電子情報通信学会無線通信システム（RCS）研究会主催

5G の根底にある無線信号処理に関する基礎ワークショップ

初学者のための無線通信信号処理入門

林和則
大阪市立大学大学院工学研究科

平成 29 年 7 月 24 日

目 次

第1章　はじめに

現代の無線通信システムは高度に専門化された信号処理技術がその様々な構成要素に適用されることで成り立っているため，この分野の初学者がその全容を理解しようとしたときにどこから手をつけて良いかすら分からないような状況にある．また，無線通信に関する多くの教科書では，等化や通信路推定，ビームフォーミング，多元接続といった要素技術毎に説明がなされていることも，様々な信号処理手法に対する統一的な理解を得ることを困難にしていると考えられる．

典型的な無線通信の信号は搬送波周波数周辺にその周波数成分が局在した帯域信号であり，その信号処理の問題は等価低域系（ベースバンド）で考えることがほとんどである．実際に送信される無線信号は当然実数の信号であり，そのスペクトルは周波数 0 の原点に関して正負の周波数で対称であるが，正の周波数領域におけるスペクトルは必ずしも搬送波周波数を中心として左右対称とはならず，等価低域系の信号は一般に複素数となる．このため，無線通信の信号処理の多くの問題では，複素数の信号に対して複素数の重み係数を用いて様々な処理を行うことになる．一方，これらの複素数の重み係数を決定する際には何らかの尺度で最も良いものを選択する必要があるが，複素数自体は素朴に順序付けできないため，通常，複素数の重み係数を引数とし実数値の出力をもつ関数（コスト関数や目的関数など）についての最適化問題を考えることになる．すなわち，通信のための信号処理では，複素のスカラー，ベクトル，あるいは行列をその引数としてもつ実数値の関数の最大化あるいは最小化を考えることが基本的である．ところが，複素数の引数に対して実数値をとる関数は一般に正則（複素微分可能）ではなく，工学部の標準的なカリキュラムで学ぶ複素関数論の知識では最適なフィルタ係数を決定できない[1]という問題がある．

本稿では，無線通信や信号処理の非専門家や初学者を対象として，典型的な問題に対する典型的な信号処理手法について解説することで，現在の無線通信システムで用いられている最先端の信号処理技術を理解するために必要な最低限の知識を，効率的かつ体系的に提供することを目的としている．まず，そのための準備として，通信の信号やシステムを確率的に取り扱う際に用いられる数学モデルである確率過程の基礎事項とその自己相関行列の基本的な性質について説明する．また，無線通信のための信号処理における目的関数やコスト関数の最大化，最小化に必要な，非正則な複素関数を複素数，複素ベクトルおよび複素行列で微分するための枠組み（ウィルティンガー微分）についても説明する．その後，典型的な問題設定として，興味のある未知ベクトル $\mathbf{x}$ に対して既知の行列 $\mathbf{A}$ で線形観測を行ったときに，白色の加法性雑音 $\mathbf{v}$ をともなって観測されるベクトル $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}$ から未知ベクトルを推定する問題を考える．これは素朴な線形回帰のモデルであるが，等化や通信路推定，MIMO

[1] “決定できない” というのは少し誇張し過ぎで，フィルタ係数の実部と虚部からなる実 2 変数関数として取り扱えば決定可能であるが，複素関数のまま取り扱うアプローチに比べて計算が非常に煩雑になってしまう．

(multi-input multi-output) 信号検出，ダイバーシティ，干渉除去，ユーザー分離など，通信における多くの信号処理の問題を表現することが可能である．また，同じモデルにおいて線形観測の行列を未知とすることで，到来方向推定やブラインド通信路推定などで利用されるマルチチャネルモデルの問題を表現することも可能である．このモデルに基づいて，ZF (zero-forcing) 推定や最小ノルム解，MMSE (minimum mean-square-error) 推定，減算型干渉除去，最大事後確率推定，最尤推定，最大比合成といった古典的な推定法から圧縮センシングなどの比較的新しい手法まで説明する．また，アレイ信号処理でよく利用される部分空間法についても説明する．さらに，これらの基本的な手法の応用例として，等化や通信路推定，MIMO 信号検出，到来方向推定などの具体例についても解説する．最後に，この分野の初学者が本稿を読んだ後にぜひ読んでほしいいくつかの文献を紹介する．

第2章　準備

ここでは準備として，本稿で使用する数学上の表記と，後述の信号処理手法を理解するために必要ないくつかの基礎事項，具体的には，確率過程の基本的性質と，非正則関数を複素数，複素ベクトル，複素行列で微分する方法（ウィルティンガー微分）について説明する．

2.1　表記

本稿では，次のような表記を用いる．$\mathbb{R}$ と $\mathbb{C}$ はそれぞれ実数と複素数の集合を表す．ボールド体の小文字は列ベクトルを，ボールド体の大文字は行列をそれぞれ表す．複素共役は * で表し，行列，ベクトルの転置の操作は $^{\mathrm{T}}$，共役転置は $^{\mathrm{H}}$ とする．$\mathrm{diag}[a_1\ \cdots\ a_N]$ は $a_1, \cdots, a_N$ を対角成分にもつ $N \times N$ の対角行列を表し，$\mathrm{tr}\{\mathbf{A}\}$ は行列 $\mathbf{A}$ のトレースを，$\det\{\mathbf{A}\}$ は行列 $\mathbf{A}$ の行列式を表す．単位行列は $\mathbf{I}$，零行列は $\mathbf{0}$ とし，文脈からサイズが分かる場合には特に明記しない．$||\mathbf{a}||_p$（$p \geq 1$）はベクトル $\mathbf{a} = [a_1\ a_2\ \cdots\ a_n]^{\mathrm{T}} \in \mathbb{R}^n$ の ℓ_p-ノルムであり，

$$||\mathbf{a}||_p = \left(\sum_{i=1}^{n} |a_i|^p\right)^{\frac{1}{p}} \tag{2.1}$$

で定義される．また，$||\mathbf{a}||_0$ はベクトル $\mathbf{a}$ の非零要素数を意味し，ℓ_0-ノルムと呼ぶことにする．虚数単位は j とし，$\Re\{\cdot\}$, $\Im\{\cdot\}$ はそれぞれ実部と虚部を表す．

2.2　確率過程とその性質

通信のための信号処理では，雑音や複雑な通信路の影響を考慮するために信号や未知パラメータを確率的に取り扱うことが多い．そのような信号などを表現するために用いられる数学モデルが確率過程である．確率過程はそれぞれに時刻[1]が与えられた確率変数の集合として定義され，離散時間の場合は番号付けされた確率変数の集合として $\{x(n);\ n = \ldots, -2, -1, 0, 1, 2, \ldots\}$ とかかれる（離散時間の場合は時系列とも呼ばれる）．したがって，時刻 n をある特定の値 n_0 に固定した場合，$x(n_0)$ は確率変数を表すことに注意する．

統計的信号処理では確率過程を特徴づけるために $x(n)$ を引数とする様々な関数の (集合) 平均を考えるが，その最も基本的なものが $x(n)$ 自体の平均

$$m_{\mathrm{x}}(n) = E[x(n)] \tag{2.2}$$

[1]インデックスであれば空間軸の座標など何でも良いが，本稿では時刻とする．

である[2]．ここで $E[\cdot]$ は期待値の操作を表し，具体的には，時刻 n を固定した $x(n)$ が確率密度関数 $p(x(n))$ をもつ連続確率変数の場合には

$$E[x(n)] = \int x(n)p(x(n))dx(n) \tag{2.3}$$

確率分布 $P(x(n))$ をもつ離散確率変数の場合には

$$E[x(n)] = \sum x(n)P(x(n)) \tag{2.4}$$

と計算される．積分および和の線形性から，いずれの場合も $E[\cdot]$ が線形性をもつことが重要である．以下では，簡単のため任意の時刻 n に固定した $x(n)$ が全て連続確率変数であるとする．

確率過程 $x(n)$ の (自己) 相関関数は

$$r_{\mathrm{x}}(n,k) = E[x(n)x^*(n-k)] \tag{2.5}$$

で定義される[3]．時刻 n を固定したときの $x(n)$ と $x(n-k)$ の同時確率分布の確率密度関数を $p(x(n), x(n-k))$ とすると，この期待値は

$$E[x(n)x^*(n-k)] = \int\int x(n)x^*(n-k)p(x(n), x(n-k))dx(n)dx(n-k) \tag{2.6}$$

と計算される．また，

$$c_{\mathrm{x}}(n,k) = E[(x(n) - m_{\mathrm{x}}(n))(x(n-k) - m_{\mathrm{x}}(n-k))^*] \tag{2.7}$$

を確率過程 $x(n)$ の (自己) 共分散関数という．特に，

$$\sigma_{\mathrm{x}}^2(n) = c_{\mathrm{x}}(n,0) = E[(x(n) - m_{\mathrm{x}}(n))(x(n) - m_{\mathrm{x}}(n))^*] \tag{2.8}$$

を分散といい，分散の有界な確率過程を 2 次過程という[4]．

(2.2) の平均は一般に時刻 n に依存するため n を引数にもち，(2.5) の自己相関や (2.7) の自己共分散は一般に二つの時刻 n と $n-k$ に依存するため n と k を引数にもつが，通信の問題では信号の統計的性質が時間的に変化しない（あるいは，観測時間内で変化が無視できる）ことが多い．平均値が時間に依存せず（$m_{\mathrm{x}}(n) = m_{\mathrm{x}}$），自己相関関数が時間差 k のみに依存する（$r_{\mathrm{x}}(n,k) = r_{\mathrm{x}}(k)$）ような，言い換えると，平均と自己相関関数が時刻の原点シフト $n \to n+l, {}^\forall l$ に依存しないような，2 次過程は広義定常過程と呼ばれる．さらに，より強い定常性の概念として狭義定常がある．確率過程 $x(n)$ の任意の有限次元分布が時刻の原点シフト $n \to n+l, {}^\forall l$ に対して不変のとき，つまり，

$$p(x(n_1), x(n_2), \cdots, x(n_p)) = p(x(n_1+l), x(n_2+l), \cdots, x(n_p+l)) \tag{2.9}$$

[2] 確率変数は大文字で，その実現値は小文字で表記するのが慣例であるが (例えば，$E[X(n)] = \int x(n)p(x(n))dx(n)$ など)，本稿では，スカラー，ベクトル，行列の区別を小文字，大文字の表記で行っているため，確率変数とその実現値を区別することなく同一の表記を用いることに注意する．

[3] 通信応用を想定して，本稿では確率過程の実現値は複素数とする．

[4] コーシー分布や t-分布など分散が有界でない確率変数もある．

が成り立つとき，$x(n)$ は狭義定常過程と呼ばれる．$m_{\mathrm{x}}(n)$ と $r_{\mathrm{x}}(n,k)$ の定義に代入することで，狭義定常の確率過程は広義定常であることが直ちに確認できる．ただし，その逆は一般に成り立たない．

通信のための信号処理で確率過程を考えるときはそのほとんど場合に広義定常性が仮定されるが，定常でない確率過程（非定常確率過程）の中には統計的性質がある規則をもって時間変化するようなものもあり，そのような確率過程も信号処理の確率モデルとしてしばしば利用される．その代表が，周期定常過程であり，定常性と同様に広義と狭義の定義がある．確率過程 $x(n)$ は，任意の整数 m に対して

$$m_{\mathrm{x}}(n) = m_{\mathrm{x}}(n+mT) \tag{2.10}$$

$$r_{\mathrm{x}}(n,k) = r_{\mathrm{x}}(n+mT, k+mT) \tag{2.11}$$

が成り立つとき，周期 T の広義周期定常過程と呼ばれる．自己相関関数が時間軸 n の方向に周期性をもつためこの軸方向のフーリエ級数展開を考えることができ，これによって定義されるサイクリック自己相関関数の性質を利用することで，2 次の統計量を用いたブラインドシステム同定 [23, 24] や広帯域信号の到来方向推定 [25] などの信号処理が可能になる．こういった研究がその後のマルチチャネル信号処理と呼ばれる分野の発展に繋がっていくが，詳細については本稿の範囲を逸脱するためこれ以上は触れない．周期定常性に興味のある読者は，[22] などを参照されたい．

無線通信のための信号処理では，離散時間確率過程 $x(n)$ の複数の観測からなるベクトル

$$\mathbf{x}(n) = [x(n)\ x(n-1)\ \cdots\ x(n-N+1)]^{\mathrm{T}} \in \mathbb{C}^N \tag{2.12}$$

を用いて何らかの量を推定することが多いため，その性質を理解することは重要である．確率過程 $\mathbf{x}(n)$ の自己相関行列は

$$\mathbf{R}_{\mathrm{x}} = E[\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)] \tag{2.13}$$

と定義される．特に $\mathbf{x}(n)$ が広義定常過程のときには

$$\begin{aligned}
&\mathbf{R}_{\mathrm{x}} \\
&= \begin{bmatrix} E[x(n)x^*(n)] & E[x(n)x^*(n-1)] & \dots & E[x(n)x^*(n-N+1)] \\ E[x(n-1)x^*(n)] & E[x(n-1)x^*(n-1)] & \dots & E[x(n-1)x^*(n-N+1)] \\ \vdots & \vdots & \ddots & \vdots \\ E[x(n-N+1)x^*(n)] & E[x(n-N+1)x^*(n-1)] & \dots & E[x(n-N+1)x^*(n-N+1)] \end{bmatrix} \\
&= \begin{bmatrix} r_{\mathrm{x}}(0) & r_{\mathrm{x}}(1) & \dots & r_{\mathrm{x}}(N-1) \\ r_{\mathrm{x}}(-1) & r_{\mathrm{x}}(0) & \dots & r_{\mathrm{x}}(N-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_{\mathrm{x}}(-N+1) & r_{\mathrm{x}}(-N+2) & \dots & r_{\mathrm{x}}(0) \end{bmatrix}
\end{aligned} \tag{2.14}$$

となり，各対角成分が等しい値をもつテプリッツ行列と呼ばれる構造をもつ．また，主対角成分は全て実数になることに注意する．

自己相関行列は多くの信号処理手法において中心的な役割を果たすため，以下に離散時間広義定常過程の自己相関行列のいくつかの性質をまとめる．なお，別途定義しない限り $E[\mathbf{x}(n)] = \mathbf{0}$ と仮定するが，$x(n)$ が広義定常であることから一般性を失わない．

性質１：離散時間広義定常過程の自己相関行列はエルミート行列

定義より

$$\begin{aligned} \mathbf{R}_{\mathrm{x}}^{\mathrm{H}} &= \left(E[\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)]\right)^{\mathrm{H}} \\ &= E[(\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n))^{\mathrm{H}}] \\ &= E[\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)] \\ &= \mathbf{R}_{\mathrm{x}} \end{aligned}$$

となることから簡単に確認出来る．ここで，$\alpha, \beta \in \mathbb{R}$, $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$ について $(\alpha\mathbf{a} + \beta\mathbf{b})^{\mathrm{H}} = \alpha\mathbf{a}^{\mathrm{H}} + \beta\mathbf{b}^{\mathrm{H}}$ が成り立つことから $(\cdot)^{\mathrm{H}}$ は線形性をもち，同じく線形性をもつ $E[\cdot]$ と交換可能であることを利用している．また，$\mathbf{A} \in \mathbb{C}^{n\times k}$，$\mathbf{B} \in \mathbb{C}^{k\times n}$ について $(\mathbf{AB})^{\mathrm{H}} = \mathbf{B}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}$ が成り立つという性質は容易に確認でき，今後もよく利用する．

成分毎に考えると，

$$\begin{aligned} r_{-k} &= E[x(n)x^*(n+k)] \\ &= (E[x^*(n)x(n+k)])^* \\ &= (E[x^*(n-k)x(n)])^* \\ &= r_k^* \end{aligned}$$

となることから

$$\mathbf{R}_{\mathrm{x}} = E[\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)] = \begin{bmatrix} r_0 & r_1 & \cdots & r_{N-1} \\ r_1^* & r_0 & \cdots & r_{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{N-1}^* & r_{N-2}^* & \cdots & r_0 \end{bmatrix}$$

によって性質 1 を示すことも可能である．

性質２：離散時間広義定常過程の自己相関行列は非負定値

$\mathbf{u} \in \mathbb{C}^N$ を任意の確定的な非零ベクトルとし，$\mathbf{x}(n)$ との内積を $y = \mathbf{u}^{\mathrm{H}}\mathbf{x}(n)$ とする．このとき，

$$\begin{aligned} E[|y|^2] &= E[yy^*] \\ &= E[\mathbf{u}^{\mathrm{H}}\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)\mathbf{u}] \\ &= \mathbf{u}^{\mathrm{H}}E[\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)]\mathbf{u} \\ &= \mathbf{u}^{\mathrm{H}}\mathbf{R}_{\mathrm{x}}\mathbf{u} \end{aligned}$$

となるが，$|y|^2$ は $\mathbf{x}(n)$ の任意の実現値に対して実数で $|y|^2 \geq 0$ であることから，その集合平均も実数で $E[|y|^2] \geq 0$ となる．したがって，上記の性質 2 が示された[5]．

なお，一般のエルミート行列 $\mathbf{A}$ と任意の $\mathbf{u} \in \mathbb{C}^N$ に対して，そのエルミート形式は

$$\begin{aligned}(\mathbf{u}^{\mathrm{H}}\mathbf{A}\mathbf{u})^{\mathrm{H}} &= \mathbf{u}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}\mathbf{u} \\ &= \mathbf{u}^{\mathrm{H}}\mathbf{A}\mathbf{u}\end{aligned}$$

となるので，エルミート形式が実数になることは自己相関行列に限らず一般のエルミート行列がもつ性質であることに注意する．

性質３：離散時間広義定常過程の自己相関行列の固有値は全て非負の実数

自己相関行列 $\mathbf{R}_{\mathrm{x}}$ の固有値を $\lambda_1, \ldots, \lambda_N$ とし，固有値 λ_i に対応する固有ベクトルを $\mathbf{q}_i = [q_{i,1} \cdots q_{i,N}]^{\mathrm{T}} (\neq \mathbf{0})$ とすると

$$\mathbf{R}_{\mathrm{x}}\mathbf{q}_i = \lambda_i \mathbf{q}_i \tag{2.15}$$

が成り立つ．両辺に左から $\mathbf{q}_i^{\mathrm{H}}$ を乗算すると

$$\mathbf{q}_i^{\mathrm{H}}\mathbf{R}_{\mathrm{x}}\mathbf{q}_i = \lambda_i \mathbf{q}_i^{\mathrm{H}}\mathbf{q}_i$$

となり，さらに $\mathbf{q}_i^{\mathrm{H}}\mathbf{q}_i$ は

$$\mathbf{q}_i^{\mathrm{H}}\mathbf{q}_i = \sum_{j=1}^{N} |q_{i,j}|^2$$

より正の実数であることからこれで両辺を割って

$$\lambda_i = \frac{\mathbf{q}_i^{\mathrm{H}}\mathbf{R}_{\mathrm{x}}\mathbf{q}_i}{\mathbf{q}_i^{\mathrm{H}}\mathbf{q}_i} \tag{2.16}$$

を得る．性質 2 より分子の $\mathbf{q}_i^{\mathrm{H}}\mathbf{R}_{\mathrm{x}}\mathbf{q}_i$ は非負の実数なので，λ_i は非負の実数となる．なお，エルミート行列のエルミート形式が実数であることから，一般のエルミート行列の固有値は実数になる (ただし，非負とは限らない) ことにも注意する．

(2.16) の右辺の形の比はベクトル $\mathbf{q}_i$ のレイリー商（Rayleigh quotient）と呼ばれる．したがって，行列 $\mathbf{R}_{\mathrm{x}}$ の固有値 λ_i は対応する固有ベクトル $\mathbf{q}_i$ のレイリー商ということができる．

性質４：離散時間広義定常過程の自己相関行列の異なる固有値に対応する固有ベクトルは直交する

自己相関行列 $\mathbf{R}_{\mathrm{x}}$ の固有値 $\lambda_1, \ldots, \lambda_N$ が全て異なるとし，固有値 λ_i に対応する固有ベクト

[5] $\mathbf{u}^{\mathrm{H}}\mathbf{R}_{\mathrm{x}}\mathbf{u}$ をエルミート形式（実数の場合は 2 次形式）といい，任意の非零ベクトル $\mathbf{u}$ についてエルミート形式が正となる行列を正定値，負となる行列を負定値，非負となる行列を非負定値（半正定値）という．

ルを $\mathbf{q}_i$ とする．(2.15) の両辺に左から $\mathbf{q}_j^{\mathrm{H}}$, $j \neq i$ を乗算することで

$$\mathbf{q}_j^{\mathrm{H}}\mathbf{R}_{\mathrm{x}}\mathbf{q}_i = \lambda_i \mathbf{q}_j^{\mathrm{H}}\mathbf{q}_i \tag{2.17}$$

を得る．一方，

$$\mathbf{R}_{\mathrm{x}}\mathbf{q}_j = \lambda_j \mathbf{q}_j \tag{2.18}$$

の両辺のエルミート転置をとると，性質 1 より $\mathbf{R}_{\mathrm{x}}$ がエルミート行列であり，また性質 3 より λ_j が実数であることから

$$\mathbf{q}_j^{\mathrm{H}}\mathbf{R}_{\mathrm{x}} = \lambda_j \mathbf{q}_j^{\mathrm{H}} \tag{2.19}$$

となる．この両辺に右から $\mathbf{q}_i$ を乗算して

$$\mathbf{q}_j^{\mathrm{H}}\mathbf{R}_{\mathrm{x}}\mathbf{q}_i = \lambda_j \mathbf{q}_j^{\mathrm{H}}\mathbf{q}_i \tag{2.20}$$

を得る．さらに，(2.17) と (2.20) を辺々引き算することで

$$(\lambda_i - \lambda_j)\mathbf{q}_j^{\mathrm{H}}\mathbf{q}_i = 0 \tag{2.21}$$

を得る．仮定より $i \neq j$ に対して $\lambda_i \neq \lambda_j$ なので，

$$\mathbf{q}_j^{\mathrm{H}}\mathbf{q}_i = 0 \quad (i \neq j) \tag{2.22}$$

が示された．

なお，性質 4 を示す際に，自己相関行列 $\mathbf{R}_{\mathrm{x}}$ がエルミート行列であることとその固有値が実数であることしか用いていない．したがって，性質 4 は自己相関行列だけでなく，一般のエルミート行列について成立する性質である．

性質５：離散時間広義定常過程の自己相関行列はユニタリ行列を用いて対角化される

自己相関行列 $\mathbf{R}_{\mathrm{x}}$ の固有値 $\lambda_1, \ldots, \lambda_N$ に対応する固有ベクトルを $\mathbf{q}_1, \ldots, \mathbf{q}_N$ とし，固有ベクトルは正規化されているもの，すなわち，$||\mathbf{q}_i||_2 = 1$ とする．

まず，固有値が全て異なる場合について考える．このとき，全ての i についての (2.15) を並べることで

$$\mathbf{R}_{\mathrm{x}}\mathbf{Q} = \mathbf{Q}\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{bmatrix} \tag{2.23}$$

となる．ただし，

$$\mathbf{Q} = \begin{bmatrix} \mathbf{q}_1 & \ldots & \mathbf{q}_N \end{bmatrix} \tag{2.24}$$

であり，性質４より

$$\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{I} \tag{2.25}$$

が成り立つことから，$\mathbf{Q}$ はユニタリ行列である．(2.23) の両辺に右から $\mathbf{Q}^{\mathrm{H}}$ を乗算することで，

$$\mathbf{Q}^{\mathrm{H}}\mathbf{R}_{\mathrm{x}}\mathbf{Q} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{bmatrix} \tag{2.26}$$

が得られる．

後述する部分空間法で考える状況のように，固有値に同じ値をもつものがある場合（固有値が縮退しているという）は，重複する固有値の数と同じ次元で，そこに含まれる任意のベクトルがその固有値をもつ固有ベクトルになるような部分空間が存在する．この部分空間内の固有ベクトルと他の固有値に対応する固有ベクトルは性質４から直交する．また，同じ固有値をもつ固有ベクトルの部分空間には，いつでも直交基底をとることができる．このため，固有値が縮退している場合にはその部分空間の直交基底を $\mathbf{q}_i$ としてとることでいつでも $\mathbf{Q}$ がユニタリ行列になるようにでき，性質５が示される．ただし，その表現は当然一意的ではない．より詳細な説明は，[21] などを参照されたい（実対称行列で説明されているが，エルミート行列の場合にも同様の議論が成り立つ）．

性質６：相異なる固有値をもつ離散時間広義定常過程の自己相関行列 $\mathbf{R}_{\mathrm{x}}$ の最大固有値 $\lambda_{\max}$ は，最適化問題

$$\lambda_{\max} = \max_{||\mathbf{q}||_2=1} \mathbf{q}^{\mathrm{H}}\mathbf{R}_{\mathrm{x}}\mathbf{q} \tag{2.27}$$

によって得られる．

自己相関行列 $\mathbf{R}_{\mathrm{x}}$ の固有値を $\lambda_1 > \lambda_2 > \cdots > \lambda_N$ とし，対応する正規化固有ベクトルを $\mathbf{q}_1, \cdots, \mathbf{q}_N$ とする．最適化問題の最大化ベクトルを $\mathbf{p}$ とすると，性質４より $\mathbf{q}_1, \cdots, \mathbf{q}_N$ は N 次元ベクトル空間の正規直交基底であることから

$$\mathbf{p} = \sum_{i=1}^{N} \alpha_i \mathbf{q}_i \tag{2.28}$$

と書ける．ここで，性質５より，

$$\begin{aligned} \mathbf{R}_{\mathrm{x}} &= \mathbf{Q} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{bmatrix} \mathbf{Q}^{\mathrm{H}} \\ &= \sum_{i=1}^{N} \lambda_i \mathbf{q}_i \mathbf{q}_i^{\mathrm{H}} \end{aligned} \tag{2.29}$$

となることから

$$\begin{aligned}\mathbf{p}^{\mathrm{H}}\mathbf{R}_{\mathrm{x}}\mathbf{p} &= \mathbf{p}^{\mathrm{H}}\left(\sum_{i=1}^{N}\lambda_i\mathbf{q}_i\mathbf{q}_i^{\mathrm{H}}\right)\mathbf{p} \\ &= \sum_{i=1}^{N}\lambda_i\mathbf{p}^{\mathrm{H}}\mathbf{q}_i\mathbf{q}_i^{\mathrm{H}}\mathbf{p} \end{aligned} \tag{2.30}$$

となる．さらに，性質 4 を用いると

$$\mathbf{p}^{\mathrm{H}}\mathbf{q}_i = \alpha_i^* \tag{2.31}$$

$$\mathbf{q}_i^{\mathrm{H}}\mathbf{p} = \alpha_i \tag{2.32}$$

となるので，

$$\mathbf{p}^{\mathrm{H}}\mathbf{R}_{\mathrm{x}}\mathbf{p} = \sum_{i=1}^{N}\lambda_i|\alpha_i|^2 \tag{2.33}$$

と書ける．ここで，$\lambda_1 > \lambda_2 > \cdots > \lambda_N$ なので，$\mathbf{p}^{\mathrm{H}}\mathbf{R}_{\mathrm{x}}\mathbf{p}$ は

$$\mathbf{p}^{\mathrm{H}}\mathbf{R}_{\mathrm{x}}\mathbf{p} \leq \lambda_1\sum_{i=1}^{N}|\alpha_i|^2 \tag{2.34}$$

と上から押さえられ，等号成立は $i \neq 1$ なる全ての i について $|\alpha_i| = 0$ のときである．また，$\mathbf{p}$ は正規化されているのでこのとき $|\alpha_1| = 1$ となり，最終的に

$$\max_{||\mathbf{q}||_2=1}\mathbf{q}^{\mathrm{H}}\mathbf{R}_{\mathrm{x}}\mathbf{q} = \lambda_1 \tag{2.35}$$

を得る．なお，ここでは簡単のため相異なる固有値をもつ場合について考えたが，固有値が縮退している場合も同様の手順で最大固有値が求まる．また，同様の手順を繰り返すことで最大固有値だけでなく，他の固有値も求めることができる．詳細は [9],[10] などを参照されたい．

2.3　ウィルティンガー微分

工学部の標準的なカリキュラムで学ぶ複素関数論では，複素微分可能な関数（正則関数）について議論されることがほとんどであるが，通信の信号処理では実微分可能であるが複素微分可能ではない関数の最大や最小について考える必要がある．ここでは，そのような関数の複素数，複素ベクトル，複素行列による微分 (ウィルティンガー（Wirtinger）微分) について説明する．

まず，最も簡単な場合として，複素のスカラーを引数にもち，複素数値を取る関数 $f(z)$ について考える．複素微分可能な関数は次で定義される．

正則関数: $D \subseteq \mathbb{C}$ を関数 $f : D \to \mathbb{C}$ の定義域とする．

$$\lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}$$

が任意の $z \in D$ について存在するとき関数 f は領域 D で複素微分可能（正則）であるという．

複素関数論の基本的な結果として，関数 f は，実微分可能[6]でありかつ次のコーシー・リーマンの方程式が成り立つとき，またそのときに限り，複素微分可能であることが知られている．

コーシー・リーマンの方程式:

$$\frac{\partial \Re\{f\}}{\partial x} = \frac{\partial \Im\{f\}}{\partial y}$$
$$\frac{\partial \Re\{f\}}{\partial y} = -\frac{\partial \Im\{f\}}{\partial x}$$

信号処理の問題では電力や 2 乗ユークリッド距離の最小化について議論することが多いが，そのような場合によく現れる関数 $f(z) = |z|^2 = zz^*$ について考えてみる．これを前述の複素微分の定義式に代入すると，

$$\begin{aligned}\lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} &= \lim_{\Delta z \to 0} \frac{|z + \Delta z|^2 - |z|^2}{\Delta z} \\ &= \lim_{\Delta z \to 0} \frac{(z + \Delta z)(z + \Delta z)^* - zz^*}{\Delta z} \\ &= \lim_{\Delta z \to 0} \frac{\Delta z z^* + z(\Delta z)^* + \Delta z(\Delta z)^*}{\Delta z} \end{aligned} \tag{2.36}$$

となるが，Δz をどのように 0 に近づけるかによって値が変わってしまうため，この極限は存在しない．実際，$\Delta z = \Delta x + \mathrm{j}\Delta y$ $(\Delta x, \Delta y \in \mathbb{R})$ として，まず $\Delta x \to 0$ とするとこの値は $z^* - z - \mathrm{j}\Delta y$ となり，さらに $\Delta y \to 0$ とすると $z^* - z$ となる．一方，先に $\Delta y \to 0$ とすると $z^* + z + \Delta x$ となり，さらに $\Delta x \to 0$ とすると $z^* + z$ となる．一般に，$z^* - z \neq z^* + z$ であるため，これは関数 $f(z) = |z|^2 = zz^*$ が複素微分可能でないことを意味する．

このような関数の微分を考える際に有効なのが，z と z^* を独立な変数として扱って関数 f の全微分を考えるというアプローチである．記号 d によって全微分を表すとすると，独立な変数 $x, y \in \mathbb{R}$ を引数にもつ関数 f の全微分は一般に

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy \tag{2.37}$$

[6]複素数 $z = x + \mathrm{j}y$ を引数にもつ関数 $f(z)$ は，これを実数 x, y を引数にもつ関数 $f(x, y)$ とみたときに x, y で微分可能である場合，実微分可能という．

とかける．ただし，$\frac{\partial f}{\partial x}$ は関数 f の変数 x についての偏微分を表す．ここで，

$$z = x + \mathrm{j}y \tag{2.38}$$
$$z^* = x - \mathrm{j}y \tag{2.39}$$

より，

$$dz = dx + \mathrm{j}dy \tag{2.40}$$
$$dz^* = dx - \mathrm{j}dy \tag{2.41}$$

となるので，

$$dx = \frac{1}{2}(dz + dz^*) \tag{2.42}$$
$$dy = \frac{1}{2\mathrm{j}}(dz - dz^*) \tag{2.43}$$

を得る．これを，(2.37) に代入すると

$$\begin{aligned} df &= \frac{\partial f}{\partial x}\frac{dz + dz^*}{2} + \frac{\partial f}{\partial y}\frac{dz - dz^*}{2\mathrm{j}} \\ &= \frac{1}{2}\left(\frac{\partial f}{\partial x} - \mathrm{j}\frac{\partial f}{\partial y}\right)dz + \frac{1}{2}\left(\frac{\partial f}{\partial x} + \mathrm{j}\frac{\partial f}{\partial y}\right)dz^* \end{aligned} \tag{2.44}$$

となる．ここで，f を独立な変数 z, z^* の関数とみると，その全微分は

$$df = \frac{\partial f}{\partial z}dz + \frac{\partial f}{\partial z^*}dz^* \tag{2.45}$$

ともかくことができる．(2.44) と (2.45) を見比べると, 次のように定義すると実数の場合と同様の表現が得られ都合が良いことが分かる．

スカラー値関数のウィルティンガー微分（引数が複素数の場合）：

$$\frac{\partial f}{\partial z} = \frac{1}{2}\left(\frac{\partial f}{\partial x} - \mathrm{j}\frac{\partial f}{\partial y}\right) \tag{2.46}$$
$$\frac{\partial f}{\partial z^*} = \frac{1}{2}\left(\frac{\partial f}{\partial x} + \mathrm{j}\frac{\partial f}{\partial y}\right) \tag{2.47}$$

簡単な例として，$z = x + \mathrm{j}y$ として $f(z) = z$ と $f(z) = z^*$ の場合を考えてみると

$$\frac{\partial z}{\partial z} = \frac{1}{2}\left(\frac{\partial(x+\mathrm{j}y)}{\partial x} - \mathrm{j}\frac{\partial(x+\mathrm{j}y)}{\partial y}\right) = \frac{1}{2}(1 - \mathrm{j}\cdot\mathrm{j}) = 1 \tag{2.48}$$
$$\frac{\partial z^*}{\partial z^*} = \frac{1}{2}\left(\frac{\partial(x-\mathrm{j}y)}{\partial x} + \mathrm{j}\frac{\partial(x-\mathrm{j}y)}{\partial y}\right) = \frac{1}{2}(1 + \mathrm{j}\cdot(-\mathrm{j})) = 1 \tag{2.49}$$
$$\frac{\partial z}{\partial z^*} = \frac{1}{2}\left(\frac{\partial(x+\mathrm{j}y)}{\partial x} + \mathrm{j}\frac{\partial(x+\mathrm{j}y)}{\partial y}\right) = \frac{1}{2}(1 + \mathrm{j}\cdot\mathrm{j}) = 0 \tag{2.50}$$
$$\frac{\partial z^*}{\partial z} = \frac{1}{2}\left(\frac{\partial(x-\mathrm{j}y)}{\partial x} - \mathrm{j}\frac{\partial(x-\mathrm{j}y)}{\partial y}\right) = \frac{1}{2}(1 - \mathrm{j}\cdot(-\mathrm{j})) = 0 \tag{2.51}$$

となる．これより，ウィルティンガー微分を計算する際に z と z^* を独立であると形式的に扱ってよいこと，すなわち $\frac{\partial}{\partial z}$ を評価する際には z^* を形式的に定数として z に関する偏微分を計算し，$\frac{\partial}{\partial z^*}$ を評価する際には z を形式的に定数として z^* に関する偏微分を計算してもよいことが確認される．また，この結果を用いて先ほどの $f(z) = |z|^2$ の場合について考えてみると

$$\frac{\partial |z|^2}{\partial z} = \frac{\partial z z^*}{\partial z} = z^* \tag{2.52}$$

$$\frac{\partial |z|^2}{\partial z^*} = \frac{\partial z z^*}{\partial z^*} = z \tag{2.53}$$

となる．

ウィルティンガー微分の定義を用いると

$$\begin{aligned} \frac{\partial f}{\partial z^*} &= \frac{1}{2}\left(\frac{\partial}{\partial x}\left(\Re\{f\} + \mathrm{j}\Im\{f\}\right) + \mathrm{j}\frac{\partial}{\partial y}\left(\Re\{f\} + \mathrm{j}\Im\{f\}\right)\right) \\ &= \frac{1}{2}\left(\frac{\partial \Re\{f\}}{\partial x} - \frac{\partial \Im\{f\}}{\partial y}\right) + \frac{\mathrm{j}}{2}\left(\frac{\partial \Re\{f\}}{\partial y} + \frac{\partial \Im\{f\}}{\partial x}\right) \end{aligned} \tag{2.54}$$

となるので，コーシー・リーマンの方程式は

$$\frac{\partial f}{\partial z^*} = 0 \tag{2.55}$$

と簡潔にかくことができる．これより，複素微分可能な関数は z^* に依存しないということが分かる．複素数を引数にもつ関数が実数値を取るためには虚部を打ち消す必要があるため，そのような関数は一般に z だけでなく z^* にも依存する．これが，通信の信号処理の問題で考える関数が一般に複素微分可能ではなく，その極値を求めるためにウィルティンガー微分のような手法が必要とされる理由である．

複素数値をとる関数 f の引数が複素ベクトル $\mathbf{z} = [z_1 \cdots z_M]^{\mathrm{T}} \in \mathbb{C}^M$ の場合も，同様にウィルティンガー微分を考えることができる．$z_m = x_m + \mathrm{j}y_m$, $(x_m, y_m \in \mathbb{R})$ とすると，関数 $f(\mathbf{z})$ の全微分は

$$df = \sum_{m=1}^{M}\left(\frac{\partial f}{\partial x_m}dx_m + \frac{\partial f}{\partial y_m}dy_m\right) \tag{2.56}$$

となるが，$dz_m = dx_m + \mathrm{j}dy_m$ 及び $dz_m^* = dx_m - \mathrm{j}dy_m$ を用いると

$$df = \sum_{m=1}^{M}\left\{\frac{1}{2}\left(\frac{\partial f}{\partial x_m} - \mathrm{j}\frac{\partial f}{\partial y_m}\right)dz_m + \frac{1}{2}\left(\frac{\partial f}{\partial x_m} + \mathrm{j}\frac{\partial f}{\partial y_m}\right)dz_m^*\right\} \tag{2.57}$$

と書き直せる．したがって，

$$\begin{aligned} \frac{\partial f}{\partial \mathbf{z}} &= \begin{bmatrix} \frac{\partial f}{\partial z_1} & \cdots & \frac{\partial f}{\partial z_M} \end{bmatrix} \\ d\mathbf{z} &= \begin{bmatrix} dz_1 & \cdots & dz_M \end{bmatrix}^{\mathrm{T}} \end{aligned} \tag{2.58}$$

および

$$\frac{\partial f}{\partial \mathbf{z}^*} = \begin{bmatrix} \frac{\partial f}{\partial z_1^*} & \cdots & \frac{\partial f}{\partial z_M^*} \end{bmatrix}$$
$$d\mathbf{z}^* = \begin{bmatrix} dz_1^* & \cdots & dz_M^* \end{bmatrix}^{\mathrm{T}} \tag{2.59}$$

としたときに

$$df = \frac{\partial f}{\partial \mathbf{z}} d\mathbf{z} + \frac{\partial f}{\partial \mathbf{z}^*} d\mathbf{z}^* \tag{2.60}$$

と比べることで，複素ベクトルを引数にもつスカラー値の関数 f のウィルティンガー微分は次で定義される．

スカラー値関数のウィルティンガー微分（引数が複素ベクトルの場合）：

$$\begin{aligned} \frac{\partial f}{\partial \mathbf{z}} &= \begin{bmatrix} \frac{\partial f}{\partial z_1} & \cdots & \frac{\partial f}{\partial z_M} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{2}\left(\frac{\partial f}{\partial x_1} - \mathrm{j}\frac{\partial f}{\partial y_1}\right) & \cdots & \frac{1}{2}\left(\frac{\partial f}{\partial x_M} - \mathrm{j}\frac{\partial f}{\partial y_M}\right) \end{bmatrix} \end{aligned} \tag{2.61}$$

$$\begin{aligned} \frac{\partial f}{\partial \mathbf{z}^*} &= \begin{bmatrix} \frac{\partial f}{\partial z_1^*} & \cdots & \frac{\partial f}{\partial z_M^*} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{2}\left(\frac{\partial f}{\partial x_1} + \mathrm{j}\frac{\partial f}{\partial y_1}\right) & \cdots & \frac{1}{2}\left(\frac{\partial f}{\partial x_M} + \mathrm{j}\frac{\partial f}{\partial y_M}\right) \end{bmatrix} \end{aligned} \tag{2.62}$$

また，複素ベクトルを引数にもつスカラー値関数 f の複素勾配は

$$\boldsymbol{\nabla} f = \begin{bmatrix} \frac{\partial f}{\partial \mathbf{z}} & \frac{\partial f}{\partial \mathbf{z}^*} \end{bmatrix} \tag{2.63}$$

で定義される．なお，これらの勾配などを行ベクトルと列ベクトルのどちらで定義するかは文献毎に流儀があるので注意すること．本稿では，$\frac{\partial f}{\partial \mathbf{z}}, \frac{\partial f}{\partial \mathbf{z}^*}$ を行ベクトルで定義し，

$$\frac{\partial f}{\partial \mathbf{z}^{\mathrm{H}}} = \left(\frac{\partial f}{\partial \mathbf{z}^*}\right)^{\mathrm{T}} \tag{2.64}$$

とする．

信号処理の教科書では多くの場合，関数 $f(\mathbf{z})$ の勾配ベクトルが

$$\boldsymbol{\nabla}_e f = \begin{bmatrix} \frac{\partial f}{\partial x_1} + \mathrm{j}\frac{\partial f}{\partial y_1} \\ \frac{\partial f}{\partial x_2} + \mathrm{j}\frac{\partial f}{\partial y_2} \\ \vdots \\ \frac{\partial f}{\partial x_M} + \mathrm{j}\frac{\partial f}{\partial y_M} \end{bmatrix} \tag{2.65}$$

で定義されている [9]．これは

$$\boldsymbol{\nabla}_e f = 2\frac{\partial f}{\partial \mathbf{z}^{\mathrm{H}}} \tag{2.66}$$

と書きなおせるが，一般の複素関数 f の勾配の定義としては適切ではない．一般に，(2.60) より

$$\frac{\partial f^*}{\partial \mathbf{z}^*} = \left(\frac{\partial f}{\partial \mathbf{z}}\right)^* \tag{2.67}$$

$$\frac{\partial f}{\partial \mathbf{z}^*} = \left(\frac{\partial f^*}{\partial \mathbf{z}}\right)^* \tag{2.68}$$

が成り立つが，関数 f が実数値をとる特別な場合には，$f = f^*$ より

$$\frac{\partial f}{\partial \mathbf{z}^*} = \left(\frac{\partial f}{\partial \mathbf{z}}\right)^* \tag{2.69}$$

となり，複素勾配は

$$\boldsymbol{\nabla} f = \begin{bmatrix} \left(\frac{\partial f}{\partial \mathbf{z}^*}\right)^* & \frac{\partial f}{\partial \mathbf{z}^*} \end{bmatrix} \tag{2.70}$$

で与えられる．これより，関数 f が実数値を取る場合は，複素勾配が $\boldsymbol{\nabla} f = \mathbf{0}$ となることと，$\frac{\partial f}{\partial \mathbf{z}} = \mathbf{0}$，$\frac{\partial f}{\partial \mathbf{z}^*} = \mathbf{0}$ がいずれも同値であり，(2.65) による勾配の定義の有効性が確認される．しかしながら，関数 f が実数値をとらない場合には一般に (2.69) が成り立たず，(2.65) による定義では不十分である．

最後に，ウィルティンガー微分の具体例をいくつか記す．典型的な通信の信号処理に現れるウィルティンガー微分はそれほど種類が多くないので，計算ルールとして覚えておくとよい．

計算ルール（引数が複素ベクトルの場合）：
$\mathbf{a}$, $\mathbf{A}$ をそれぞれ定係数の列ベクトル，および行列とすると

$$\frac{\partial}{\partial \mathbf{z}^{\mathrm{H}}}\left(\mathbf{z}^{\mathrm{H}}\mathbf{a}\right) = \mathbf{a} \tag{2.71}$$

$$\frac{\partial}{\partial \mathbf{z}^{\mathrm{H}}}\left(\mathbf{z}^{\mathrm{H}}\mathbf{A}\mathbf{z}\right) = \mathbf{A}\mathbf{z} \tag{2.72}$$

これは実ベクトルによる微分の場合と異なっていることに注意する．$\mathbf{x}$ が実数の列ベクトルの場合は，それぞれ

$$\frac{\partial}{\partial \mathbf{x}}\left(\mathbf{x}^{\mathrm{T}}\mathbf{a}\right) = \mathbf{a} \tag{2.73}$$

$$\frac{\partial}{\partial \mathbf{x}}\left(\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}\right) = \mathbf{A}\mathbf{x} + \mathbf{A}^{\mathrm{T}}\mathbf{x} \tag{2.74}$$

となる．実数のベクトル，行列による微分については，[20] が工学者向けに書かれてありお勧めである．

一方，引数 $\mathbf{Z}$ が複素行列の場合は次の通りである．

計算ルール（引数が複素行列の場合）：

$$\frac{\partial}{\partial \mathbf{Z}^{\mathrm{H}}}\left(\mathrm{tr}\left\{\mathbf{Z}^{\mathrm{H}}\mathbf{A}\right\}\right) = \mathbf{A} \tag{2.75}$$

$$\frac{\partial}{\partial \mathbf{Z}^{\mathrm{H}}}\left(\mathrm{tr}\left\{\mathbf{Z}^{\mathrm{H}}\mathbf{A}\mathbf{Z}\right\}\right) = \mathbf{A}\mathbf{Z} \tag{2.76}$$

ウィルティンガー微分や複素信号の解析に関する説明は，以前は [8] などの一部の解析学の教科書や [9, 11] などの信号処理の教科書の付録にしか見られなかったが，最近では [15, 16] などの優れた文献がある．ここでは，スカラー値の関数しか取り扱わなかったが，より一般のベクトル値や行列値をとる関数についてのウィルティンガー微分の詳細などについてはこれらの文献を参照されたい．また，上述のように無線通信の信号処理の文脈では非正則関数を取り扱うことがほとんどであるが，やはり複素関数論の主役は等角写像や高階微分，一致の定理，解析接続といった単なる実 2 変数関数には無い興味深い性質をもつ正則関数である[7]．これを機会に複素関数論について復習されたい方には，初学者にも大変分かりやすく書かれている [19] がお勧めである．

[7] 最大電力制約下で通信路容量を達成する分布が離散分布であることを証明する際に，複素関数論の一致の定理が利用されている例もあり [17, 18]，正則関数に関する知見が通信や信号処理に役に立たないという訳では無い．

第3章　種々の信号処理手法

ここでは，素朴な問題設定に対する基本的かつ典型的な信号処理手法について説明する．

3.1　問題設定

本稿では，興味のある未知ベクトル $\mathbf{x} = [x_1 \ \cdots \ x_N]^{\mathrm{T}} \in \mathbb{C}^N$ に対して，既知の行列 $\mathbf{A} = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_N] \in \mathbb{C}^{M\times N}$ を用いて線形観測を行なったときに，得られる観測ベクトル $\mathbf{y} = [y_1 \ \cdots \ y_M]^{\mathrm{T}} \in \mathbb{C}^M$ から $\mathbf{x}$ を推定する問題を考える．すなわち，

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v} \tag{3.1}$$

において，$\mathbf{y}$ と $\mathbf{A}$ から $\mathbf{x}$ を推定する問題である．ここで，$\mathbf{v} = [v_1 \ \cdots \ v_M]^{\mathrm{T}} \in \mathbb{C}^M$ は加法性白色雑音ベクトルとする．ただし，信号処理の手法によっては目的が $\mathbf{A}$ の推定であったり，複素数でない成分が仮定されたりするため，注意されたい．また，各手法によってこれらのベクトルを確率的に取扱う場合や確定的に取扱う場合があるが，確率的に扱う場合には，$\mathbf{y}$，$\mathbf{x}$，$\mathbf{v}$ の相関行列をそれぞれ，$\mathbf{R}_{\mathrm{y}} = E[\mathbf{y}\mathbf{y}^{\mathrm{H}}]$，$\mathbf{R}_{\mathrm{x}} = E[\mathbf{x}\mathbf{x}^{\mathrm{H}}]$，$\mathbf{R}_{\mathrm{v}} = E[\mathbf{v}\mathbf{v}^{\mathrm{H}}] = \sigma_{\mathrm{v}}^2\mathbf{I}$ とする．

3.2　ZF 推定

最もシンプルな推定法として，ZF 推定がある．これは，雑音が無い場合に誤差無く $\mathbf{x}$ を推定できる手法であり，$\mathbf{x}$ の推定値は

$$\begin{aligned}\hat{\mathbf{x}}_{\mathrm{zf}} &= \mathbf{W}_{\mathrm{zf}}^{\mathrm{H}}\mathbf{y} \\ &= \mathbf{x} + \mathbf{W}_{\mathrm{zf}}^{\mathrm{H}}\mathbf{v}\end{aligned} \tag{3.2}$$

で与えられる[1]．ここで $\mathbf{W}_{\mathrm{zf}}^{\mathrm{H}}$ は ZF 推定を与える $N \times M$ の重み行列であり，上述の ZF 推定の性質から

$$\mathbf{W}_{\mathrm{zf}}^{\mathrm{H}}\mathbf{A} = \mathbf{I} \tag{3.3}$$

[1] このように，観測ベクトル $\mathbf{y}$ に対して行列を乗算 (より正確には，さらにそれに定数ベクトルを加算) することで推定値を得る手法を一般に線形推定という．ここで，推定値を得るために乗算される行列として $\mathbf{W}_{\mathrm{zf}}$ ではなく $\mathbf{W}_{\mathrm{zf}}^{\mathrm{H}}$ を用いていることに違和感を覚えたかもしれない．もちろん，$\mathbf{W}_{\mathrm{zf}}$ を用いても良いが，その場合，以下の計算過程で $\mathbf{W}_{\mathrm{zf}}$，$\mathbf{W}_{\mathrm{zf}}^{*}$，$\mathbf{W}_{\mathrm{zf}}^{\mathrm{T}}$，$\mathbf{W}_{\mathrm{zf}}^{\mathrm{H}}$ の全てのパターンが出てきてしまう．一方，$\mathbf{W}_{\mathrm{zf}}^{\mathrm{H}}$ としておくと，$\mathbf{W}_{\mathrm{zf}}$，$\mathbf{W}_{\mathrm{zf}}^{\mathrm{H}}$ の 2 種類しか出てこず，表記がシンプルになる．また，特別な場合として重み行列がベクトルとなる場合にその共役転置が乗算されることにすると，ベクトルを全て列ベクトルで定義するというルールにも適合する．つまり，単に表記上の理由で $\mathbf{W}_{\mathrm{zf}}^{\mathrm{H}}$ を用いているのであって，それ以上の特別な理由はない．

を満足する．従って，$\mathbf{A}$ が正則な正方行列 $(M = N)$ の場合には

$$\mathbf{W}_{\mathrm{zf}}^{\mathrm{H}} = \mathbf{A}^{-1} \tag{3.4}$$

と一意に定まる．

一方，$\mathbf{A}$ が $M > N$ の場合には，一般に (3.3) から $\mathbf{W}_{\mathrm{zf}}$ が一意に定まらないため，(3.3) を満足する重み行列のうちで推定値 (3.2) の信号対雑音電力比 (SNR: signal-to-noise power ratio) を最大にするものを $\mathbf{W}_{\mathrm{zf}}$ として選ぶことにする．ここで，(3.2) の信号成分は重み行列に依存せず，また，雑音成分の電力は

$$\begin{aligned} E\left[(\mathbf{W}_{\mathrm{zf}}^{\mathrm{H}}\mathbf{v})^{\mathrm{H}}\mathbf{W}_{\mathrm{zf}}^{\mathrm{H}}\mathbf{v}\right] &= \mathrm{tr}\left\{\mathbf{W}_{\mathrm{zf}}^{\mathrm{H}}E[\mathbf{v}\mathbf{v}^{\mathrm{H}}]\mathbf{W}_{\mathrm{zf}}\right\} \\ &= \sigma_{\mathrm{v}}^2\mathrm{tr}\left\{\mathbf{W}_{\mathrm{zf}}^{\mathrm{H}}\mathbf{W}_{\mathrm{zf}}\right\} \end{aligned} \tag{3.5}$$

で与えられることから，SNR を最大にする $\mathbf{W}_{\mathrm{zf}}$ は

$$\mathbf{W}_{\mathrm{zf}} = \arg\min_{\mathbf{W}\in\mathbb{C}^{M\times N}} \mathrm{tr}\left\{\mathbf{W}^{\mathrm{H}}\mathbf{W}\right\} \quad s.t. \;\; \mathbf{W}^{\mathrm{H}}\mathbf{A} = \mathbf{I} \tag{3.6}$$

なる最適化問題を解くことで得られる[2]．ここで，雑音成分の電力の計算の際に，行列のトレースの性質 $\mathrm{tr}\{\mathbf{AB}\} = \mathrm{tr}\{\mathbf{BA}\}$ を用いて，期待値内の行列の乗算の順番を入れ替えていることに注意する．このテクニックは信号処理で頻繁に用いられており，これを利用することで確定的な重み行列を期待値操作の外に出すことができ，コスト関数の重み行列に関するウィルティンガー微分が計算可能になる．

ラグランジュ関数を

$$\begin{aligned} \mathcal{L}_{\mathrm{zf}}(\mathbf{W}) &= \mathrm{tr}\left\{\mathbf{W}^{\mathrm{H}}\mathbf{W}\right\} + \sum_{n=1}^{N}\boldsymbol{\phi}_n^{\mathrm{H}}(\mathbf{W}^{\mathrm{H}}\mathbf{a}_n - \mathbf{e}_n) \\ &= \mathrm{tr}\left\{\mathbf{W}^{\mathrm{H}}\mathbf{W}\right\} + \sum_{n=1}^{N}\mathrm{tr}\left\{(\mathbf{W}^{\mathrm{H}}\mathbf{a}_n - \mathbf{e}_n)\boldsymbol{\phi}_n^{\mathrm{H}}\right\} \end{aligned} \tag{3.7}$$

と定義すると

$$\begin{aligned} \frac{\partial\mathcal{L}_{\mathrm{zf}}(\mathbf{W})}{\partial\mathbf{W}^{\mathrm{H}}} &= \mathbf{W} + \sum_{n=1}^{N}\mathbf{a}_n\boldsymbol{\phi}_n^{\mathrm{H}} \\ &= \mathbf{W} + \mathbf{A}\boldsymbol{\Phi}^{\mathrm{H}} \end{aligned} \tag{3.8}$$

となる．ただし，$\boldsymbol{\phi}_n$ はラグランジュ乗数からなる N 次元列ベクトルであり，$\mathbf{e}_n$ は n 番目の成分のみが 1 で他が 0 の N 次元列ベクトル，$\boldsymbol{\Phi} = [\boldsymbol{\phi}_1 \;\cdots\; \boldsymbol{\phi}_N]$ である．さらに，

$$\frac{\partial\mathcal{L}_{\mathrm{zf}}(\mathbf{W})}{\partial\mathbf{W}^{\mathrm{H}}} = \mathbf{0} \tag{3.9}$$

---

[2] 例えば，$\mathbf{w}_{\mathrm{opt}} = \arg\min_{\mathbf{w}} J(\mathbf{w})$ $s.t.$ $\mathbf{f}(\mathbf{w}) = \mathbf{0}$ と書いたとき，arg は次の関数の引数 (ここでは min の引数 $\mathbf{w}$)，$\min_{\mathbf{w}} J(\mathbf{w})$ は $J(\mathbf{w})$ を $\mathbf{w}$ について最小化，$s.t.$ は制約条件，をそれぞれ意味するので，これは「$\mathbf{f}(\mathbf{w}) = \mathbf{0}$ という制約条件を満足する $\mathbf{w}$ のうちでコスト関数 $J(\mathbf{w})$ を最小にするものを最適解 $\mathbf{w}_{\mathrm{opt}}$ としなさい」という意味となる．

を解くと

$$\mathbf{W}_{\mathrm{zf}} = -\mathbf{A}\mathbf{\Phi}^{\mathrm{H}} \tag{3.10}$$

となり，これを (3.6) の制約式に代入して，$\mathbf{A}$ が列フルランクと仮定すると

$$\mathbf{\Phi} = -(\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1} \tag{3.11}$$

を得る．従って，最終的に

$$\mathbf{W}_{\mathrm{zf}}^{\mathrm{H}} = (\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}} \tag{3.12}$$

を得る．この式で $M = N$ とすると $\mathbf{W}_{\mathrm{zf}}^{\mathrm{H}} = \mathbf{A}^{-1}$ となり，(3.4) に一致することから，これは $M = N$ の場合にも有効である．

ZF 推定は文脈によって最小 2 乗推定とも呼ばれる．これは

$$\hat{\mathbf{x}}_{\mathrm{ls}} = \arg\min_{\mathbf{x}\in\mathbb{C}^N} ||\mathbf{A}\mathbf{x} - \mathbf{y}||_2^2 \tag{3.13}$$

なる最適化問題を考えると，

$$\begin{aligned} ||\mathbf{A}\mathbf{x} - \mathbf{y}||_2^2 &= (\mathbf{A}\mathbf{x} - \mathbf{y})^{\mathrm{H}}(\mathbf{A}\mathbf{x} - \mathbf{y}) \\ &= \mathbf{x}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}\mathbf{A}\mathbf{x} - \mathbf{x}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}\mathbf{y} - \mathbf{y}^{\mathrm{H}}\mathbf{A}\mathbf{x} + \mathbf{y}^{\mathrm{H}}\mathbf{y} \end{aligned}$$

となることから

$$\frac{\partial}{\partial \mathbf{x}^{\mathrm{H}}} ||\mathbf{A}\mathbf{x} - \mathbf{y}||_2^2 = \mathbf{A}^{\mathrm{H}}\mathbf{A}\mathbf{x} - \mathbf{A}^{\mathrm{H}}\mathbf{y} = \mathbf{0} \tag{3.14}$$

を解くことで，ZF 推定と同じ推定値

$$\hat{\mathbf{x}}_{\mathrm{ls}} = (\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{y} \tag{3.15}$$

が得られるからである．

ZF 推定は希望信号成分を忠実に再現することを優先した推定法であるため，(3.6) のように雑音電力を抑えるような重み行列を選択していても，$\mathbf{A}$ の性質によっては雑音成分の影響で推定値が大きく劣化してしまうことがある．この問題は，ZF 推定の雑音強調 (noise enhancement) として知られている．ZF 推定値の雑音成分の電力は，(3.5) に (3.12) を代入することで

$$\sigma_{\mathrm{v}}^2 \mathrm{tr}\left\{\mathbf{W}_{\mathrm{zf}}^{\mathrm{H}}\mathbf{W}_{\mathrm{zf}}\right\} = \sigma_{\mathrm{v}}^2 \mathrm{tr}\left\{(\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\right\} \tag{3.16}$$

となる．$\mathbf{A}$ が列フルランクとすると，その特異値分解は

$$\mathbf{A} = \mathbf{U}\begin{bmatrix} \mathbf{\Xi} \\ \mathbf{0}_{(M-N)\times N} \end{bmatrix}\mathbf{V}^{\mathrm{H}} \tag{3.17}$$

で与えられる．ただし，$\mathbf{U} \in \mathbb{C}^{M \times M}$, $\mathbf{V} \in \mathbb{C}^{N \times N}$ はユニタリ行列，$\mathbf{\Xi}$ は $\mathbf{A}$ の特異値（すべて非零）からなる対角行列 $\mathbf{\Xi} = \mathrm{diag}[\xi_1 \ \cdots \ \xi_N]$ である．このとき，

$$\mathbf{A}^{\mathrm{H}}\mathbf{A} = \mathbf{V}\mathbf{\Xi}^2\mathbf{V}^{\mathrm{H}} \tag{3.18}$$

となるので，

$$(\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1} = \mathbf{V}\mathbf{\Xi}^{-2}\mathbf{V}^{\mathrm{H}} \tag{3.19}$$

である．従って，雑音電力は

$$\sigma_{\mathrm{v}}^2 \mathrm{tr}\left\{\mathbf{W}_{\mathrm{zf}}^{\mathrm{H}}\mathbf{W}_{\mathrm{zf}}\right\} = \sigma_{\mathrm{v}}^2 \sum_{n=1}^{N} \frac{1}{|\xi_n|^2} \tag{3.20}$$

となる．これより，$\mathbf{A}$ の特異値 ξ_n に一つでも（0ではないが）0に非常に近い値のものがあれば，推定値中の雑音電力が非常に大きくなってしまうことが分かる．

3.3 最小ノルム解

$M < N$ の場合には，(3.3) を満足する $\mathbf{W}_{\mathrm{zf}}$ が存在しないため ZF 推定が行なえない．これは，$M < N$ の場合には $\mathbf{y} = \mathbf{A}\mathbf{x}$ を満足する $\mathbf{x}$ が無限に存在し，雑音が無い場合であっても線形観測された真の $\mathbf{x}$ を $\mathbf{y}$ と $\mathbf{A}$ から特定できないためである[3]．このような場合によく用いられるのが，正則化と呼ばれるアプローチである．正則化では $\mathbf{y} = \mathbf{A}\mathbf{x}$ を満足する $\mathbf{x}$ のうちそのノルムが最小のものを解として選択する（最小ノルム解）．例えば，ノルムとして ℓ_2-ノルムを採用すると，最小ノルム解は

$$\hat{\mathbf{x}}_{\mathrm{mn}} = \arg\min_{\mathbf{x} \in \mathbb{C}^N} ||\mathbf{x}||_2^2 \quad s.t. \quad \mathbf{y} = \mathbf{A}\mathbf{x} \tag{3.21}$$

を解くことで与えられる．ラグランジュ関数を

$$\mathcal{L}_{\mathrm{mn}}(\mathbf{x}) = ||\mathbf{x}||_2^2 + (\mathbf{A}\mathbf{x} - \mathbf{y})^{\mathrm{H}}\boldsymbol{\phi} \tag{3.22}$$

とすると，$\frac{\partial \mathcal{L}_{\mathrm{mn}}(\mathbf{x})}{\partial \mathbf{x}^{\mathrm{H}}} = \mathbf{0}$ より

$$\hat{\mathbf{x}}_{\mathrm{mn}} = -\mathbf{A}^{\mathrm{H}}\boldsymbol{\phi} \tag{3.23}$$

となり，これを制約式に代入して，$\mathbf{A}$ が行フルランクと仮定すると

$$\boldsymbol{\phi} = -(\mathbf{A}\mathbf{A}^{\mathrm{H}})^{-1}\mathbf{y} \tag{3.24}$$

となる．これより，

$$\hat{\mathbf{x}}_{\mathrm{mn}} = \mathbf{A}^{\mathrm{H}}(\mathbf{A}\mathbf{A}^{\mathrm{H}})^{-1}\mathbf{y} \tag{3.25}$$

を得る．

[3] このような連立方程式を劣決定という．

3.4 MMSE 推定

ZF 推定はノイズの影響を直接考慮していないため雑音強調という問題があったが，これに対処するために MMSE 推定が広く用いられている．MMSE 推定には，非線形演算も許す一般の MMSE 推定と線形演算のみによる線形 MMSE 推定があり，よく用いられるのは後者であるが，ここでは一般の MMSE 推定についても簡単に説明する．

一般の MMSE 推定において，ベクトル値の関数 $\mathbf{f}$ を用いて

$$\hat{\mathbf{x}}_{\mathrm{mmse}} = \mathbf{f}(\mathbf{y}) \tag{3.26}$$

によって推定値を与えるとすると，$\mathbf{f}$ は 2 乗誤差の条件付き期待値

$$J_{\mathrm{mmse}}[\mathbf{f}] = E\left[||\mathbf{f}(\mathbf{y}) - \mathbf{x}||_2^2 \middle| \mathbf{y} \right] \tag{3.27}$$

を最小にするように決定される．ここで，$\mathbf{y}$ が与えられたときの $\mathbf{x}$ の条件付き確率密度関数を $p(\mathbf{x}|\mathbf{y})$ とし，$\mathbf{y}$ に関する $\mathbf{x}$ の条件付き期待値を

$$\bar{\mathbf{x}}(\mathbf{y}) = E[\mathbf{x}|\mathbf{y}] = \int \mathbf{x} p(\mathbf{x}|\mathbf{y}) d\mathbf{x} \tag{3.28}$$

と定義すると[4],

$$\begin{aligned} J_{\mathrm{mmse}}[\mathbf{f}] =& E\left[||\mathbf{f}(\mathbf{y}) - \mathbf{x}||_2^2 | \mathbf{y} \right] \\ =& E\left[||\mathbf{f}(\mathbf{y}) - \bar{\mathbf{x}}(\mathbf{y}) + \bar{\mathbf{x}}(\mathbf{y}) - \mathbf{x}||_2^2 | \mathbf{y} \right] \\ =& ||\mathbf{f}(\mathbf{y}) - \bar{\mathbf{x}}(\mathbf{y})||_2^2 + E\left[||\bar{\mathbf{x}}(\mathbf{y}) - \mathbf{x}||_2^2 | \mathbf{y} \right] \\ &+ \{\mathbf{f}(\mathbf{y}) - \bar{\mathbf{x}}(\mathbf{y})\}^{\mathrm{H}} E\left[\{\bar{\mathbf{x}}(\mathbf{y}) - \mathbf{x}\} | \mathbf{y} \right] \\ &+ E\left[\{\bar{\mathbf{x}}(\mathbf{y}) - \mathbf{x}\}^{\mathrm{H}} | \mathbf{y} \right] \{\mathbf{f}(\mathbf{y}) - \bar{\mathbf{x}}(\mathbf{y})\} \\ =& ||\mathbf{f}(\mathbf{y}) - \bar{\mathbf{x}}(\mathbf{y})||_2^2 + E\left[||\bar{\mathbf{x}}(\mathbf{y}) - \mathbf{x}||_2^2 | \mathbf{y} \right] \\ \geq& E\left[||\bar{\mathbf{x}}(\mathbf{y}) - \mathbf{x}||_2^2 | \mathbf{y} \right] \end{aligned} \tag{3.29}$$

となり，等号成立は $\mathbf{f}(\mathbf{y}) = \bar{\mathbf{x}}(\mathbf{y})$ のときである．従って，一般の MMSE 推定は

$$\hat{\mathbf{x}}_{\mathrm{mmse}} = \bar{\mathbf{x}}(\mathbf{y}) \tag{3.30}$$

によって与えられる．この推定値は誤差の分散を最小にすることから最小分散推定値とも呼ばれる．

次に，線形推定を行うという制約下での MMSE 推定について考える．各ベクトルの平均値が $\mathbf{0}$ のとき，線形 MMSE 推定は行列 $\mathbf{W}_{\mathrm{lmmse}}$ を用いて

$$\hat{\mathbf{x}}_{\mathrm{lmmse}} = \mathbf{W}_{\mathrm{lmmse}}^{\mathrm{H}} \mathbf{y} \tag{3.31}$$

で与えられる．ここで，$\mathbf{W}_{\mathrm{lmmse}}$ は

$$\mathbf{W}_{\mathrm{lmmse}} = \arg \min_{\mathbf{W} \in \mathbb{C}^{M \times N}} E\left[||\mathbf{W}^{\mathrm{H}}(\mathbf{A}\mathbf{x} + \mathbf{v}) - \mathbf{x}||_2^2 \right] \tag{3.32}$$

[4] $\mathbf{x}$ が離散確率変数のときは確率測度を用いた式に適宜修正する．

によって決定される．

$$
\begin{aligned}
J_{\mathrm{lmmse}}(\mathbf{W}) &= E\left[||\mathbf{W}^{\mathrm{H}}(\mathbf{A}\mathbf{x}+\mathbf{v})-\mathbf{x}||_2^2\right] \\
&= E\left[(\mathbf{W}^{\mathrm{H}}\mathbf{A}\mathbf{x}+\mathbf{W}^{\mathrm{H}}\mathbf{v}-\mathbf{x})^{\mathrm{H}}(\mathbf{W}^{\mathrm{H}}\mathbf{A}\mathbf{x}+\mathbf{W}^{\mathrm{H}}\mathbf{v}-\mathbf{x})\right] \\
&= E\left[\mathrm{tr}\{(\mathbf{W}^{\mathrm{H}}\mathbf{A}\mathbf{x}+\mathbf{W}^{\mathrm{H}}\mathbf{v}-\mathbf{x})(\mathbf{W}^{\mathrm{H}}\mathbf{A}\mathbf{x}+\mathbf{W}^{\mathrm{H}}\mathbf{v}-\mathbf{x})^{\mathrm{H}}\}\right] \\
&= \mathrm{tr}\left\{\mathbf{W}^{\mathrm{H}}\mathbf{A}E[\mathbf{x}\mathbf{x}^{\mathrm{H}}]\mathbf{A}^{\mathrm{H}}\mathbf{W}\right\} + \mathrm{tr}\left\{\mathbf{W}^{\mathrm{H}}\mathbf{A}E[\mathbf{x}\mathbf{v}^{\mathrm{H}}]\mathbf{W}\right\} \\
&\quad - \mathrm{tr}\left\{\mathbf{W}^{\mathrm{H}}\mathbf{A}E[\mathbf{x}\mathbf{x}^{\mathrm{H}}]\right\} + \mathrm{tr}\left\{\mathbf{W}^{\mathrm{H}}E[\mathbf{v}\mathbf{x}^{\mathrm{H}}]\mathbf{A}^{\mathrm{H}}\mathbf{W}\right\} \\
&\quad + \mathrm{tr}\left\{\mathbf{W}^{\mathrm{H}}E[\mathbf{v}\mathbf{v}^{\mathrm{H}}]\mathbf{W}\right\} - \mathrm{tr}\left\{\mathbf{W}^{\mathrm{H}}E[\mathbf{v}\mathbf{x}^{\mathrm{H}}]\right\} \\
&\quad - \mathrm{tr}\left\{E[\mathbf{x}\mathbf{x}^{\mathrm{H}}]\mathbf{A}^{\mathrm{H}}\mathbf{W}\right\} - \mathrm{tr}\left\{E[\mathbf{x}\mathbf{v}^{\mathrm{H}}]\mathbf{W}\right\} + \mathrm{tr}\left\{E[\mathbf{x}\mathbf{x}^{\mathrm{H}}]\right\} \\
&= \mathrm{tr}\left\{\mathbf{W}^{\mathrm{H}}\mathbf{A}\mathbf{R}_{\mathrm{x}}\mathbf{A}^{\mathrm{H}}\mathbf{W}\right\} - \mathrm{tr}\left\{\mathbf{W}^{\mathrm{H}}\mathbf{A}\mathbf{R}_{\mathrm{x}}\right\} + \sigma_{\mathrm{v}}^2\mathrm{tr}\left\{\mathbf{W}^{\mathrm{H}}\mathbf{W}\right\} \\
&\quad - \mathrm{tr}\left\{\mathbf{R}_{\mathrm{x}}\mathbf{A}^{\mathrm{H}}\mathbf{W}\right\} + \mathrm{tr}\left\{\mathbf{R}_{\mathrm{x}}\right\}
\end{aligned} \tag{3.33}
$$

となることから，

$$
\frac{\partial J_{\mathrm{lmmse}}(\mathbf{W})}{\partial \mathbf{W}^{\mathrm{H}}} = \mathbf{A}\mathbf{R}_{\mathrm{x}}\mathbf{A}^{\mathrm{H}}\mathbf{W} - \mathbf{A}\mathbf{R}_{\mathrm{x}} + \sigma_{\mathrm{v}}^2\mathbf{W} = \mathbf{0} \tag{3.34}
$$

を解いて

$$
\mathbf{W}_{\mathrm{lmmse}}^{\mathrm{H}} = \mathbf{R}_{\mathrm{x}}\mathbf{A}^{\mathrm{H}}\left(\mathbf{A}\mathbf{R}_{\mathrm{x}}\mathbf{A}^{\mathrm{H}} + \sigma_{\mathrm{v}}^2\mathbf{I}\right)^{-1} \tag{3.35}
$$

を得る．

一般に線形 MMSE 推定値は最小分散推定値（一般の MMSE 推定値）とは一致しないが，$(\mathbf{x},\mathbf{y})$ の結合分布がガウス分布の場合にはこれらが一致する [1].

3.5　減算型干渉除去

ここまでに述べた線形の信号検出法は，観測ベクトル $\mathbf{y}$ に何らかの行列を一度乗算することで $\mathbf{x}$ の推定値を得るものであった．$\mathbf{x}$ の各成分にとってそれ以外の $\mathbf{x}$ の成分は干渉であるとみると，何らかの方法（通常，線形推定）で得られた仮の推定値を用いてその干渉成分のレプリカを生成して推定値から干渉成分を減算することで，より信頼度の高い推定値を得られる可能性がある．このようなアイデアで繰り返し処理によって信号検出を行う方法が減算型干渉除去 (subtractive interference cancellation) である．情報理論の分野ではかなり前から議論されていたが [26, 27]，通信のための信号処理の文脈で広く知られるようになったのは符号分割多元接続 (CDMA: code division multiple access) のマルチユーザ検出法 [28] としてであると思われる．減算型干渉除去は大きく 2 つのアプローチに分けられ，それぞれ逐次干渉除去 (SIC: successive interference cancellation)，並列干渉除去 (PIC: parallel interference cancellation) と呼ばれる．

SIC は，$\mathbf{x}$ の成分毎の SNR が大きく異なる場合，例えば $\mathbf{x}$ と $\mathbf{v}$ の成分毎の電力が同一として行列 $\mathbf{A}$ の各列のノルムが大きく異なる場合など，に効果的な信号検出法であり，観測信号電力が大きい $\mathbf{x}$ の成分から順に検出を行う点に特徴がある．具体的には，簡単のため観測ベクトル $\mathbf{y}$ における $\mathbf{x}$ の成分毎の SNR が $x_1,\ x_2,\ldots,x_N$ の順に小さくなっていくとす

ると，まず，$x_2,\dots,x_N$ を全て雑音成分であるとみなして何らかの方法で x_1 の推定を行う．すなわち，

$$\begin{aligned}\mathbf{y}_{\mathrm{sic}}^{(1)} &= \mathbf{y}\\ &= \mathbf{A}\mathbf{x}+\mathbf{v}\\ &= \mathbf{a}_1 x_1 + \left(\sum_{i=2}^{N}\mathbf{a}_i x_i + \mathbf{v}\right)\end{aligned} \tag{3.36}$$

として3行目の右辺第2項を雑音とみなし，例えば線形推定を行うとすると x_1 の推定値 $\hat{x}_{\mathrm{sic},1}$ は

$$\hat{x}_{\mathrm{sic},1} = \mathbf{w}_1^{\mathrm{H}}\mathbf{y}_{\mathrm{sic}}^{(1)} \tag{3.37}$$

と書かれる．ただし，$\mathbf{w}_1 \in \mathbb{C}^N$ は線形推定の重みベクトルであり，ZF 基準や MMSE 基準などで決定される．

ここまでの処理は通常の線形推定と違いが無いが，次に $\mathbf{y}$ から x_1 の成分をキャンセルした観測ベクトル $\mathbf{y}_{\mathrm{sic}}^{(2)}$ を生成する．具体的には，推定値 $\hat{x}_{\mathrm{sic},1}$ が正しいと信じて，

$$\begin{aligned}\mathbf{y}_{\mathrm{sic}}^{(2)} &= \mathbf{y} - \mathbf{a}_1\hat{x}_{\mathrm{sic},1}\\ &= \mathbf{a}_2 x_2 + \left(\sum_{i=3}^{N}\mathbf{a}_i x_i + \mathbf{v}\right) + \mathbf{a}_1(x_1 - \hat{x}_{\mathrm{sic},1})\\ &\approx \mathbf{a}_2 x_2 + \left(\sum_{i=3}^{N}\mathbf{a}_i x_i + \mathbf{v}\right)\end{aligned} \tag{3.38}$$

とする．ここで，実際に $\hat{x}_{\mathrm{sic},1} = x_1$ であれば $\mathbf{y}_{\mathrm{sic}}^{(2)}$ は3行目の右辺に等しくなり，$\mathbf{y}$ から直接 x_2 を検出するよりも $\mathbf{y}_{\mathrm{sic}}^{(2)}$ から検出する方が x_1 による干渉がなくなった分だけ特性が改善する．x_2 を $\mathbf{w}_2 \in \mathbb{C}^N$ を用いて線形推定する場合は，

$$\hat{x}_{\mathrm{sic},2} = \mathbf{w}_2^{\mathrm{H}}\mathbf{y}_{\mathrm{sic}}^{(2)} \tag{3.39}$$

である．以上の処理を繰り返して，x_N の推定値まで求めるのが SIC の基本的な手順である．観測における SNR が最も大きな信号成分にとっては古典的な検出法に比べてメリットはないが，観測 SNR が小さい信号ほどメリットが大きいという点が重要である．ここでは最も簡単な手法として成分毎の推定値をそのまま用いて干渉成分を除去したが判定値を利用することも可能であり，行列 $\mathbf{A}$ についての情報が正確である場合には硬判定値を用いることで特性を改善できることが知られている [28]．また，SIC は準最適な信号検出法ではあるが，ガウス放送通信路（正確には劣化型放送通信路）において重畳符号化と組み合わせて用いることで通信路容量領域の境界を達成できることが知られており，情報理論の文脈でも重要な手法である [27]．

一方，PIC は $\mathbf{x}$ の成分毎の推定値ではなく，ベクトル $\mathbf{x}$ 全体の仮推定値を用いて干渉を除去する方法である．具体的には，線形推定などの方法を用いて観測ベクトル $\mathbf{y}_{\mathrm{pic}}^{(1)} = \mathbf{y}$ か

ら $\mathbf{x}$ の仮推定値 $\hat{\mathbf{x}}_{\mathrm{pic}}^{(1)}$ が得られたとすると，

$$\begin{aligned}\mathbf{y}_{\mathrm{pic}}^{(2)} &= \mathbf{y} - \mathbf{A}_{\mathrm{offdiag}}\hat{\mathbf{x}}_{\mathrm{pic}}^{(1)} \\ &= \mathbf{A}\mathbf{x} - \mathbf{A}_{\mathrm{offdiag}}\hat{\mathbf{x}}_{\mathrm{pic}}^{(1)} + \mathbf{v} \\ &\approx \mathbf{A}_{\mathrm{diag}}\mathbf{x} + \mathbf{v}\end{aligned} \tag{3.40}$$

とする．ここで，$\mathbf{A}_{\mathrm{diag}}$ は $\mathbf{A}$ の主対角成分のみからなる対角行列であり，$\mathbf{A}_{\mathrm{offdiag}} = \mathbf{A} - \mathbf{A}_{\mathrm{diag}}$ である．次に，$\mathbf{y}_{\mathrm{pic}}^{(2)}$ から線形推定などで推定値 $\hat{\mathbf{x}}_{\mathrm{pic}}^{(2)}$ を求め，これを用いて上記の式と同様の方法で干渉成分を除去する．この処理を繰り返すことで，推定精度の向上を図るのが PIC の基本的な考え方である．PIC は SIC と異なり，$\mathbf{x}$ の成分間で SNR の違いがあまり無い場合に有効であることが知られている．

3.6　最大事後確率推定と最尤推定

$\mathbf{x}$ の各成分が有限のシンボル集合 $\mathcal{S}$ から生成されるとする．すなわち $\mathbf{x} \in \mathcal{S}^N$ である．観測ベクトル $\mathbf{y}$ が得られたとき，$\mathbf{x}$ の推定値として選んだ任意の $\hat{\mathbf{x}} \in \mathcal{S}^N$ が真の $\mathbf{x}$ と一致する確率は $P(\hat{\mathbf{x}}|\mathbf{y})$ であり，これは事後確率と呼ばれる．従って，判定が正しい確率を最大にするような，言い換えると，誤り確率が最小になるような判定ルールは，事後確率を最大にする $\hat{\mathbf{x}} \in \mathcal{S}^N$ を選択すること，すなわち

$$\hat{\mathbf{x}}_{\mathrm{map}} = \arg\max_{\mathbf{x}\in\mathcal{S}^N} P(\mathbf{x}|\mathbf{y}) \tag{3.41}$$

である．これは最大事後確率推定と呼ばれ，誤り確率を最小にするという意味で最適な検出法である．

一方，ベイズの定理を用いると[5]，事後確率は

$$P(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})P(\mathbf{x})}{p(\mathbf{y})} \tag{3.42}$$

と書き直せる．ここで，$\mathbf{x}$ の候補が同様に確からしく，$P(\mathbf{x})$ が一様分布であるとすると，事後確率を最大化する問題は

$$\hat{\mathbf{x}}_{\mathrm{ml}} = \arg\max_{\mathbf{x}\in\mathcal{S}^N} p(\mathbf{y}|\mathbf{x}) \tag{3.43}$$

に帰着される．これは尤度 $p(\mathbf{y}|\mathbf{x})$ を最大化する問題であることから，最尤推定と呼ばれる．

線形観測モデル (3.1) において，$\mathbf{v}$ の各成分が独立同一分布に従う複素ガウス雑音であるとすると，尤度関数は

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{\pi^M \det\{\mathbf{R}_{\mathrm{v}}\}} \exp\left(-\frac{||\mathbf{y} - \mathbf{A}\mathbf{x}||_2^2}{\sigma_{\mathrm{v}}^2}\right) \tag{3.44}$$

[5] これは連続確率変数と離散確率変数についてのベイズ則 [29] であることに注意する

となることから，最尤推定の問題は

$$\hat{\mathbf{x}}_{\mathrm{ml}} = \arg\min_{\mathbf{x}\in\mathcal{S}^N} ||\mathbf{y} - \mathbf{A}\mathbf{x}||_2^2 \tag{3.45}$$

と書き直せる．これは，加法性雑音がガウス雑音の場合には，最尤推定は受信信号からのユークリッド距離が最小となる $\mathbf{Ax}$ を見つけ出す問題に帰着されることを意味する．この場合，評価関数が簡単化されているが，$\mathcal{S}$ の要素数を $|\mathcal{S}|$ とすると評価すべき候補数は $|\mathcal{S}|^N$ であり，N が大きい場合にはやはり実現が困難である．このため，様々な計算量削減法が提案されているが，一般に最尤推定に基づく手法はその良好な特性と引き換えに要求演算量が非常に大きい．

3.7　最大比合成

同一の信号に対する複数の観測が得られたときに，これらの観測をどのように利用（合成）するとよいかという問題に出くわすことがある (ダイバーシティなど)．ここではまず，(3.1) の特別な場合として

$$\mathbf{y} = \mathbf{a}x + \mathbf{v} \tag{3.46}$$

を考える．ただし，$x \in \mathbb{C}$, $\mathbf{a} = [a_1 \ \cdots \ a_M]^{\mathrm{T}} \in \mathbb{C}^M$ である．素朴な方法としては，$y_1, \cdots, y_M$ のうち最も SNR の高いものだけを利用する方法 (選択合成) や，$y_1, \cdots, y_M$ を同じ重みで足し合わせて合成する方法 (等利得合成) があるが，よく用いられるのは合成後の SNR が最大になるように重み付けをして足し合わせる最大比合成（MRC: maximal ratio combining）である．

最大比合成の重みベクトルを $\mathbf{w}_{\mathrm{mrc}}$ とすると，合成後の信号は

$$\begin{aligned}\hat{x}_{\mathrm{mrc}} &= \mathbf{w}_{\mathrm{mrc}}^{\mathrm{H}}\mathbf{y} \\ &= \mathbf{w}_{\mathrm{mrc}}^{\mathrm{H}}\mathbf{a}x + \mathbf{w}_{\mathrm{mrc}}^{\mathrm{H}}\mathbf{v}\end{aligned} \tag{3.47}$$

で与えられる．2 行目右辺第 1 項が信号成分，第 2 項が雑音成分を表すため，合成後の SNR は

$$\begin{aligned}\gamma_{\mathrm{mrc}} &= \frac{E[|\mathbf{w}_{\mathrm{mrc}}^{\mathrm{H}}\mathbf{a}x|^2]}{E[|\mathbf{w}_{\mathrm{mrc}}^{\mathrm{H}}\mathbf{v}|^2]} \\ &= \frac{\sigma_{\mathrm{x}}^2\mathbf{w}_{\mathrm{mrc}}^{\mathrm{H}}\mathbf{a}\mathbf{a}^{\mathrm{H}}\mathbf{w}_{\mathrm{mrc}}}{\sigma_{\mathrm{v}}^2\mathbf{w}_{\mathrm{mrc}}^{\mathrm{H}}\mathbf{w}_{\mathrm{mrc}}}\end{aligned} \tag{3.48}$$

となる．ただし，$E[|x|^2] = \sigma_{\mathrm{x}}^2$ としている．

いま，(3.48) より合成後の SNR が行列 $\mathbf{a}\mathbf{a}^{\mathrm{H}}$ のレイリー商（の定数倍）として表現されていることから，これを最大にする $\mathbf{w}_{\mathrm{mrc}}$ は 2.2 節の性質 6 から $\mathbf{a}\mathbf{a}^{\mathrm{H}}$ の最大固有値に対応する固有ベクトルによって与えられる．さらに，$\mathbf{a}\mathbf{a}^{\mathrm{H}}$ はランク 1 の行列であることから

$$\mathbf{w}_{\mathrm{mrc}} = \mathbf{a} \tag{3.49}$$

とすれば良いことが直ちに分かる．このとき，

$$
\begin{aligned}
\gamma_{\mathrm{mrc}} &= \frac{\sigma_{\mathrm{x}}^2 \mathbf{a}^{\mathrm{H}} \mathbf{a} \mathbf{a}^{\mathrm{H}} \mathbf{a}}{\sigma_{\mathrm{v}}^2 \mathbf{a}^{\mathrm{H}} \mathbf{a}} \\
&= \frac{\sigma_{\mathrm{x}}^2 \mathbf{a}^{\mathrm{H}} \mathbf{a}}{\sigma_{\mathrm{v}}^2} \\
&= \frac{|a_1|^2 \sigma_{\mathrm{x}}^2}{\sigma_{\mathrm{v}}^2} + \frac{|a_2|^2 \sigma_{\mathrm{x}}^2}{\sigma_{\mathrm{v}}^2} + \cdots + \frac{|a_M|^2 \sigma_{\mathrm{x}}^2}{\sigma_{\mathrm{v}}^2}
\end{aligned} \tag{3.50}
$$

となることから，最大比合成後の SNR は個々の観測の SNR の和となることが確認できる．

これまで，送信信号がスカラーのモデル (3.46) について考えてきたが，(3.1) のモデルの場合にも同様の手法が適用できる．重みベクトルを $\mathbf{w}_{\mathrm{mrc}}$ とすると，このときの合成後の信号は

$$
\begin{aligned}
\hat{x}_{\mathrm{mrc}} &= \mathbf{w}_{\mathrm{mrc}}^{\mathrm{H}} \mathbf{y} \\
&= \mathbf{w}_{\mathrm{mrc}}^{\mathrm{H}} \mathbf{A} \mathbf{x} + \mathbf{w}_{\mathrm{mrc}}^{\mathrm{H}} \mathbf{v}
\end{aligned} \tag{3.51}
$$

で与えられる．ただし，$\hat{x}_{\mathrm{mrc}}$ はスカラーであり，未知ベクトル $\mathbf{x}$ の直接的な推定値ではないことに注意する．2 行目右辺第 1 項を信号成分，第 2 項を雑音成分とすると，合成信号の SNR は

$$
\begin{aligned}
\gamma_{\mathrm{mrc}} &= \frac{E[|\mathbf{w}_{\mathrm{mrc}}^{\mathrm{H}} \mathbf{A} \mathbf{x}|^2]}{E[|\mathbf{w}_{\mathrm{mrc}}^{\mathrm{H}} \mathbf{v}|^2]} \\
&= \frac{\mathbf{w}_{\mathrm{mrc}}^{\mathrm{H}} \mathbf{A} \mathbf{R}_{\mathrm{x}} \mathbf{A}^{\mathrm{H}} \mathbf{w}_{\mathrm{mrc}}}{\sigma_{\mathrm{v}}^2 \mathbf{w}_{\mathrm{mrc}}^{\mathrm{H}} \mathbf{w}_{\mathrm{mrc}}}
\end{aligned} \tag{3.52}
$$

となる．したがって，$\mathbf{w}_{\mathrm{mrc}}$ として，行列 $\mathbf{A}\mathbf{R}_{\mathrm{x}}\mathbf{A}^{\mathrm{H}}$ の最大固有値に対応する固有ベクトルを選べば良い．また，雑音成分が干渉信号で構成されている等，白色でない雑音を仮定すると，合成信号の SNR は

$$
\gamma_{\mathrm{mrc}} = \frac{\mathbf{w}_{\mathrm{mrc}}^{\mathrm{H}} \mathbf{A} \mathbf{R}_{\mathrm{x}} \mathbf{A}^{\mathrm{H}} \mathbf{w}_{\mathrm{mrc}}}{\mathbf{w}_{\mathrm{mrc}}^{\mathrm{H}} \mathbf{R}_{\mathrm{v}} \mathbf{w}_{\mathrm{mrc}}} \tag{3.53}
$$

と表現される．これより，$\mathbf{w}_{\mathrm{mrc}}$ として，行列 $\mathbf{A}\mathbf{R}_{\mathrm{x}}\mathbf{A}^{\mathrm{H}}$ と $\mathbf{R}_{\mathrm{v}}$ の一般固有値問題，すなわち

$$
\mathbf{A} \mathbf{R}_{\mathrm{x}} \mathbf{A}^{\mathrm{H}} \mathbf{w} = \lambda \mathbf{R}_{\mathrm{v}} \mathbf{w} \tag{3.54}
$$

における最大一般固有値に対応する一般固有ベクトルを選べばよいことが分かる．

3.8　部分空間法

ここでは $\mathbf{x}$ の推定ではなく，行列 $\mathbf{A}$ の成分あるいは $\mathbf{A}$ に含まれるパラメータを観測ベクトル $\mathbf{y}$(の相関行列) から推定する問題を考える．

$\mathbf{x}$ と $\mathbf{v}$ が無相関でかつ $\mathbf{v}$ が白色雑音のとき，$\mathbf{y}$ の相関行列は

$$\begin{aligned}\mathbf{R}_{\mathrm{y}} &= E[\mathbf{y}\mathbf{y}^{\mathrm{H}}] \\ &= \mathbf{A}\mathbf{R}_{\mathrm{x}}\mathbf{A}^{\mathrm{H}} + \sigma_{\mathrm{v}}^2\mathbf{I}\end{aligned} \tag{3.55}$$

で与えられる．$\mathbf{R}_{\mathrm{y}}$ の M 個の固有値をその大きさにより $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_M$ とし, $\mathbf{A}\mathbf{R}_{\mathrm{x}}\mathbf{A}^{\mathrm{H}}$ の M 個の固有値を同様に $\nu_1 \geq \nu_2 \geq \cdots \geq \nu_M$ とする．$\mathbf{R}_{\mathrm{y}}$ はエルミート行列なので M 個の互いに直交する固有ベクトルをもち，固有値 λ_m に対応する正規化固有ベクトルを $\mathbf{q}_m$ とすると

$$\begin{aligned}\lambda_m\mathbf{q}_m &= \mathbf{R}_{\mathrm{y}}\mathbf{q}_m \\ &= (\mathbf{A}\mathbf{R}_{\mathrm{x}}\mathbf{A}^{\mathrm{H}} + \sigma_{\mathrm{v}}^2\mathbf{I})\mathbf{q}_m \\ &= (\nu_m + \sigma_{\mathrm{v}}^2)\mathbf{q}_m\end{aligned}$$

となる．よって λ_m と ν_m の間には

$$\lambda_m = \nu_m + \sigma_{\mathrm{v}}^2 \ \ , \ \ m = 1, 2, \ldots, M \tag{3.56}$$

なる関係がある．

ここで $M > N$ とし，$\mathbf{A}$ がフル列ランク，$\mathbf{R}_{\mathrm{x}}$ がフルランクと仮定すると，$\mathbf{A}\mathbf{R}_{\mathrm{x}}\mathbf{A}^{\mathrm{H}}$ の固有値は小さい方から $M - N$ 個が全て 0 となる．よって，(3.56) はさらに

$$\lambda_m = \begin{cases} \nu_m + \sigma_{\mathrm{v}}^2, & m = 1, \ldots, N \\ \sigma_{\mathrm{v}}^2, & m = N+1, \ldots, M \end{cases} \tag{3.57}$$

となる．実際，rank $\mathbf{A}^{\mathrm{H}} = N$ より，$\mathbf{A}^{\mathrm{H}}$ のカーネル空間 $\mathcal{N}(\mathbf{A}^{\mathrm{H}})$ の次元は $M - N$ であり，$\mathbf{q} \in \mathcal{N}(\mathbf{A}^{\mathrm{H}})$ に対しては $\mathbf{R}_{\mathrm{y}}\mathbf{q} = \sigma_{\mathrm{v}}^2\mathbf{q}$ となる．すなわち，$\mathbf{q}$ は固有値 σ_{v}^2 に対応する固有ベクトルであり，このカーネル空間の次元は $M - N$ なので，(3.57) のように重複度は $M - N$ である．よって，$\mathbf{q}_{N+1}, \cdots, \mathbf{q}_M$ は $\mathcal{N}(\mathbf{A}^{\mathrm{H}})$ の基底であり，

$$\mathbf{q}_m^{\mathrm{H}}\mathbf{A} = \mathbf{0}, \ \ m = N+1, \ldots, M \tag{3.58}$$

となる．ここで，$\mathbf{Q}_{\mathrm{S}} = [\mathbf{q}_1, \cdots, \mathbf{q}_N]$，$\mathbf{Q}_{\mathrm{N}} = [\mathbf{q}_{N+1}, \cdots, \mathbf{q}_M]$ と定義すると，$\mathcal{R}(\mathbf{Q}_{\mathrm{S}})$ は信号部分空間，$\mathcal{R}(\mathbf{Q}_{\mathrm{N}})$ は雑音部分空間と呼ばれる [2] ($\mathcal{R}(\cdot)$ は行列の列空間)．これは，(3.58) より $\mathcal{R}(\mathbf{Q}_{\mathrm{N}}) = \mathcal{N}(\mathbf{A}^{\mathrm{H}})$ であり，また $\mathbf{q}_1, \cdots, \mathbf{q}_M$ は正規直交基底であることから $\mathcal{R}(\mathbf{Q}_{\mathrm{S}}) = \mathcal{R}(\mathbf{Q}_{\mathrm{N}})^{\perp}$ であるが ($^{\perp}$ は直交補空間を表す)，一般に $\mathcal{R}(\mathbf{A}) = \mathcal{N}(\mathbf{A}^{\mathrm{H}})^{\perp}$ であることから，$\mathcal{R}(\mathbf{Q}_{\mathrm{S}}) = \mathcal{R}(\mathbf{A})$ 及び $\mathcal{R}(\mathbf{Q}_{\mathrm{N}}) = \mathcal{R}(\mathbf{A})^{\perp}$ が成立するからである．

直交条件 (3.58) を用いることで $\mathbf{A}$ の成分や $\mathbf{A}$ に含まれるパラメータを推定することが出来る．このような手法は信号部分空間と雑音分空間の直交性を利用してることから一般に部分空間法と呼ばれる．また，雑音部分空間への直交射影を考えているとみなすことができるため，信号部分空間への直交射影を考える主成分分析に対して，マイナー成分分析とも呼ばれる．

3.9　圧縮センシング

圧縮センシングは，真の解 $\mathbf{x}$ がスパースであること，すなわち $\mathbf{x}$ の非零要素の個数が見かけの次元 N に比べてずっと少ないこと，があらかじめ分かっているときに，そのスパース性を利用して $M < N$ の連立方程式 $\mathbf{y} = \mathbf{A}\mathbf{x}$ から真の解を見つけ出すための理論的枠組みである [3, 4, 5].

$M < N$ の場合に $\mathbf{y} = \mathbf{A}\mathbf{x}$ から $\mathbf{x}$ を推定する方法としては，既に正則化による最小ノルム解について述べたが，ここで追加された真の $\mathbf{x}$ がスパースであるという事前知識を考慮すると，素朴でかつ自然な方法は，最適化問題

$$\hat{\mathbf{x}}_{\ell_0} = \arg\min_{\mathbf{x}} ||\mathbf{x}||_0 \quad s.t. \quad \mathbf{y} = \mathbf{A}\mathbf{x} \tag{3.59}$$

を解くことで $\mathbf{x}$ を推定する方法である．これは ℓ_0 再構成と呼ばれ，非常に強力であることが知られているが，ℓ_0-ノルムの離散性と非凸性から一般に NP 困難である．そこで，ℓ_0-ノルムを ℓ_1-ノルムに置き換えることで緩和した問題

$$\hat{\mathbf{x}}_{\ell_1} = \arg\min_{\mathbf{x}} ||\mathbf{x}||_1 \quad s.t. \quad \mathbf{y} = \mathbf{A}\mathbf{x} \tag{3.60}$$

を考えると，これは線形計画問題に帰着されるため現実的な計算量で解くことが出来る．さらに重要なことに，$\mathbf{A}$ に関するある条件の下では，$M < N$ であってもこの最適化問題を解くことで真のスパースな $\mathbf{x}$ が完全再構成されることが理論的に示されている．

雑音を考慮する場合には，$\epsilon > 0$ として等式制約を不等式制約に置き換えることで

$$\hat{\mathbf{x}}_{c\ell_1} = \arg\min_{\mathbf{x}} ||\mathbf{x}||_1 \quad s.t. \quad ||\mathbf{A}\mathbf{x} - \mathbf{y}||_2^2 \leq \epsilon \tag{3.61}$$

を考える．あるいは，ある $\mu > 0$ に対して (3.61) と同じ解を与える等価な制約無し最適化問題

$$\hat{\mathbf{x}}_{\ell_1-\ell_2} = \arg\min_{\mathbf{x}} \left(\mu||\mathbf{x}||_1 + \frac{1}{2}||\mathbf{A}\mathbf{x} - \mathbf{y}||_2^2 \right) \tag{3.62}$$

を考えることもある．(3.62) はコスト関数に ℓ_1-ノルムと ℓ_2-ノルムの項が含まれていることから $\ell_1 - \ell_2$ 再構成と呼ばれる．また，よく知られた Lasso (least absolute shrinkage and selection operator)[6]，すなわち

$$\hat{\mathbf{x}}_{\mathrm{lasso}} = \arg\min_{\mathbf{x}} ||\mathbf{A}\mathbf{x} - \mathbf{y}||_2^2 \quad s.t. \quad ||\mathbf{x}||_1 \leq t \tag{3.63}$$

も (3.61), (3.62) と等価な再構成法である．

一般に圧縮センシングでは閉形式の解が得られないため，これらの最適化問題を解くための様々アルゴリズムが提案されている．圧縮センシングの基本的な考え方や，アルゴリズムの具体例，通信の問題への応用例については [7] などを参照されたい．

第4章　具体例

ここでは，前章で説明した基本的な信号処理手法が応用されている通信の問題のいくつかの例について説明する．

4.1　等化

等化は周波数選択性の畳み込み通信路によって歪んだ受信信号を補償するための技術であり，その受信信号モデルは典型的に (3.1) で表現される．具体的には，$\mathbf{s}$ を送信信号ベクトル，$\mathbf{r}$ を受信信号ベクトル，$\mathbf{H}$ を通信路行列，$\mathbf{v}$ を付加雑音ベクトルとして

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{v} \tag{4.1}$$

と表される．通信路行列 $\mathbf{H}$ の構造やサイズは伝送方式によって異なるが，最近様々な通信規格で広く採用されている周波数領域等化を前提としたシステムにおいては，ガード区間であるサイクリックプレフィックスの付加と除去の操作を含めた通信路行列は次のような巡回行列と呼ばれる構造をもつ．

$$\mathbf{H} = \begin{bmatrix} h_0 & 0 & \dots & 0 & h_L & \dots & h_1 \\ \vdots & h_0 & \ddots & & \ddots & \ddots & \vdots \\ h_L & & \ddots & \ddots & & \ddots & h_L \\ 0 & \ddots & & \ddots & \ddots & & 0 \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & & \ddots & 0 \\ 0 & \dots & \dots & 0 & h_L & \dots & h_0 \end{bmatrix} \in \mathbb{C}^{M \times M}$$

ここで，$h_0, \cdots, h_L$ は通信路のインパルス応答である．任意の巡回行列は，離散フーリエ変換 (DFT: discrete Fourier transform) 行列

$$\mathbf{D} = \frac{1}{\sqrt{M}} \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & e^{-j\frac{2\pi \times 1 \times 1}{M}} & \dots & e^{-j\frac{2\pi \times 1 \times (M-1)}{M}} \\ \vdots & \vdots & & \vdots \\ 1 & e^{-j\frac{2\pi (M-1) \times 1}{M}} & \dots & e^{-j\frac{2\pi (M-1) \times (M-1)}{M}} \end{bmatrix}$$

によってユニタリ相似変換されるという非常に有用な性質をもっている．すなわち，$\mathbf{H}$ はその成分 $\{h_0, h_1, \dots, h_L\}$ の値によらず

$$\mathbf{H} = \mathbf{D}^{\mathrm{H}} \mathbf{\Lambda} \mathbf{D} \tag{4.2}$$

と書くことができる．ただし，$\boldsymbol{\Lambda} = \mathrm{diag}[\lambda_1 \ \cdots \ \lambda_M]$ であり，その対角成分は

$$\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_M \end{bmatrix} = \sqrt{M}\mathbf{D} \begin{bmatrix} h_0 \\ \vdots \\ h_L \\ \mathbf{0}_{(M-L-1)\times 1} \end{bmatrix} \tag{4.3}$$

で定義される[1]．これより，受信信号ベクトルは

$$\mathbf{r} = \mathbf{D}^{\mathrm{H}}\boldsymbol{\Lambda}\mathbf{D}\mathbf{s} + \mathbf{v} \tag{4.4}$$

と書くことができる．

線形等化を仮定し，等化器出力を

$$\hat{\mathbf{s}} = \mathbf{W}^{\mathrm{H}}\mathbf{r} \tag{4.5}$$

とすると，ZF 等化の場合には (3.4) より

$$\begin{aligned} \mathbf{W}^{\mathrm{H}} &= (\mathbf{D}^{\mathrm{H}}\boldsymbol{\Lambda}\mathbf{D})^{-1} \\ &= \mathbf{D}^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{D}^{-\mathrm{H}} \\ &= \mathbf{D}^{\mathrm{H}}\boldsymbol{\Lambda}^{-1}\mathbf{D} \end{aligned} \tag{4.6}$$

となる．また，MMSE 等化の場合には，(3.35) より

$$\begin{aligned} \mathbf{W}^{\mathrm{H}} &= \sigma_{\mathrm{s}}^2\mathbf{H}^{\mathrm{H}}\left(\sigma_{\mathrm{s}}^2\mathbf{H}\mathbf{H}^{\mathrm{H}} + \sigma_{\mathrm{v}}^2\mathbf{I}\right)^{-1} \\ &= \sigma_{\mathrm{s}}^2\mathbf{D}^{\mathrm{H}}\boldsymbol{\Lambda}^{\mathrm{H}}\mathbf{D}\left(\sigma_{\mathrm{s}}^2\mathbf{D}^{\mathrm{H}}\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\mathrm{H}}\mathbf{D} + \sigma_{\mathrm{v}}^2\mathbf{I}\right)^{-1} \\ &= \mathbf{D}^{\mathrm{H}}\boldsymbol{\Lambda}^{\mathrm{H}}\left(\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\mathrm{H}} + \frac{\sigma_{\mathrm{v}}^2}{\sigma_{\mathrm{s}}^2}\mathbf{I}\right)^{-1}\mathbf{D} \end{aligned}$$

となる．ただし，$E[\mathbf{s}\mathbf{s}^{\mathrm{H}}] = \sigma_{\mathrm{s}}^2\mathbf{I}$ とし，$\mathbf{D}$ がユニタリ行列であるという性質を用いている．等化器行列 $\mathbf{W}^{\mathrm{H}}$ は，ZF 等化と MMSE 等化の何れの場合も，対角行列を IDFT の操作を表す $\mathbf{D}^{\mathrm{H}}$ と DFT の操作を表す $\mathbf{D}$ で挟んだ形になっているため，高速フーリエ変換 (FFT: fast Fourier transform) を用いた周波数領域での 1 タップ等化によって効率的に等化を実現できる．

ここでは，周波数領域等化を前提として ZF 推定と MMSE 推定の例を示したが，最尤推定に基づく最尤系列推定 (MLSE: maximum likelihood sequence estimation) も一般に等化基準としてよく用いられている．

4.2　通信路推定

通信路推定は，等化器重みの決定や様々な信号処理に必要な通信路に関する情報，すなわち，通信路のインパルス応答や周波数応答を推定する問題である．元になるモデルは等化の

[1] $[\lambda_1, \ldots, \lambda_M]^{\mathrm{T}}$ は通信路のインパルス応答の離散フーリエ変換であり, 通信路の周波数応答に他ならない．

場合と同様に

$$\mathbf{r} = \mathbf{H}\mathbf{p} + \mathbf{v} \tag{4.7}$$

で表される．ただし，ここでは送信信号ベクトルとして既知のパイロット信号ベクトル $\mathbf{p} = [p_1 \ \cdots \ p_M]^{\mathrm{T}}$ を仮定している．さらに，等化の例と同様に通信路行列 $\mathbf{H}$ が巡回行列であるとすると，巡回行列の性質から (4.7) は

$$\mathbf{r} = \mathbf{P}\mathbf{h} + \mathbf{v} \tag{4.8}$$

と書き直される．ここで，$\mathbf{P}$ はパイロット信号で構成される巡回行列

$$\mathbf{P} = \begin{bmatrix} p_1 & p_M & \cdots & p_2 \\ p_2 & p_1 & & p_3 \\ \vdots & & \ddots & \vdots \\ p_M & p_{M-1} & \cdots & p_1 \end{bmatrix} \tag{4.9}$$

であり，$\mathbf{h}$ は $\mathbf{H}$ の 1 列目のベクトルである．

(4.8) から $\mathbf{h}$ を推定する問題は，$\mathbf{H}$ が巡回行列であるときに (4.1) から $\mathbf{s}$ を推定する問題と本質的に同じである．このため，基本的には前節と同様の推定法が同じように適用可能であるが，等化の場合と異なり，$E[\mathbf{h}\mathbf{h}^{\mathrm{H}}]$ は通常事前に分からないため，MMSE 推定よりも ZF 推定に相当する $\mathbf{P}$ の逆行列を用いた推定法を用いることが多い[2]．ただし，通信路推定の場合には，通常 ZF 推定ではなく最小 2 乗推定と呼ばれる．

ここでは，巡回構造をもつ通信路行列を仮定し，通信路行列とパイロット信号ベクトルの役割の入れ替えを行なうことで，線形観測モデル (4.8) を得たが，テプリッツ構造をもつ通常の畳み込み通信路行列の場合にも同様の入れ替えが可能である．また，インパルス応答 $\mathbf{h}$ の成分がスパースである場合には圧縮センシングの手法がそのまま適用でき，これによって必要なパイロット信号数を減らす検討も行なわれている．

4.3 MIMO 通信

MIMO 通信は複数アンテナを備えた送受信機間で行なう通信であり，その受信信号モデルとして，そのまま (3.1) が利用できる．送信アンテナ数を N，受信アンテナ数を M とすると，$\mathbf{s} \in \mathbb{C}^N$ を送信信号ベクトル，$\mathbf{r} \in \mathbb{C}^M$ を受信信号ベクトル，$\mathbf{H} \in \mathbb{C}^{M \times N}$ を通信路行列，$\mathbf{v} \in \mathbb{C}^M$ を付加雑音ベクトルとして

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{v} \tag{4.10}$$

となる．ここで，各アンテナ間の通信路は複素数のゲインで表されるフラットフェージング通信路を仮定している．畳み込み通信路の場合は，前述の周波数領域等化によるブロック伝送方式と組み合わせることで，(4.10) と同様のモデルが得られる．

[2]何らかの方法で $E[\mathbf{h}\mathbf{h}^{\mathrm{H}}]$ の情報を得ることで，MMSE 推定を利用して推定精度を向上させる方法も検討されている．

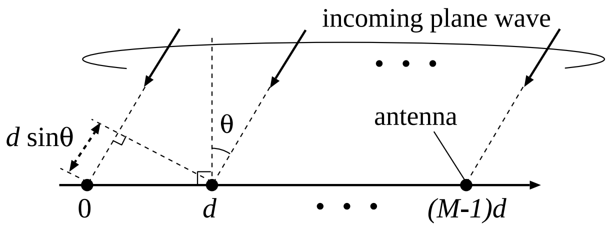

図 4.1: 等間隔直線アンテナアレー

MIMO 通信のモデル (4.10) は，等化や通信路推定の場合のそれと式の上では同一の形をしているが，MIMO 通信の場合には通信路行列 $\mathbf{H}$ に巡回性やテプリッツ性などの確定的な構造を仮定することがほとんどなく，多くの場合ランダム行列として扱われる．このため，MIMO 通信の信号検出法はその多くが一般的で他の問題にも応用しやすく，特に [12] などの MIMO 通信のテキストは通信のための信号処理手法を効率的に学ぶのに適している．MIMO 通信では，本稿で述べたほとんどの信号推定法がその検出法として検討されている．最近では送受信アンテナ数が数十〜数百といった大規模な MIMO 通信が検討されており，そこでは単にアンテナ数が多くなるだけでなく，まさに “more is different” というべき大規模特有の興味深い性質が見られる．大規模 MIMO については [30] が詳しい．

4.4　到来方向推定

複数のアンテナで観測される信号を用いて，信号（電波）の到来する方向を推定する問題を考える．簡単のため，到来する N 個の信号は同一の中心周波数をもつ狭帯域の平面波であるとし，M 個のアンテナ素子が直線上に等間隔 d で並んでいるとする（図 4.1 参照）．

n 番目の到来波がアンテナ素子の並びに垂直な方向から θ_n の角度で到来するとき，この狭帯域信号に対する隣り合うアンテナ素子での受信信号の相違は，行路差 $d\sin\theta_n$ に起因する位相差

$$\phi_n = 2\pi\frac{d\sin\theta_n}{\eta} \tag{4.11}$$

のみとなる．ここで，η は到来波の波長である．これより，1 番目のアンテナ素子での各到来波の受信信号成分を $\{s_1, s_2, \cdots, s_N\}$ とすると，m 番目のアンテナ素子での受信信号は

$$r_m = \sum_{n=1}^{N} s_n e^{\mathrm{j}\phi_n(m-1)} + v_m \tag{4.12}$$

となる．ここで，v_m は平均 0, 分散 σ_{v}^2 の白色雑音である．これをまとめると

$$\mathbf{r} = [r_1\ \cdots\ r_M]^{\mathrm{T}} = \sum_{n=1}^{N} s_n\mathbf{a}(\theta_n) + \mathbf{v} \tag{4.13}$$

と書ける．ただし，$\mathbf{v} = [v_1 \ \cdots \ v_M]^{\mathrm{T}}$ であり，

$$\mathbf{a}(\theta) = \left[1, e^{\mathrm{j}2\pi\frac{d\sin\theta}{\eta}}, \cdots, e^{\mathrm{j}2\pi\frac{(N-1)d\sin\theta}{\eta}}\right]^{\mathrm{T}} \tag{4.14}$$

である．さらに，$\mathbf{A} = [\mathbf{a}(\theta_1) \ \cdots \ \mathbf{a}(\theta_N)]$，$\mathbf{s} = [s_1 \ \cdots \ s_N]^{\mathrm{T}}$ と定義することで，(3.1) と同じ形のモデル

$$\mathbf{r} = \mathbf{A}\mathbf{s} + \mathbf{v} \tag{4.15}$$

を得る．ただし，ここでの目的は $\mathbf{s}$ の推定ではなく，到来方向の推定，言い換えると行列 $\mathbf{A}$ に含まれるパラメータの推定である．そこで，部分空間法の利用を考える．到来波数 N とアンテナ素子数 M が $M > N$ であるとすると，(3.58) から

$$\mathbf{q}_m^{\mathrm{H}}\mathbf{A} = \mathbf{0}, \quad m = N+1, \ldots, M \tag{4.16}$$

なる関係が成り立つ．ここで，$\mathbf{q}_m$ は $\mathbf{R} = E[\mathbf{r}\mathbf{r}^{\mathrm{H}}]$ の小さい $M-N$ 個の固有値に対応する固有ベクトルである．このとき，

$$S(\theta) = \frac{1}{\sum_{m=N+1}^{M} |\mathbf{a}^{\mathrm{H}}(\theta)\mathbf{q}_m|^2} \tag{4.17}$$

で θ を変化させると, 到来波に対応した到来角 $\theta = \theta_n$ $(n = 1, \cdots, N)$ で分母が 0 となり，そのグラフに鋭いピークができるため，これにより信号の到来角を推定できる[3]．この方法は MUSIC (multiple signal classification) 法 [13] と呼ばれる．到来方向推定は古典的な信号処理の問題の一つであるが，最近大きな進展があり再注目されている．最近の研究動向については [14] なども参照されたい．

---

[3] そのまま計算すると分母が 0 になり発散してしまうという恐れがあるが，実際には $\mathbf{R}$ は有限のサンプルを用いた時間平均によって近似的に計算されるサンプル相関行列であるため，実際の応用では問題にならない

第5章　文献紹介

前章までの本文中でも適宜文献の紹介を行ったが，最後に，本稿で直接説明しなかった内容に関するものも含めて，無線通信分野の初学者[1]が勉強する際に参考になると思われる文献をいくつか紹介する．なお，ここで紹介する文献は，筆者のごく限られた経験に基づくものであり，世の中に数多くある優れた文献のうちのごく一部であることをあらかじめご理解いただきたい．

5.1　数学関連

この分野に足を踏み入れようとする初学者にとって，数学を勉強する（勉強しなおす）ことは遠回りのように感じるかもしれないが，実はむしろその逆で，数学の基礎事項についての理解をおろそかにしたまま信号処理や無線通信の勉強をする方が圧倒的に遠回りになる．実際，筆者の友人の数学者は，本業の傍らの空き時間に 3 ヶ月程度勉強しただけで，情報理論をほぼマスターしてしまった．もちろんこれは極端な例であるが，基礎的な数学の素養を身につけておけば，あっという間にこの分野の第一線にたどり着くことができる．

工学系の研究者が学ぶべき数学は理学部数学科のそれとは明らかに異なっており，より少ない分量に必要事項が凝縮された文献が望ましいという読者も多いと思われる．そのような方にお勧めなのが

千葉逸人，工学部で学ぶ数学，プレアデス出版，2009.

である．改訂版で「微積分学のまとめとその応用」が追加されたことで，確率・統計以外の必要最小限の内容をこれ一冊でほぼ網羅しているといってよく，非常にコンパクトにまとめられている．さらに特筆すべきことは，工学部の読者を意識して，一般的な数学の教科書に書かれていないような直感的な説明も与えられている点である．ちなみに，本書は著者（千葉逸人氏，九州大学）が京都大学の学生時代に書かれたものであるが，実際に読まれた方がその事実を知れば大変驚かれると思う．

信号処理を学ぶ上で最も重要な数学の分野はやはり線形代数である．もちろん，微積分学も非常に重要であるが，実際の研究の中で触れる計算の頻度から考えると線形代数に軍配が上がる[2]．また，線形代数で学ぶ事柄はある意味で数学の言葉のようのものであり，線形代数を理解していないと数学の他の分野では何が書かれているかその意味すら分からない．さ

[1]工学部の標準的なカリキュラムを学んだ学部卒程度の学生を想定している．

[2]正直に言うと筆者は微積分学が苦手であるが，なんとかこの分野で食べていけていることからも線形代数の重要性が理解される．

らに，数学や物理，あるいは工学の他の分野の研究者と議論をする際の共通言語としても，最低限線形代数をしっかり身につけておくことが必要である．このような事情から，線形代数については工学向けの文献ではなく，理学部数学科の学生が学ぶような教科書，例えば，

　佐武一郎，線型代数学，裳華房，1974.

などをじっくりと何度もチャレンジして読んでみてほしい (最近，新装版が出て文字が読みやすくなった)．実際の通信のための信号処理で Jordan 標準形が登場することは稀であるが，この分野の研究者を志すのあれば，将来，読んでいて良かったと思うときが来ることを保証する．

　確率・統計に関する文献は，初等的な確率論に基づく内容がやや平易すぎるものと測度論に基づく難解なものに二極化している印象があるが，この分野の初学者には，

　渡辺澄夫，村田昇，確率と統計–情報学への架け橋–，コロナ社，2005.

が大変お勧めである．非常に密度が高く，情報系分野の研究者に必要とされる内容がコンパクトにまとめられていながら，工学部の学生が読み進めていけるように配慮されている点が素晴らしい．統計的信号処理のほとんどの場面で，この文献の内容を理解すれば十分であると思われるが，やはりきちんと確率論を学びたいという向きは

　高橋幸雄，基礎数理講座 2　確率論，朝倉書店，2008.

が橋渡しに最適である．特に，σ-集合体を導入する必要性について詳しく書かれた 9 章は，工学系の研究者にとって大変有益であると思われる．ただし，これを理解するには選択公理などの集合論の知識が多少必要になる．集合論は工学部の標準的なカリキュラムには含まれていないが，IEEE Trans. Information Theory や IEEE Trans. Signal Processing の論文を理解するにはその知識が必要になることも多いので，どこかの段階で集合や位相について勉強されることを勧める．

　松坂和夫，集合・位相入門，岩波書店，1968.

あたりが標準的な教科書であるが，まずは読み物として

　野口廣，トロポジー　基礎と方法，ちくま学芸文庫，筑摩書房，2007.

でイメージを掴んでから，数学 30 講シリーズ（朝倉書店）や集合・位相の標準的な教科書で勉強されるとよい．この本は文庫本ではあるが，数学の重要事項が大変分かりやすく説明されており，例えばこの本を読んでから微積分学を勉強すると収束や連続の概念などがかなりクリアに理解できると思われる．

　他に，工学部のカリキュラムに含まれないが高度な信号処理で利用されている数学として

関数解析がある．筆者は内容に基づいて文献を紹介できるほど関数解析について理解していないが，

山田功，工学のための関数解析，数理工学社，2009.

は信号処理分野の第一線の研究者が工学者向けに著した文献である．

5.2　信号処理関連

信号処理の定番の教科書として外せないのは，やはり

S. Haykin, Adaptive Filter Theory (5th edition), Pearson, 2013.

である．筆者の手元にも，学生時代に購入して今ではボロボロになった第 3 版がある．また，

B. F.-Boroujeny，Adaptive Filters, Theory and Applications, John Wiley & Sons, 1998.

も同様の構成でお勧めできるが，より適応フィルタに特化した内容になっている．和書では，

片山徹，新版応用カルマンフィルタ，朝倉書店，2000.

が素晴らしい．タイトルにはカルマンフィルタとあるが，ベイズ推定やウィナーフィルタについても大変詳しく書かれている．信号処理の教科書としてはやや高度な内容になっており初学者には少しハードルが高いかもしれないが，例えば，エルゴード性の説明においても平均値と相関関数のエルゴード定理が証明されていたり，ウィナーフィルタも因果律の制約付きのものが説明されているなど，信号処理についてより深い理解を得たい人にお勧めできる．最近では，統計的信号処理についての和文の入門書もあり，

関原謙介，統計的信号処理，共立出版，2011.
関原謙介，ベイズ信号処理，共立出版，2015.

などは，本稿が目指していたものにかなり近い内容であるように思われる．

確率推論については

C. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.

が定番であり，和訳本も出ている．「パターン認識」や「機械学習」といったタイトルがついているため，無線通信や信号処理とは関係がないと思う読者もいるかもしないが，その内容は全ての無線通信の研究者が知っておくべきものである．筆者の研究室でも，前述の適応フィルタの教科書とあわせて輪講のテキストとしてよく使用している．また，統計科学関連

の文献は和書が充実しており，統計科学のフロンティアシリーズ（岩波書店）や予測と発見の科学シリーズ（朝倉書店）などはいずれも大変面白い．

5.3　通信・情報理論関連

情報理論の定番は

T. M. Cover, J. A. Thomas, Elements of Information Theory, John Wiley & Sons, 2006.
R. G. Gallager, Information Theory and Reliable Communication, John Wiley & Sons, 1968.

などであるが，工学系の初学者に分かりやすいように書かれているのは Cover の本であると思われ，和訳本も出ている．情報理論や符号理論には和書，洋書を問わず他にも優れた文献が数多くあり，例えば，

D. J. C. Mackay, Information Theory, Inference, and Learning Algorithms, Cambridge University Press, 2003.
T. Richardson, Modern Coding Theory, Cambridge University Press, 2008.
韓太舜，小林欣吾，情報と符号化の数理，培風館，1999.
和田山正，誤り訂正技術の基礎，森北出版，2010.

などは是非一度目を通していただきたい文献である．ただし，情報理論は応用数学の色彩が強いので，何れの文献もハードルが高いと感じるかもしれない．そういった向きには，まず読み物として

甘利俊一，情報理論，ちくま学芸文庫，筑摩書房，2011

を読んでイメージをつかむことをお勧めする．特に，情報の量を定義するところの説明は秀逸であり，信号空間に関する説明は情報理論だけではなく信号処理の理解にも大いに役立つと思われる．

無線通信関連の文献は，

J. G. Proakis, M. Salehi, Digital Communications, McGraw-Hill, 2008.

が定番である．1000 ページ以上あり読み応え満点であるが，本稿で説明した基礎事項を理解した後に是非チャレンジしてほしい．情報理論寄りの立場から書かれた

D. Tse, P. Viswanath, Fundamentals of Wireless Communication, Cambridge University Press, 2005.

R. G. Gallager, Principles of Digital Communication, Cambridge University Press, 2008.

などは，もう少しページ数も少なく手に取りやすいかもしれない．通信理論，通信方式関係の和書は比較的古典的な内容のものが多い印象があり，最新の内容を網羅した和文の入門書の登場が期待されるが，前章でも述べたように，

菊間信良，アダプティブアンテナ技術，オーム社，2003.
大鐘武雄，小川恭孝，わかりやすい MIMO システム技術，オーム社，2009.

などのアレイ信号処理や MIMO 通信のテキストは通信のための信号処理手法を効率的に学ぶのに適している．

第6章　むすび

本稿では，無線通信で頻繁に使用される基本的かつ典型的な信号処理技術について，できる限り素朴な問題設定を用いて説明した．現在の無線通信システムでは非常に複雑な信号処理手法が用いられているが，ほとんどの場合，本稿で説明した何れかの手法あるいはその導出に用いたアプローチが基礎となっている．また，準備の章で説明したウィルティンガー微分は，実際の信号処理で頻繁に利用されているにもかかわらず通信や信号処理の入門書できちんと説明されていることは極めて稀であり，筆者も学生時代にここで躓いて大変困った記憶があるため，特に重点的に説明を行った．突然複素関数論の話が出てきて戸惑った読者もいるかもしれないが，ここは通信のための信号処理を理解する上で鍵と言ってもよいところなので，本当に分かったと思えるまで実際に手を動かして計算してみてほしい．

本稿の冒頭で述べた目的が達成できたかどうかは甚だ怪しいところではあるが，この分野に足を踏み入れようとする初学者にとって，本稿が無線通信のための信号処理技術を理解する一助になれば幸いである．

最後に，筆者にこのような機会を与えて頂いた，「5G の根底にある無線信号処理に関する基礎ワークショップ」運営委員長 田野　哲氏（岡山大学），運営委員　岡本英二氏（名古屋工業大学）をはじめ他の運営委員のメンバー，並びに，無線通信システム研究専門委員会の委員の皆様に心より御礼申し上げます．

関連図書

[1] 片山徹, 新版応用カルマンフィルタ, 朝倉書店, 2000.

[2] 酒井英昭, "主成分分析と独立成分分析," システム／制御／情報, vol. 43, no. 4, pp. 188-195, 1999.

[3] D.L. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory, vol.52, no.4, pp.1289-1306, April 2006.

[4] E.J. Candes and T. Tao, "Decoding by linear programming," IEEE Trans. Inf. Theory, vol.51, no.12, pp.4203-4215, Dec. 2005.

[5] E.J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, " IEEE Trans. Inf. Theory, vol.52, no.2, pp.489-509, Feb. 2006.

[6] R. Tibshirani, "Regression shrinkage and selection via the lasso, " J. R. Statist. Soc. B, vol.58, no.1, pp.267-288, 1996.

[7] K. Hayashi, M. Nagahara, T. Tanaka, "A User's Guide to Compressed Sensing for Communications Systems," IEICE Trans. Commun., Vol. E96-B, No. 03, pp.685-712, Mar. 2013

[8] L. Schwartz, シュヴァルツ解析学 6 複素関数, 東京図書株式会社, 1971.

[9] S. Haykin, Adaptive Filter Theory, 3rd Edition, Prentice Hall, 1996.

[10] B. F.-Boroujeny，Adaptive Filters, Theory and Applications, John Wiley & Sons, 1998.

[11] T. Kailath, A. Sayed, B. Hassibi, Linear Estimation, Prentice Hall, 2000.

[12] 大鐘武雄，小川恭孝，わかりやすい MIMO システム技術，オーム社，2009.

[13] R. O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," IEEE Trans. Antennas and Propag., Vol. AP-34, Vo. 3, pp. 276-280, 1986.

[14] 林 和則, "狭帯域信号の到来方向推定," 電子情報通信学会 基礎・境界ソサイエティ Fundamentals Review, Vol. 8, No. 3, pp. 143-150, Jan. 2015.

[15] P. J. Schreier and L. L. Scharf, Statistical Signal Processing of Complex-Valued Data, Cambridge University Press, 2010.

[16] A. Hjørungnes, Complex-Valued Matrix Derivatives, Cambridge University Press, 2011.

[17] 池田思朗, “通信路容量と確率測度の最適化，” IEICE Fundamentals Review, Vol.5, No.3, pp. 230-238, Jan. 2012.

[18] J. G. Smith, “The information capacity of amplitude- and variance-constrained scalar Gaussian channels,” Inf. Control, Vol.18, pp. 203-219, 1971.

[19] 山本直樹，複素関数論の基礎，裳華房，2015.

[20] 小島紀男，矢沢志雄作，本間光一，マトリクスとシステム，東海大学出版会，1990.

[21] 町田東一，駒崎友和，松浦武信，マトリクスの固有値と対角化，東海大学出版会，1990.

[22] W. A. Gardner, Cyclostationarity in Communications and Signal Processing, IEEE Press, 1994.

[23] L. Tong, G. Xu, and T. Kailath, “Blind Channel Identification Based on Second-Order Statistics : A Time Domain Approach,” IEEE Trans. Inform. Theory, vol.41, pp. 340-349, Mar. 1994.

[24] L. Tong, G. Xu, B. Hassibi, and T. Kailath, “Blind Channel Identification Based on Second-Order Statistics : A Frequency Domain Approach,” IEEE Trans. Inform. Theory, vol.40, pp. 340-349, Mar. 1994.

[25] G. Xu, T. Kailath, “Direction-of-arrival estimation via exploitation of cyclostationarity - A combination of temporal and spatial processing, ” IEEE Trans. Signal Processing, vol. 40, no. 7, pp. 1775-1786, July 1992.

[26] Thomas M. Cover and Joy A. Thomas, Elements of Information Theory, 2nd Edition, Wiley-Interscience, 2006.

[27] D. Tse, P. Viswanath, Fundamentals of Wireless Communication, Cambridge University Press, 2005.

[28] S. Moshan, “Multi-user detection for DS-CDMA communications, ” IEEE Communications Magazine,pp. 124-136, Oct. 1996.

[29] D. P. Bertsekas and J. N. Tsitsiklis, Introduction to Probability, Athena Scientific, 2008.

[30] A. Chockalingam, B. S. Rajan, Large MIMO Systems, Cambridge University Press, 2014.