbb{E} \left[\text{env}_{\frac{\alpha \lambda}{\beta \sqrt{\Delta}}} f \left(X + \frac{\alpha}{\sqrt{\Delta}} H \right) \right] \quad (17)$$

に関する最適化問題 $\min_{\alpha > 0} \max_{\beta \geq 0} F(\alpha, \beta)$ がただ一つの解 (α^*, β^*) をもつとする。また、 $X \sim p_X$ および $H \sim p_G$ である。このとき、

$$\frac{1}{N} \|\hat{\mathbf{x}} - \mathbf{x}\|_2^2 \xrightarrow{P} (\alpha^*)^2 - \sigma_v^2 \quad (M, N \rightarrow \infty) \quad (18)$$

が成り立つ。

定理 2 を用いて、式 (3) の最適化問題によって得られる $\hat{\mathbf{x}}$ の MSE $\|\hat{\mathbf{x}} - \mathbf{x}\|_2^2 / N$ の漸近的な値を評価できる。図 1 に、 $\Delta = 0.6, p_0 = 0.9, \sigma_v^2 = 0.001$ のときの最適化問題 (3) の MSE 特性を示す。横軸は正則化パラメータ λ である。‘empirical’ は ADMM を用いて求めた推定値 $\hat{\mathbf{x}}$ の MSE を表し、‘prediction’ は定理 2 を用いて求めた漸近的な MSE $(\alpha^*)^2 - \sigma_v^2$ を表す。推定値の MSE は、独立に行った 500 回の推定の結果を平均して求めた。問題のサイズ N が十分大きい場合は、実際に推定値を求めたときの特性とその理論解析結果が近い値となっていることがわかる。また、この場合に MSE を最小にする正則化パラメータの値はおおよそ $\lambda = 0.02$ であることもわかる。

式 (3) の最適化問題以外にも、様々な凸最適化問題の漸近的な MSE や NSE (Normalized Squared Error) が CGMT によって得られている [28], [30]~[33]。また、 ± 1 を等確率でとる BPSK (Binary Phase Shift Keying) 信号の推定における範囲制約付き最適化問題の漸近的な SER (Symbol Error Rate) の導出 [34] や、その PAM (Pulse Amplitude Modulation) 信号への一般化 [25] も行われている。さらなる一般化として、離散値をとるベクトルの推定に対する SOAV (Sum of Absolute Values) 最適化 [12] の漸近的な特性も得られている [29]。CGMT を用いた特性解析は非線形観測からの推定にも適用されており [35]、同様のアプ

ローチが低分解能 ADC (Analog-to-Digital Converters) を用いた MIMO (Multiple-Input Multiple-Output) 信号検出の解析に応用されている [36]. 観測行列の情報に誤差が含まれている場合についても検討されており [37], 同様の方法によって BPSK 信号を用いた MIMO 伝送における電力割り当てについて議論されている [38]. CGMT は一定の条件下で複素数領域での問題にも応用可能であり, 例えば MIMO 伝送における PSK (Phase Shift Keying) 信号の推定に関する解析結果が得られている [39].

4. CGMT を用いた特性解析の応用

本章では, CGMT を用いた特性解析の応用例として, 雑音分散 σ_v^2 が未知の場合に σ_v^2 を推定する手法 [26] と, 式 (3) の最適化問題に対する ADMM の特性解析 [27] の概要について述べる.

4.1 未知の雑音分散の推定 [26]

式 (3) の最適化問題に含まれる正則化パラメータ λ の適切な値は雑音分散 σ_v^2 の値によって変化するため, 雑音分散の情報は正則化パラメータの設定において重要となる. 本節では, 雑音分散 σ_v^2 が未知であるとし, CGMT を用いた特性解析結果を用いて σ_v^2 を推定する手法である ARM (Asymptotic Residual Matching) [26] について概説する.

準備として, 推定値 $\hat{\mathbf{x}}$ の残差 $\text{Res}(\hat{\mathbf{x}}) := \|\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}\|_2^2 / N$ の漸近的な評価に関する定理を述べておく.

[定理 3] 定理 2 と同じ状況で,

$$\text{Res}(\hat{\mathbf{x}}) := \frac{1}{N} \|\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}\|_2^2 \xrightarrow{P} (\beta^*)^2 \quad (M, N \rightarrow \infty) \quad (19)$$

が成り立つ.

ARM では, 推定値 $\hat{\mathbf{x}}$ の残差とその理論解析結果を比較することで雑音分散を推定する. まず, 正則化パラメータ λ を適当に設定して, 様々な σ_v^2 に対する $(\beta^*)^2$ の値を事前に計算しておく. その λ を用いた式 (3) の最適化問題を解くことで, 推定値の残差 $\text{Res}(\hat{\mathbf{x}})$ は雑音分散 σ_v^2 の情報を使わずに計算できるため, $\text{Res}(\hat{\mathbf{x}})$ と近い $(\beta^*)^2$ に対応する σ_v^2 を得ることで雑音分散を推定する. すなわち,

$$\hat{\sigma}_v^2 = \arg \min_{\sigma^2 > 0} \left| \beta^*(\sigma^2)^2 - \text{Res}(\hat{\mathbf{x}}) \right| \quad (20)$$

を σ_v^2 の推定値とする. ただし, $\beta^*(\sigma^2)$ は式 (17) で $\sigma_v^2 = \sigma^2$ としたときの $F(\alpha, \beta)$ を用いて求めた β^* の値を表す. σ_v^2 の推定精度は λ の設定値に依存するため適切な λ の値を選ぶ必要があるが, 一度求めた推定値 $\hat{\sigma}_v^2$ の値から λ を再設定し, あらためて σ_v^2 を推定することで推定精度を改善できる [26].

計算機シミュレーションにより, 雑音分散推定の NMSE (Normalized MSE) 特性 $(\hat{\sigma}_v^2 - \sigma_v^2)^2 / (\sigma_v^2)^2$ を評価する. 図 2 に, $N = 200, \Delta = 0.8, p_0 = 0.9$ としたときの NMSE 特性を示す. 横軸は雑音分散 σ_v^2 の真の値である. ARM ではまず $\lambda = 0.01$ として σ_v^2 を推定し, その結果を用いて λ を再設定してからもう一度 σ_v^2 を推定した (詳細は [26] を参照). 簡単のため p_0 は既知としているが, $\hat{p}_0 = 1 - \|\mathbf{y}\|_2^2 / M$ として推定した値を用いることでほぼ同等の特性を得られる. ‘ridge-regularization

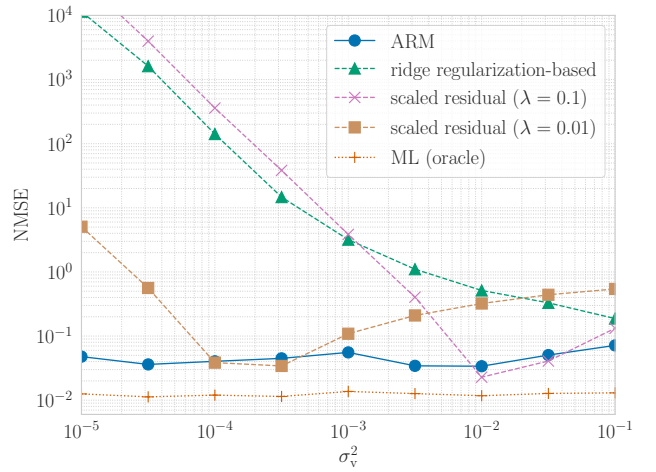


図 2 雑音分散 σ_v^2 の推定における NMSE 特性
($N = 200, \Delta = 0.8, p_0 = 0.9$)

based’ は, リッジ正則化を用いた最適化の漸近的な特性解析を用いた手法 [40] である. ‘scaled residual’ は, 推定値 $\hat{\mathbf{x}}$ を用いて $\hat{\sigma}_v^2 = \|\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}\|_2^2 / (M - \|\hat{\mathbf{x}}\|_0)$ とする手法 [41], [42] である. 正則化パラメータ λ の値は適切に定める必要があり, 交差検証に基づくアプローチなどが検討されているが, 今回は簡単のため $\lambda = 0.1, 0.01$ と設定した. ‘ML (oracle)’ は, \mathbf{x} が既知であるとしたときの最尤推定値 $\hat{\sigma}_v^2 = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 / M$ の特性である. これは比較のために評価したものであり, 他の手法では \mathbf{x} は未知であることに注意されたい. 図 2 から, ‘ridge-regularization based’ や ‘scaled residual’ では雑音分散をうまく推定できない場合がある一方で, ARM はどの σ_v^2 に対しても 10^{-1} 以下の NMSE を達成していることがわかる.

4.2 ADMM の特性解析 [27]

CGMT を用いた特性解析は, 式 (3) のような最適化問題の解析だけでなく, それを解くための最適化アルゴリズムの解析にも応用できる場合がある. 本節では, 式 (4)–(8) で与えられる ADMM の更新式の漸近的な特性を解析した結果 [27] について述べる.

ADMM における $\mathbf{s}^{(k)}$ の更新式 (4) の最適化問題は, 第 3 章で述べた手続きによって CGMT の主問題 (PO) の形に変形可能である. 式 (4) の最適化問題を CGMT を用いて解析し, さらに $\mathbf{z}^{(k)}$ と $\mathbf{w}^{(k)}$ の更新が成分ごとに行われることを考慮すると, 以下の定理が得られる.

[定理 4] 観測行列 \mathbf{A} と雑音ベクトル \mathbf{v} は定理 2 における仮定をみたすとする. まず, 確率過程

$$S_{k+1} = \hat{S}_{k+1}(\alpha_k^*, \beta_k^*), \quad (21)$$

$$Z_{k+1} = \text{prox}_{\frac{\lambda}{\rho} f}(S_{k+1} + W_k), \quad (22)$$

$$W_{k+1} = W_k + S_{k+1} - Z_{k+1}, \quad (23)$$

を考える. ただし, $Z_0 = W_0 = 0$ とする. ここで, $\hat{S}_{k+1}(\alpha, \beta)$ は確率変数 $X \sim p_X$ と $H \sim p_G$ を用いて

$$\hat{S}_{k+1}(\alpha, \beta)$$

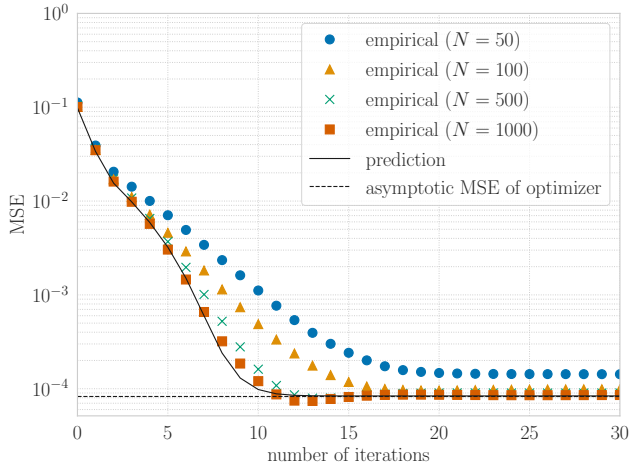


図3 ADMM の MSE 特性とその理論解析結果
($\Delta = 0.7$, $p_0 = 0.9$, $\sigma_v^2 = 0.0001$, $\rho = 0.1$)

$$= \frac{1}{\frac{\beta\sqrt{\Delta}}{\alpha} + \rho} \left(\frac{\beta\sqrt{\Delta}}{\alpha} \left(X + \frac{\alpha}{\sqrt{\Delta}} H \right) + \rho(Z_k - W_k) \right) \quad (24)$$

と定義される。さらに、 (α_k^*, β_k^*) は最適化問題

$$\min_{\alpha > 0} \max_{\beta \geq 0} \left\{ \frac{\alpha\beta\sqrt{\Delta}}{2} + \frac{\beta\sigma_v^2\sqrt{\Delta}}{2\alpha} - \frac{\beta^2}{2} + \mathbb{E} \left[J^{(k+1)}(\alpha, \beta) \right] \right\} \quad (25)$$

のただ一つの解であるとする。ここで、

$$\begin{aligned} J^{(k+1)}(\alpha, \beta) &= \frac{\beta\sqrt{\Delta}}{2\alpha} \left(\hat{S}_{k+1}(\alpha, \beta) - X \right)^2 - \beta H \left(\hat{S}_{k+1}(\alpha, \beta) - X \right) \\ &\quad + \frac{\rho}{2} \left(\hat{S}_{k+1}(\alpha, \beta) - Z_k + W_k \right)^2 \end{aligned} \quad (26)$$

である。

このとき、 $k = 0, 1, 2, \dots$ に対して

$$\frac{1}{N} \left\| \mathbf{s}^{(k+1)} - \mathbf{x} \right\|_2^2 \xrightarrow{P} (\alpha_k^*)^2 - \sigma_v^2 \quad (M, N \rightarrow \infty) \quad (27)$$

が成り立つ。

定理 4 により、ADMM の各反復における \mathbf{x} の推定値 $\mathbf{s}^{(k)}$ の漸近的な MSE がどのように変化するか分かる。図 3 に、 $\Delta = 0.7$, $p_0 = 0.9$, $\sigma_v^2 = 0.0001$ のときの ADMM の MSE 特性を示す。横軸は ADMM の更新式の反復回数である。図 1 と同様に、‘empirical’ が実際に \mathbf{x} を推定したときの MSE 特性 $\left\| \mathbf{s}^{(k+1)} - \mathbf{x} \right\|_2^2 / N$ を表し、‘prediction’ が定理 4 を用いて求めた漸近的な MSE $(\alpha_k^*)^2 - \sigma_v^2$ を表す。ADMM のパラメータは $\rho = 0.1$ とした。また、‘asymptotic MSE of optimizer’ は定理 2 から求まる最適解 $\hat{\mathbf{x}}$ の漸近的な MSE を表す。図 3 から、問題のサイズ N が十分大きい場合は、実際の特性と理論解析結果がよく一致していることがわかる。さらに、ADMM の反復回数が多くなるにつれて、推定値 $\mathbf{s}^{(k)}$ の MSE やその理論解析結果が最適解の漸近的な MSE に収束していることもわかる。

5. まとめ

本稿では、線形逆問題のための凸最適化問題の特性を CGMT によって解析する方法について概説した。CGMT を用いた特性解析では、解析対象の最適化問題を主問題 (PO) の形で表現し、対応する補助問題 (AO) の最適解や最適値について調べることで最適化問題の漸近的な特性を得る。また、CGMT を用いた解析の応用例として、未知の雑音分散の推定や ADMM の特性解析についても述べた。CGMT は様々な最適化問題の特性解析に適用可能であり、今後は最適化問題のパラメータ設定や最適化アルゴリズムの設計に応用されていくことが期待できる。

謝辞 本研究の一部は、科学研究費補助金 (研究課題番号 20K23324) の支援を受けたものである。

文献

- [1] E. J. Candès and T. Tao, “Decoding by linear programming,” *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
- [2] E. J. Candès, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [3] D. L. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [4] E. J. Candès and M. B. Wakin, “An introduction to compressive sampling,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [5] K. Hayashi, M. Nagahara, and T. Tanaka, “A user’s guide to compressed sensing for communications systems,” *IEICE TRANSACTIONS on Communications*, vol. E96-B, no. 3, pp. 685–712, Mar. 2013.
- [6] J. W. Choi, B. Shim, Y. Ding, B. Rao, and D. I. Kim, “Compressed sensing for wireless communications: Useful tips and tricks,” *IEEE Communications Surveys Tutorials*, vol. 19, no. 3, pp. 1527–1550, thirdquarter 2017.
- [7] M. Lustig, D. L. Donoho, and J. M. Pauly, “Sparse MRI: The application of compressed sensing for rapid MR imaging,” *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.
- [8] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, “Compressed sensing MRI,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 72–82, Mar. 2008.
- [9] E. J. Candès and B. Recht, “Exact matrix completion via convex optimization,” *Found Comput Math*, vol. 9, no. 6, pp. 717–772, Dec. 2009.
- [10] E. J. Candès and T. Tao, “The power of convex relaxation: Near-optimal matrix completion,” *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2053–2080, May 2010.
- [11] A. Aïssa-El-Bey, D. Pastor, S. M. A. Sbaï, and Y. Fadlallah, “Sparsity-based recovery of finite alphabet solutions to underdetermined linear systems,” *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 2008–2018, Apr. 2015.
- [12] M. Nagahara, “Discrete signal reconstruction by sum of absolute values,” *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1575–1579, Oct. 2015.
- [13] I. Daubechies, M. Defrise, and C. D. Mol, “An iterative thresholding algorithm for linear inverse problems with a sparsity constraint,” *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004.
- [14] P. L. Combettes and V. R. Wajs, “Signal recovery by proximal forward-backward splitting,” *Multiscale Model. Simul.*, vol. 4, no. 4, pp. 1168–1200, Jan. 2005.
- [15] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, “Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems,” *IEEE Journal of Selected Topics in Signal*

- Processing*, vol. 1, no. 4, pp. 586–597, Dec. 2007.
- [16] A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, Jan. 2009.
- [17] D. Gabay and B. Mercier, “A dual algorithm for the solution of non-linear variational problems via finite element approximation,” *Computers & Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, Jan. 1976.
- [18] J. Eckstein and D. P. Bertsekas, “On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators,” *Mathematical Programming*, vol. 55, no. 1, pp. 293–318, Apr. 1992.
- [19] P. L. Combettes and J.-C. Pesquet, “Proximal splitting methods in signal processing,” in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, ser. Springer Optimization and Its Applications, H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Eds. New York, NY: Springer New York, 2011, vol. 49, pp. 185–212.
- [20] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.
- [21] D. L. Donoho, A. Maleki, and A. Montanari, “The noise-sensitivity phase transition in compressed sensing,” *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6920–6941, Oct. 2011.
- [22] M. Bayati and A. Montanari, “The LASSO risk for Gaussian matrices,” *IEEE Transactions on Information Theory*, vol. 58, no. 4, pp. 1997–2017, Apr. 2012.
- [23] D. L. Donoho, I. Johnstone, and A. Montanari, “Accurate prediction of phase transitions in compressed sensing via a connection to minimax denoising,” *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3396–3433, Jun. 2013.
- [24] C. Thrampoulidis, S. Oymak, and B. Hassibi, “Regularized linear regression: A precise analysis of the estimation error,” in *Proc. Conference on Learning Theory*, Jun. 2015, pp. 1683–1709.
- [25] C. Thrampoulidis, E. Abbasi, and B. Hassibi, “Precise error analysis of regularized M -estimators in high dimensions,” *IEEE Transactions on Information Theory*, vol. 64, no. 8, pp. 5592–5628, Aug. 2018.
- [26] R. Hayakawa, “Noise variance estimation using asymptotic residual in compressed sensing,” *arXiv:2009.13678*, Sep. 2020.
- [27] —, “Asymptotic analysis of ADMM for compressed sensing,” *arXiv: 2009.08545*, Sep. 2020.
- [28] C. Thrampoulidis, W. Xu, and B. Hassibi, “Symbol error rate performance of box-relaxation decoders in massive MIMO,” *IEEE Transactions on Signal Processing*, vol. 66, no. 13, pp. 3377–3392, Jul. 2018.
- [29] R. Hayakawa and K. Hayashi, “Asymptotic performance of discrete-valued vector reconstruction via box-constrained optimization with sum of ℓ_1 regularizers,” *IEEE Transactions on Signal Processing*, vol. 68, pp. 4320–4335, 2020.
- [30] C. Thrampoulidis, E. Abbasi, and B. Hassibi, “LASSO with non-linear measurements is equivalent to one with linear measurements,” in *Proc. the 28th International Conference on Neural Information Processing Systems*, 2015, pp. 3420–3428.
- [31] C. Thrampoulidis, A. Panahi, D. Guo, and B. Hassibi, “Precise error analysis of the LASSO,” in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2015, pp. 3467–3471.
- [32] C. Thrampoulidis, A. Panahi, and B. Hassibi, “Asymptotically exact error analysis for the generalized ℓ_2^2 -LASSO,” in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Jun. 2015, pp. 2021–2025.
- [33] I. B. Atitallah, C. Thrampoulidis, A. Kammoun, T. Y. Al-Naffouri, M. Alouini, and B. Hassibi, “The BOX-LASSO with application to GSSK modulation in massive MIMO systems,” in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Jun. 2017, pp. 1082–1086.
- [34] C. Thrampoulidis, E. Abbasi, W. Xu, and B. Hassibi, “BER analysis of the box relaxation for BPSK signal recovery,” in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2016, pp. 3776–3780.
- [35] C. Thrampoulidis and B. Hassibi, “Isotropically random orthogonal matrices: Performance of LASSO and minimum conic singular values,” in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Jun. 2015, pp. 556–560.
- [36] C. Thrampoulidis and W. Xu, “The performance of box-relaxation decoding in massive MIMO with low-resolution ADCs,” in *Proc. IEEE Statistical Signal Processing Workshop (SSP)*, Jun. 2018, pp. 821–825.
- [37] A. M. Alrashdi, I. B. Atitallah, T. Y. Al-Naffouri, and M. Alouini, “Precise performance analysis of the LASSO under matrix uncertainties,” in *Proc. IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Nov. 2017, pp. 1290–1294.
- [38] A. M. Alrashdi, I. B. Atitallah, T. Ballal, C. Thrampoulidis, A. Chaaban, and T. Y. Al-Naffouri, “Optimum training for MIMO BPSK transmission,” in *Proc. IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2018, pp. 1–5.
- [39] E. Abbasi, F. Salehi, and B. Hassibi, “Performance analysis of convex data detection in MIMO,” in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 4554–4558.
- [40] M. A. Suliman, A. M. Alrashdi, T. Ballal, and T. Y. Al-Naffouri, “SNR estimation in linear systems with Gaussian matrices,” *IEEE Signal Processing Letters*, vol. 24, no. 12, pp. 1867–1871, Dec. 2017.
- [41] J. Fan, S. Guo, and N. Hao, “Variance estimation using refitted cross-validation in ultrahigh dimensional regression,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 74, no. 1, pp. 37–65, 2012.
- [42] S. Reid, R. Tibshirani, and J. Friedman, “A study of error variance estimation in LASSO regression,” *Statistica Sinica*, vol. 26, no. 1, pp. 35–67, 2016.