

修士論文

時系列解析を用いた大気揺らぎ予測補正による補償光学
の性能向上

Performance improvement of adaptive optics systems with
atmospheric turbulence predictive correction based on
time series analysis

東北大学大学院 理学研究科天文学専攻修士2年
B9SM3003 飯塚悠太

令和2年

要旨

地球の大気ゆらぎの影響をリアルタイムで測定し補正することで、地上望遠鏡の回折限界分解能を達成しようというのが補償光学のコンセプトである。具体的には、大気揺らぎの変動が無視できる程度の短時間に、可変形鏡を変形させて、乱れた波面を平面波に戻すという動作原理である。この概念は 1953 年に H.W.Babcock によって発表されたものであり、今現在でもこの原理が採用されている。

上記のコンセプトに対し本研究では、大気揺らぎの将来を予測し、波面補正を行うことが目的である。手法として、東北大学 50cm 望遠鏡で大気ゆらぎを測定し、そのデータを用いて時系列解析 (ARIMA モデルフィッティング) を行った。次にそのデータを用いて、Tip-Tilt 成分の補正シミュレーションを行い、その性能を従来の方法と比較し、予測補正の可能性について評価、議論する。

目次

第 1 章	導入	1
1.1	補償光学	1
1.2	補償光学の原理	1
1.2.1	波面センサー (Wavefront Sensor ; WFS)	1
1.2.2	可変形鏡 (Deformable Mirror ; DM)	3
1.2.3	制御計算機	4
1.3	ガイド星 (Guide Star; GS)	4
1.4	コーン効果 (the cone effect/focal anisoplanatism)	4
1.5	補償光学の性能の評価	5
1.5.1	fitting error	6
1.5.2	reconstruction error	6
1.5.3	temporal error	6
1.6	補償光学の多様化	7
1.6.1	接地境界層補償光学 (Ground-Layer Adaptive Optics; GLAO)	7
1.6.2	多層共役補償光学 (Multi-Congugate Adaptive Optics; MCAO)	7
1.6.3	多天体補償光学 (Multi Object Adaptive Optics; MOAO)	7
1.6.4	極限補償光学 (Extreme Adaptive Optics; ExAO)	8
1.6.5	レーザートモグラフィー補償光学 (Laser Tomography Adaptive Optics; LTAO)	8
1.7	補償光学を用いたサイエンス	11
1.7.1	球状星団 M13 の観測	11
1.7.2	AGN(活動銀河核) によるフィードバック現象の観測	12
1.7.3	系外惑星の直接撮像	12
第 2 章	大気揺らぎの物理	14
2.1	Kolmogorov の乱流理論	14
2.2	大気ゆらぎを評価する様々なパラメーター	14
2.2.1	屈折率構造関数と $C_N^2(z)$	14

2.2.2	平坦波面長 r_0 (Fried length)	15
2.2.3	等波面離角 θ_0 (isoplanatic angle)	15
2.2.4	コヒーレンス時間 τ_0 (coherence time)	15
2.3	可視光補償光学の難しさ	15
2.4	frozen-flow 仮定	16
2.5	Tip-Tilt 成分	18
第 3 章	時系列解析	21
3.1	時系列データ	21
3.2	時系列モデル	22
3.2.1	AR モデル	22
3.2.2	MA モデル	24
3.2.3	ARMA モデル	26
3.2.4	ARIMA モデル	27
3.3	統計的モデリング	28
3.3.1	最尤法	28
3.3.2	赤池情報量基準	28
3.4	Box-Jenkins 法	29
3.5	Python による時系列データ解析	29
第 4 章	制御工学	30
4.1	制御系の分類	30
4.1.1	フィードフォワード制御 (feedforward control)	30
4.1.2	フィードバック制御 (feedback control)	30
4.2	Laplace 変換	31
4.3	過渡応答	31
4.4	伝達関数	32
4.5	セッティング	34
第 5 章	データセット	35
5.1	東北大学 50cm 望遠鏡を用いた観測	35
5.2	MATLAB	36
第 6 章	結果	38
6.1	コレログラム	38
6.2	予測	45
6.3	過渡応答	47
6.4	MATLAB を用いたシミュレーション	48

6.5	周波数解析	52
第 7 章	議論	53
7.1	評価	53
7.1.1	分散	53
7.1.2	パラメーターの推定	54
7.1.3	VAR モデル	54
7.1.4	グレンジャー因果	55
7.2	機械学習	57
第 8 章	結論	58
付録 A	付録	60
A.1	コレログラム	60
A.2	統計的モデリング	61
A.2.1	確率分布と統計的モデル	61
A.2.2	KL 情報量とエントロピー最大化原理	61
A.2.3	KL 情報量の推定と対数尤度	62
A.2.4	最尤法によるパラメータ推定	63
A.2.5	状態空間モデル	63
A.2.6	カルマンフィルタ	64
A.2.7	長期予測	65
A.2.8	尤度の表現とパラメータ推定	66
参考文献		67

第 1 章

導入

1.1 補償光学

望遠鏡が発明されて以来、天文学者たちが得られる星像は大気揺らぎの影響を受けてぼやけてしまっていた。この原因は、それぞれの場所の空気の塊は異なった温度と密度を持っており、従って異なる屈折率を持つことである。地上からの観測において、視線方向に存在する空気の塊は、視線方向に沿った光路長の空間的および時間的変動をもたらす。過去、天文学者たちは山頂に望遠鏡を建設することや宇宙空間に望遠鏡を打ち上げることによって、大気の影響を緩和しようとしてきた。

しかし 1953 年、H.W.Babcock によって画期的なアイデアが提唱された [1]。それは、大気揺らぎの情報をリアルタイムで測定し、形状を局所的に変化させることのできる鏡に測定情報をフィードバックすることで回折限界に近い星像を獲得できるのではないか、というものであった。この技術が現在、補償光学 (Adaptive Optics ; AO) と呼ばれているものである。Babcock の提案から 20 年以上経ち、実用化するための技術が成熟した 1970 年代、アメリカ軍が人工衛星監視システムとして AO への投資を開始し、1982 年にマウイ島のハレアケアある 1.6m 望遠鏡に最初の AO システム (当時は Compensated Imaging System と呼ばれていた) を導入した。その後 1980 年代後半から天文学用として独立に開発が進められている。

補償光学の概念図を図1.1に示す。

1.2 補償光学の原理

上記のコンセプトを踏まえ、望遠鏡の回折限界像を達成するため、補償光学システムには主に以下の三つの構成要素が重要である。

1.2.1 波面センサー (Wavefront Sensor ; WFS)

実際の観測において、望遠鏡に入射した光の波面は、小開口 (Subaperture) に分割され、各場所の波面情報が波面センサーによって観測される。その測定の更新頻度は数 ms であり、センサーには非常に高感度なものが要求される。代表的な波面センサーにジャックハルトマン波面センサー

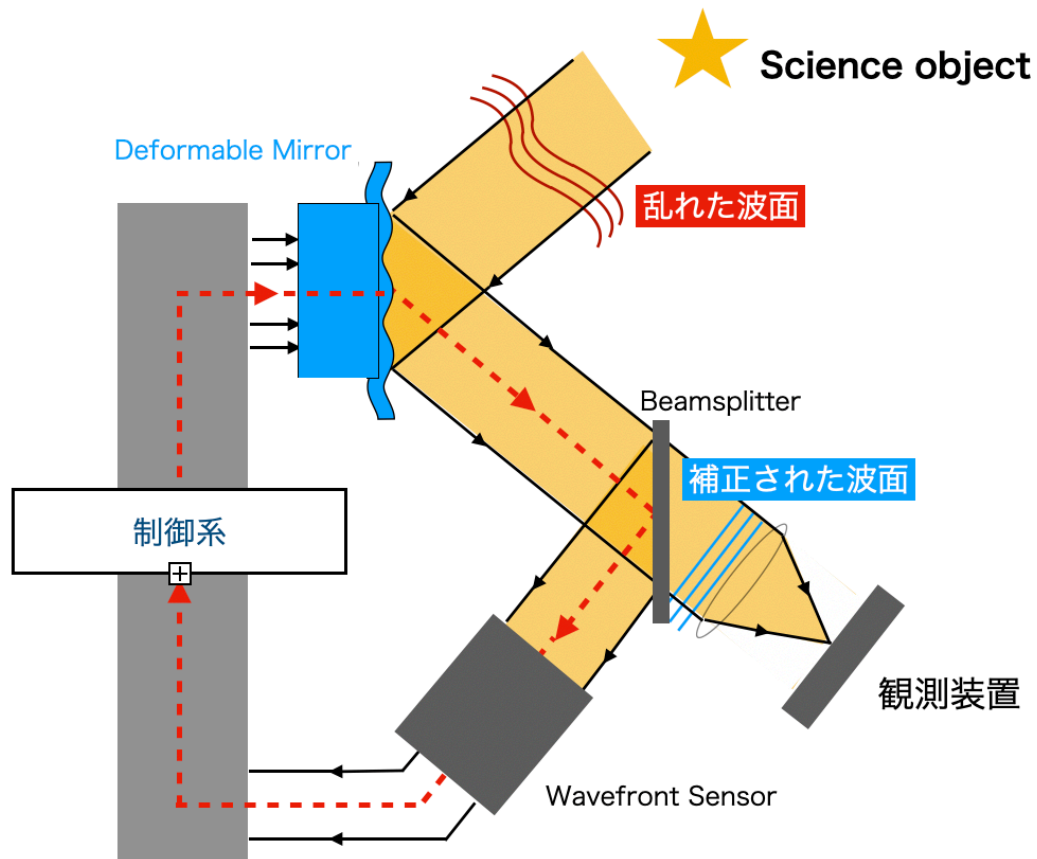


図1.1 補償光学の概念図

(Shack-Hartmann wavefront sensor; SHWFS) や曲率センサーなどがあり、それぞれ波面の局所的な傾き、曲率を測定する。

ここでは代表的なものとして、SHWFSを紹介する。SHWFSは最も一般的に使用されている波面測定方式の一つである。図1.2、図1.3にその概念図を示す。これは、小さなレンズが格子状に並んだ構造物（マイクロレンズアレイ）と検出器から構成されており、乱れた波面が入射すると、各マイクロレンズアレイによって複数の像が検出器上に結像される。平面波が入射したときのそれぞれの星像の基準位置と、乱れた波面の星像の位置の差を測定することで、波面全体の形状を推定することができる。

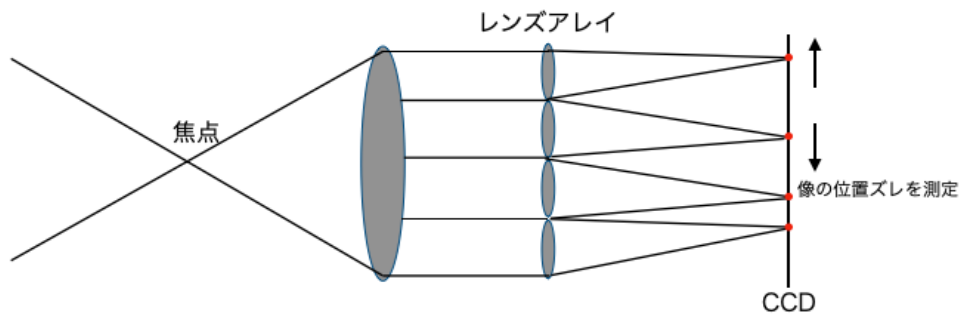


図1.2 シャックハルトマン波面センサー (横から)

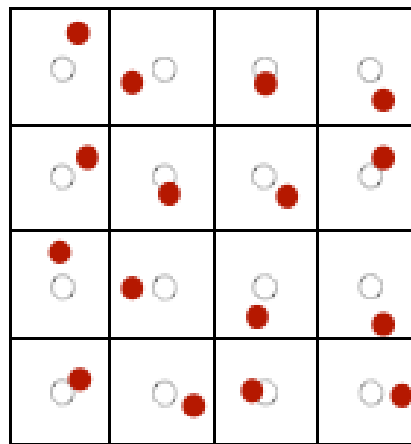


図1.3 シャックハルトマン波面センサー (正面から)

1.2.2 可変形鏡 (Deformable Mirror ; DM)

可変形鏡は、光の波面の位相ずれを補正する機能を担っており、反射、透過の両タイプがあるが、天文学で使われているのは反射型である。このタイプには、圧電素子や静電力を用いて鏡を変形させる方式があり、現在主に利用されているのはフェースプレート鏡とバイモルフ鏡である (1.4)。フェースプレート鏡は、連続した薄い鏡に複数のアクチュエーターを取り付けて、それぞれの素子を独立に動作させることによって、鏡を変形させる。

バイモルフ鏡は、垂直方向の電圧に対して、横方向に伸縮特性が正反対の薄い圧電素子板を2枚貼りあわせて作成する。

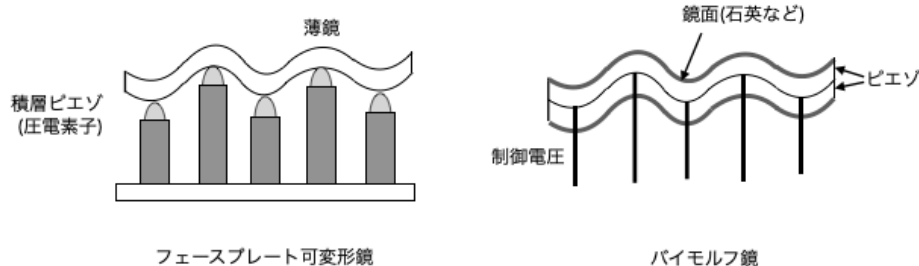


図1.4 可変形鏡

1.2.3 制御計算機

この部分の役割として、波面センサーで取得した波面情報をもとに、大気揺らぎによって乱された波面を平面波に戻すように可変形鏡を変形させ、安定的に駆動制御する。波面センサーの測定速度に応じて、毎秒数百 Hz～数 kHz 程度の波面演算を行う。

図1.5は波面補正のプロセスの概念図である。上記の補正器を用いて、図のような手順で波面補正が行われる。

1.3 ガイド星 (Guide Star; GS)

補償光学を実行するにあたって、波面測定参照光源として利用可能な、明るい星が必要である。その参照源として、観測天体の近くにある自然の星 (Natural Guide Star; NGS) を用いる方法とレーザーを打ち上げて人工的に生成したガイド星 (Laser Guide Star; LGS) を用いる方法を組み合わせられて行われている。

現在稼働中の近赤外線補償光学システムでは、可視光で 12～14mag までガイド星として使用することができ、それよりも暗いと光量不足により補償光学システムが十分に駆動しない。そこで、任意の方向にガイド星を作ることによって観測可能視野を大幅に拡大しようという目的のもと、LGSの技術が発明された。具体的な手法としては、ナトリウム D 線 (589nm) に波長を合わせたレーザーを照射することで、高度 90 km、厚さ 10 km のナトリウム層を励起発光させることで、人工のガイド星を生成する。

1.4 コーン効果 (the cone effect/focal anisoplanatism)

LGS からの光は、有限の距離にある点からの光であり、無限遠に位置する天体からの光の円柱とは正確に重ならないような円錐を形成している。そのため、LGS からの光線は NGS からの光線と同じ大気の乱流構造を横切らない。つまり、高い高度に位置する大気ゆらぎの影響が NGS

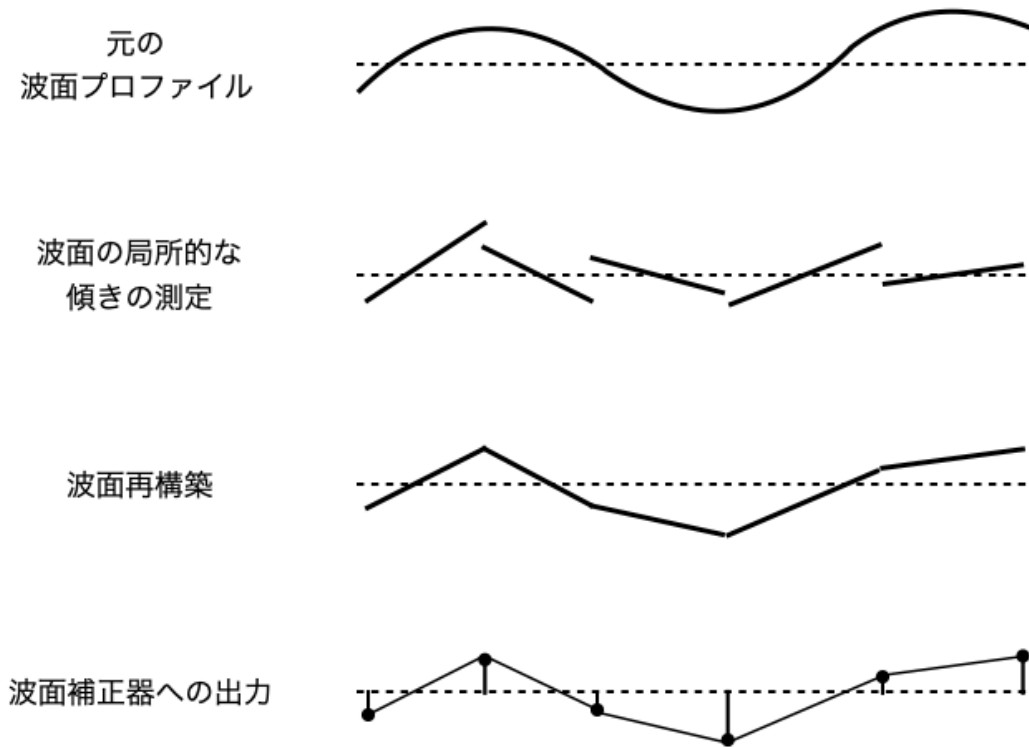


図1.5 波面測定のプロセス

を用いた場合よりも小さくなる。この現象をコーン効果といい、コーン効果による影響を小さくするため、複数の LGS を用いた補償光学 (後述) が提案されている。

1.5 補償光学の性能の評価

補償光学システムの性能を定量的に評価する一般的な指標としては、望遠鏡の開口部上の波面残差の分散 σ_{WFS}^2 、もう一つ一般的に用いられる指標はストレル比 SR (Strehl Ratio) である。これら 2 つの値の関係は、Marechal 近似を用いて次のように表される。[2]

$$S \sim e^{-\sigma_{WFS}^2} \quad (1.1)$$

波面残差の分散は、それらが無相関であると仮定して、個々の成分に分割することができる。

$$\sigma_{WFS}^2 = \sigma_{fit}^2 + \sigma_{rec}^2 + \sigma_{temp}^2 + other \quad (1.2)$$

式を見てわかる通り、生じてしまう主な誤差は次の三つである。

- fitting error; σ_{fit}^2
- reconstruction error; σ_{rec}^2
- temporal error; σ_{temp}^2

1.5.1 fitting error

波面の測定は、望遠鏡の開口を分割してそれぞれの波面の傾きを測定しているが、それより小さいスケールは測定できないので残差として残ってしまう。これを fitting error という。(図1.5)

乱流は mm のオーダーの空間スケールまで存在する。望遠鏡の開口が数メートルの場合、完全な補正を行うためには、入力収差をより細かく測定する必要があることを意味し、ガイド星からの光をより多くの部分に分割することになり、ガイド星の明るさの要求が高くなる。

1.5.2 reconstruction error

波面再構成の精度を低下させるすべての効果（測定ノイズ、キャリブレーションノイズ、サンプリング誤差、エイリアシング、色分散など）を組み合わせた誤差を reconstruction error という。

1.5.3 temporal error

実際の補償光学系では波面測定してから可変形鏡を動かすので、測定した時刻の波面と残差を測定する時刻が異なっていることによるエラーである。さらに、波面センサーの露光時間、WFSの読み出しやコマンド計算などによる遅延が加算される。これらの影響により、補正が可変形鏡に適用される時刻には、大気揺らぎの摂動が変化してしまい、不完全な補正になってしまう。これを temporal error という。

大気の特徴的な進化時間を τ_0 とすると（後述）、この時間内で補正ループを回す必要がある。

図1.6は、様々な種類の補償光学を用いて得られた残差を分解したものである [13]。

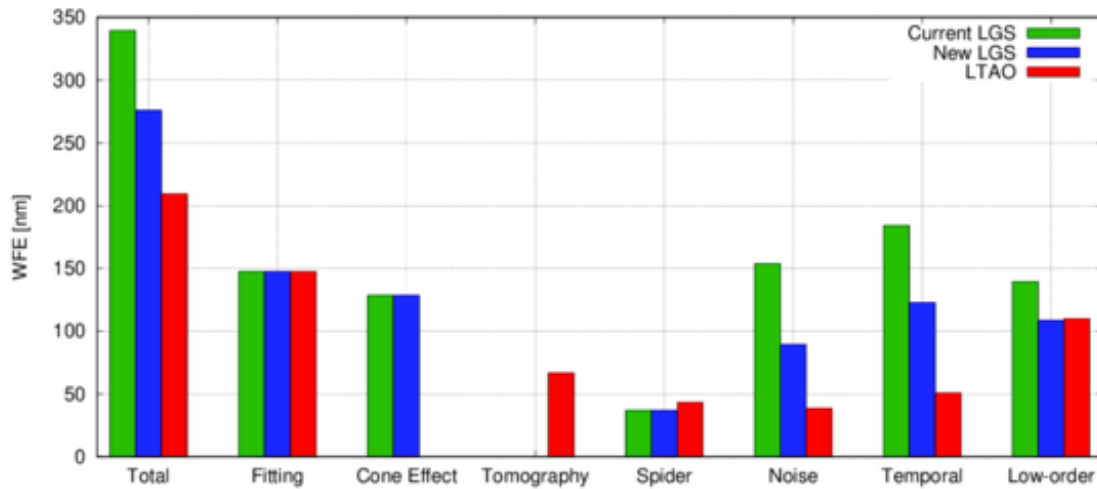


図1.6 すばる望遠鏡の補償光学系を仮定したシミュレーションで評価されたそれぞれの誤差成分の大きさ

1.6 補償光学の多様化

最近では、一つのガイド星による測定を使った補償光学システムは古典的補償光学と呼ばれている。現在、サイエンス天体によってさまざまな種類の補償光学が使い分けられており、多様化している。ここでは様々なタイプの補償光学について端的に説明する。それぞれの模式図は図1.7、図1.8、図1.9を参考にされたい。

1.6.1 接地境界層補償光学 (Ground-Layer Adaptive Optics; GLAO)

一般に大気ゆらぎの大きさは、高度が低くなるほど大きくなることわかっている(後述)。そこで、最も像に影響を与えてしまう地表層付近のゆらぎのみを補正して、ある程度の像質を担保しようとするものである。

1.6.2 多層共役補償光学 (Multi-Conjugate Adaptive Optics; MCAO)

複数のガイド星 (NGS, LGS) を使用して、断層写真法の手法で大気揺らぎの高さ方向の情報を取得し、複数の大気揺らぎの高さに光学的に共役 (Conjugate) な位置に可変形鏡を配置し、それぞれが対応した高さにある大気ゆらぎを補正する。

1.6.3 多天体補償光学 (Multi Object Adaptive Optics; MOAO)

複数の補償光学システムを各天体ごとに用意するものである。全体として MCAO より広い視野をカバーするのを目的としている。

1.6.4 極限補償光学 (Extreme Adaptive Optics; ExAO)

素子数が1000以上の超多素子補償光学で、視野が狭く明るいNGSが必要であるが、ストレル比で0.8以上の高精度な補正が可能となる。

1.6.5 レーザートモグラフィ補償光学 (Laser Tomography Adaptive Optics; LTAO)

高輝度なLGSを用いて、トモグラフィ補償の原理を用いて近赤外での定常的な回折限界の達成しようとするものである。(図1.7)

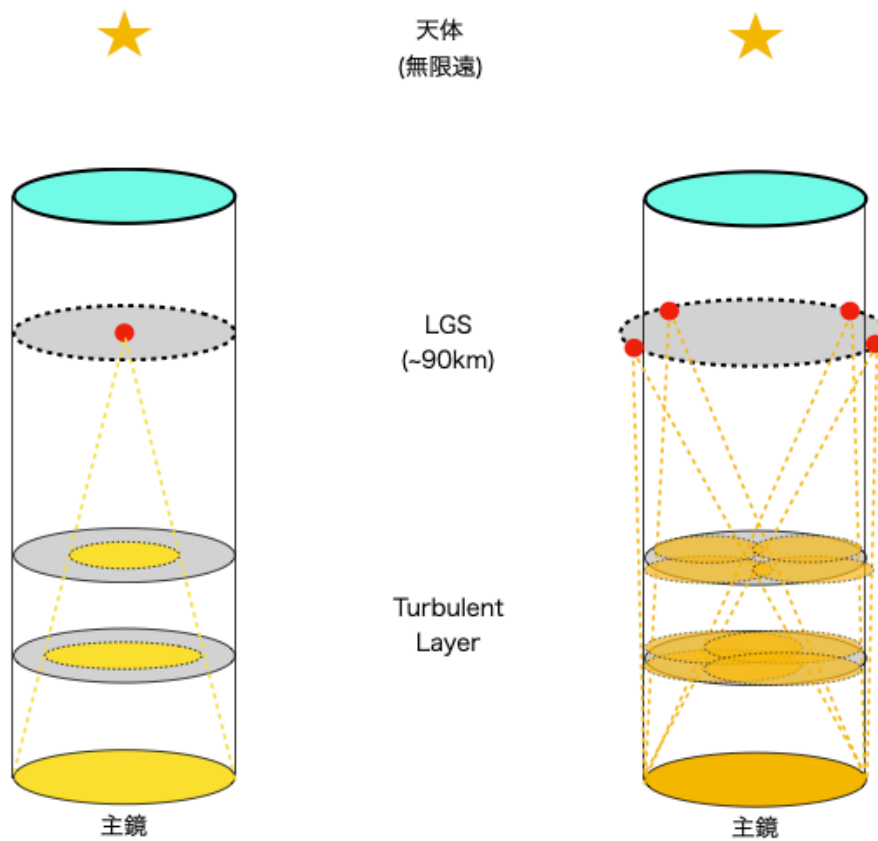


図1.7 LTAO

図1.9は、各種類の AO における SR-FoV(Field of View; 視野)をまとめたものである [3]。SR と FoV はトレードオフの関係にあり、サイエンス天体によって用いるべき AO システムが変わってくる。

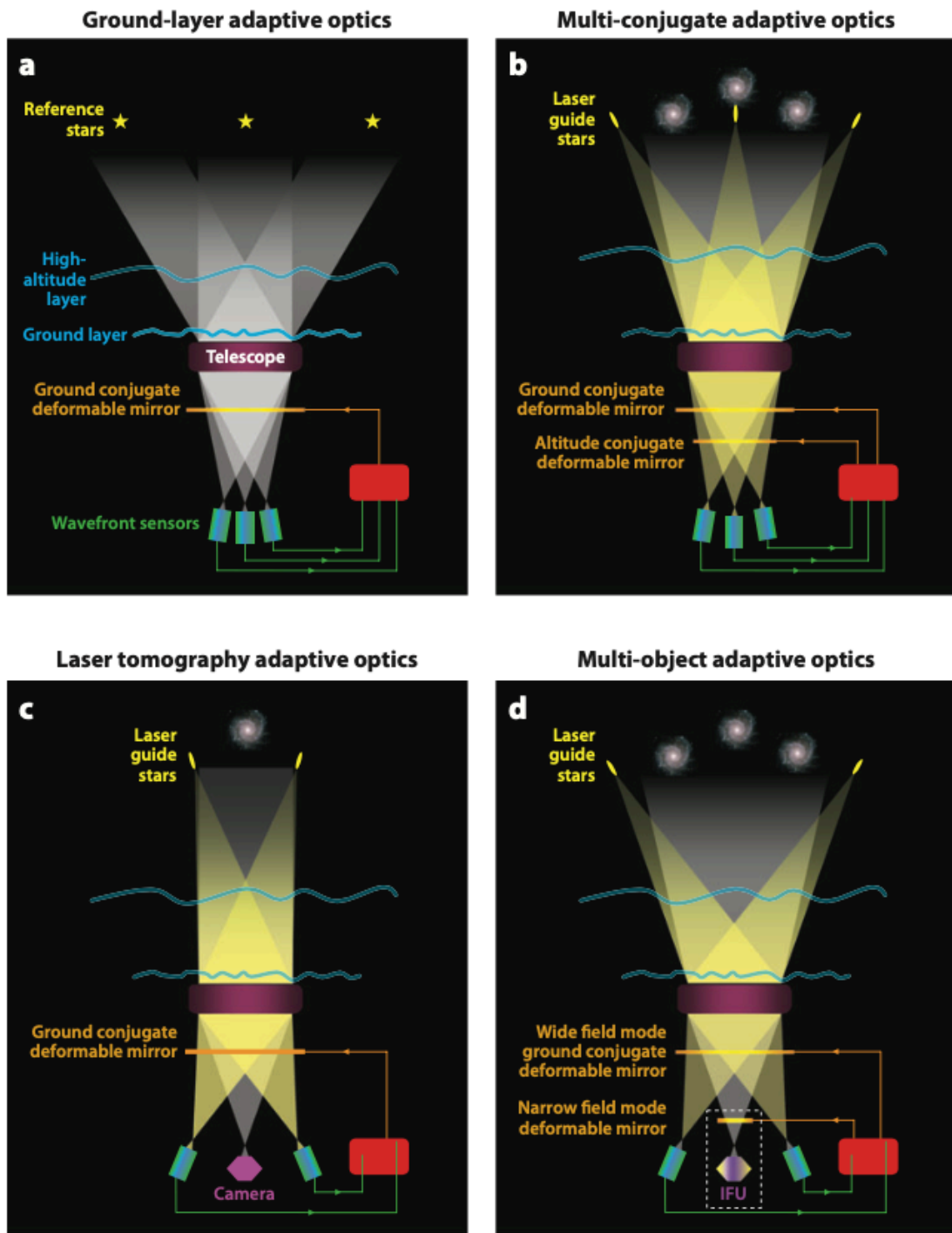


図1.8 補償光学の種類 [18]

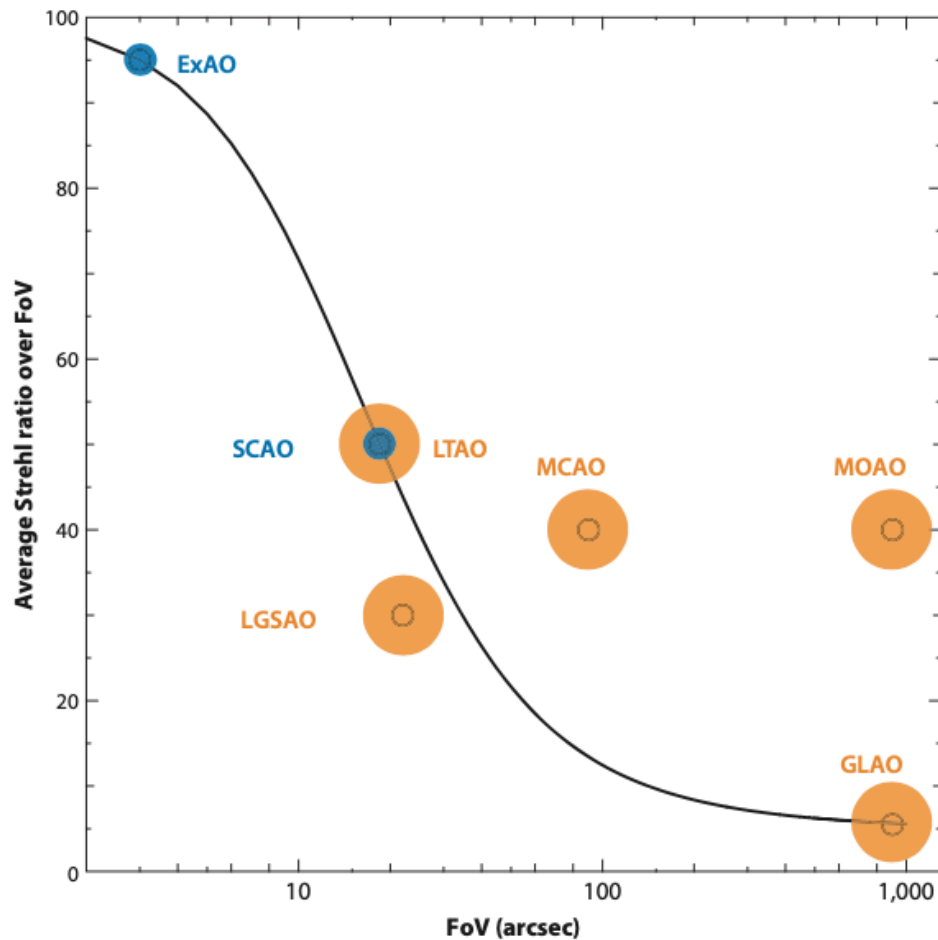


図1.9 補償光学の種類 2[3]

1.7 補償光学を用いたサイエンス

1.7.1 球状星団 M13 の観測

補償光学を用いた観測の一例を図1に示す。これはジェミニ南望遠鏡で観測球状星団 M13 の観測したもので、AO を使用しない場合と使用する場合がそれぞれ画像 a(左)、b(右) になっている。画像 a は、V バンドの acquisition camera で撮影されたもので、FWHM(Full Width at Half Maximum; 半値全幅) は 0.85 秒である。画像 b の AO 補正画像は K' バンド (2.12 μm) で、FWHM は 92 mas、視野はそれぞれ 22 秒である。比較画像から AO システムにより、角度分解能の向上がもたらされることは明らかである。

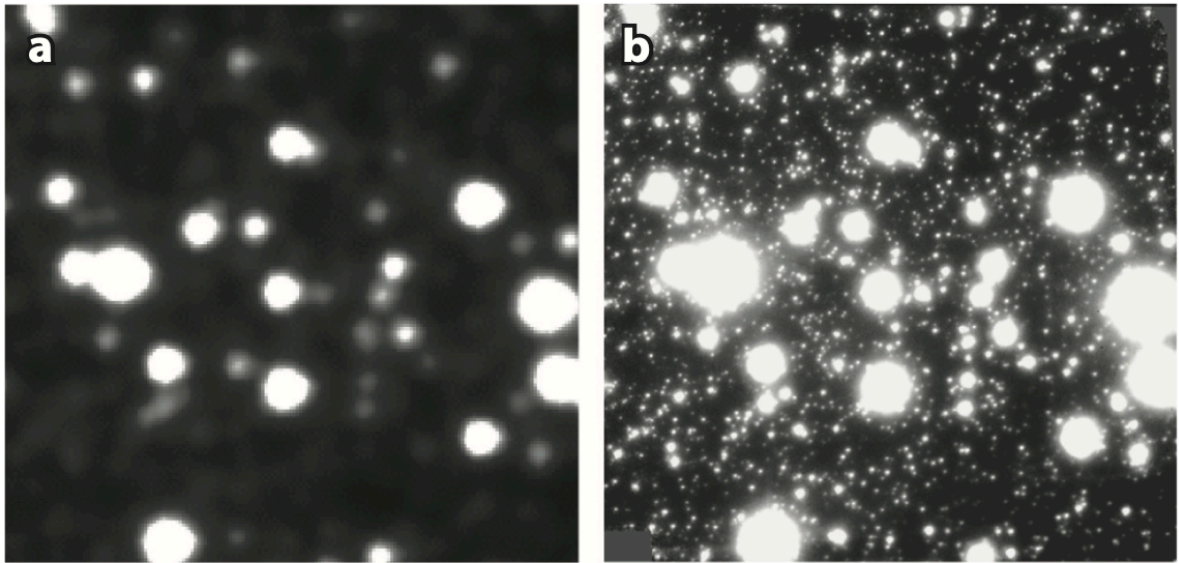


図1.10 補償光学を用いたサイエンスケース

1.7.2 AGN(活動銀河核) によるフィードバック現象の観測

AGN フィードバックとは、銀河中心付近での AGN から吹き出すジェットによって、銀河の周囲のガスを吹き飛ばすことで、ガスの降着や星の形成を妨げるメカニズムである。AGN フィードバックが起こっていることの観測的証拠はまだ乏しかった。

Kawaguchi et al.2018[8] では、近接した ($z = 0.039$) [O III] 輝線が顕著な AGN を観測し、すばる望遠鏡 AO188 を用いた光学 IFU(京都 3D II with AO188) を用いて、AGN のホスト銀河へのフィードバック過程が調べられた。銀河進化に対する AGN フィードバックの観測的検証を行い、そこで AGN アウトフローの空間構造や力学的エネルギーの推定が行われた。

1.7.3 系外惑星の直接撮像

補償光学の中でも、視野を犠牲にする代わりに高いストレル比を達成しようというのが前述の極限補償光学 (ExAO) である。これの使い道として代表的なサイエンスが系外惑星の直接観測である。惑星を直接観測することができれば、その化学的情報、例えば惑星の温度や色、大気の組成、表層環境などの情報が得られる。そして惑星の化学的情報は、惑星の形成や進化に関する知見をもたらす。究極的には生命の存在可能な環境や生命の存在有無を調査することが可能となる。

2008 年に行われた Keck 望遠鏡と Gemini 望遠鏡による高コントラスト観測 [12] で、初めて直接観測が成功し HR 8799 星の周りを回る 3 つの惑星が、24、38、68 天文単位の距離にあることが明

らかになった。

第 2 章

大気揺らぎの物理

2.1 Kolmogorov の乱流理論

大気乱流とその構造についての理論について説明する (Roddier 1981[4])。地球大気を流体とみなすと、流体の流れは、液体のレイノルズ数 $Re = V_0 L_0 / \mu_0$ (ただし、 V_0, L_0 は流れの特徴的な速度とスケールで、 μ_0 は流体の動粘性係数である) で特徴付けられる。地球大気の場合は $\mu_0 = 15 \times 10^{-6} m^2 s^{-1}, V_0 = 1 m/s, L_0 = 15 m$ を適用すると、 $Re = 10^6$ となり、これは乱流とみなされる。

この乱流によって生じる大気屈折率の非一様性が波面の位相を乱し、集光したときに像の劣化を引き起こす。この非一様性のことを”大気ゆらぎ”と呼ぶ。

2.2 大気ゆらぎを評価する様々なパラメーター

ここでは、大気ゆらぎによって乱された波面の性質を表す重要な指標を紹介する。

2.2.1 屈折率構造関数と $C_N^2(z)$

大気ゆらぎの評価には構造関数が用いられる。詳細な説明は Hardy 1998[5]、付録を参照されたい。構造関数とは、ある物理量や物理量同士の相関の度合いを表す関数である。相関関数と異なるのは、値が小さければ相関がより強いという点である。

大気屈折率の構造関数は Obukhov 1949[6] によって解析的に求められており、次のように定義される。

$$D_N = \langle |N(x) - N(x+r)|^2 \rangle = C_N^2 r^{2/3} \quad (2.1)$$

この式から、 C_N^2 の値が大きいほど構造関数が急激に大きくなる、すなわち相関長が小さくなることがわかる。 C_N^2 はゆらぎ強度の指標となるパラメータである。

2.2.2 平坦波面長 r_0 (Fried length)

フリード長とも呼ばれる。大気揺らぎを通して来ることによって光の波面は乱される。小さな空間スケールでは局所的に波面の乱れを無視できるのだが、その乱れが無視できる程度の大きさを平坦波面長と呼ぶ。平坦波面長は次式で表される。

$$r_0 = \left[\frac{0.423k^2}{\cos\gamma} \int C_N^2(h)dh \right]^{-\frac{3}{5}} \quad (2.2)$$

$$\propto \left[\frac{1}{\lambda^2 \cos\gamma} \int C_N^2(h)dh \right]^{-\frac{3}{5}} \quad (2.3)$$

ここで、 k は光の波数で $k \propto \lambda^{-1}$ の関係があることから、 $r_0 \propto \lambda^{6/5}$ という波長依存性があり、長い波長ほど大きくなる。マウナケア山頂の平坦波面長は波長 500nm で約 17cm、波長 2.2 μ m では約 90cm となる [8]。

2.2.3 等波面離角 θ_0 (isoplanatic angle)

大気ゆらぎが共通とみなせる角度領域のことをいう。これは平坦波面長 r_0 と次のような関係がある。

$$\theta_0 \propto \cos\gamma \frac{r_0}{h} \quad (2.4)$$

ここで、 γ は天頂角、 h は乱流の特徴的な高さを表している。

2.2.4 コヒーレンス時間 τ_0 (coherence time)

これは、光路変動が互いに 1 ラジアン位の相収差未満にある時間間隔を表す。これにより、必要な AO 時間補正幅が定義され、可視光では数ミリ秒である。

これはフリード長と大気の典型的な風速 v を用いて次のように表される。

$$\tau_0 = 0.31 \frac{r_0}{v} \quad (2.5)$$

2.3 可視光補償光学の難しさ

以上のパラメータの定義より、口径 D の望遠鏡を $\frac{\pi}{4} \left(\frac{D}{r_0}\right)^2$ 個の小開口に分割すれば、小開口内の波面は、小開口ごとにそれぞれ傾きは異なっているが、それぞれ平面波とみなすことができる。したがって、補正に必要な波面の分割数、つまり補償光学系の素子数は r_0 の 2 乗すなわち波長の 12/5 乗に反比例する。リアルタイム補正に必要な速度は τ_0 の定義から r_0 に反比例するため、それぞれの小開口の波面の測定に使用できる小開口ごとの光量は r_0 の 3 乗すなわち波長の 18/5 乗に反比例する。そのため、波長が短くなると波面の測定が急激に困難になることがわかる。

2.4 frozen-flow 仮定

地球の大気揺らぎは層状に分布していることが知られている(下図)。frozen-flow 仮定とは、大気の屈折率の変動は、乱流が望遠鏡の視野を通過するのにかかる時間に匹敵する時間スケールで空間的に安定しているというものである (Taylor 1938)。この仮定は、一般的に予測制御や AO シミュレーション等で使用されていますが、この仮説を検証するために、いくつかの実験的研究 (引用多数) がなされたが、この仮定の下で波面が予測できる時間スケールの結果は一致していない。ただし、数十ミリ秒間の安定性は一般的に受け入れられている。層の速度ベクトルの安定性については、時間スケールは数秒から数分、中には数時間にも及ぶ。

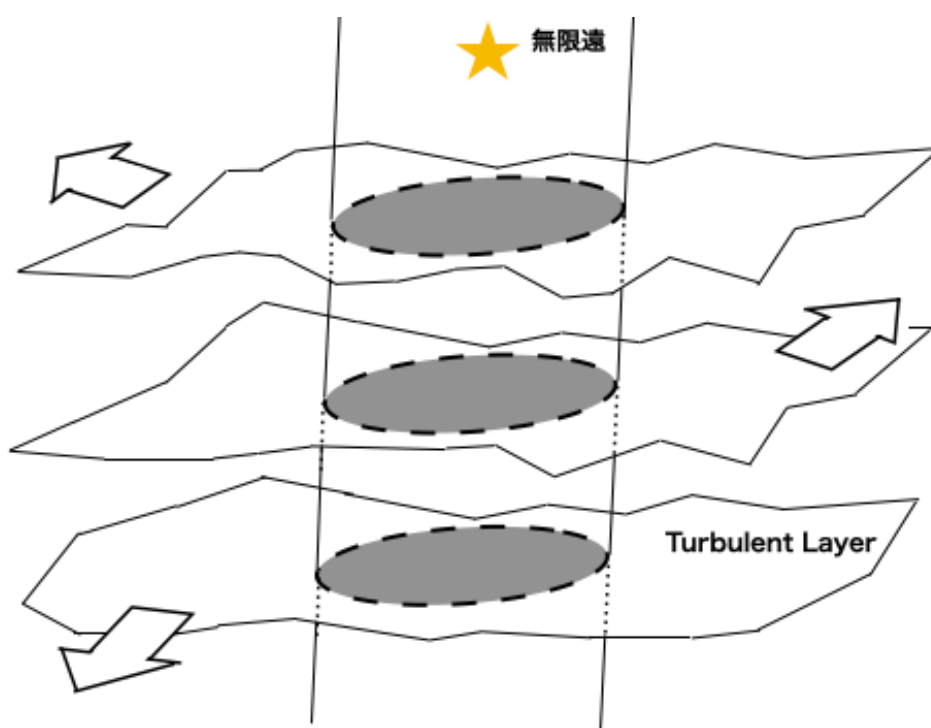


図2.1 地球大気概念図

図2.2に、Hardy 1998[5]における典型的な大気ゆらぎの高さ分布を示す。大気ゆらぎ強度の分布の特徴として、地上付近で最も強いことが分かる。これは大気が地表と摩擦を起こすことに起因する。一方、高度が上がるにつれて弱くなっており、典型的な大気圧下では上空 20km よりも高い位置にある大気ゆらぎの影響は無視できるほど小さくなる。補償光学において、これ以上の標高の大気揺らぎは無視している。また、上空 10~12km 付近にはゆらぎ強度が局所的に高くなる場所がある。これは対流圏 (Stratosphere) と成層圏 (Troposphere) との境界領域 (対流圏界面; tropopause) において吹いているジェット気流に起因するものである。これらの典型的な特徴に加え、緯度や地形、気温などのパラメータが複雑に絡み合うことで観測サイトや季節、昼夜によって

異なる大気の特徴が作り出されている。

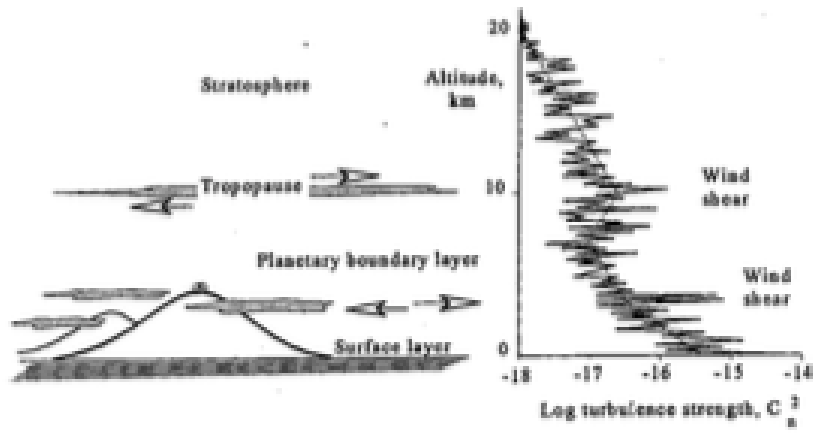


図2.2 大気揺らぎの高さ分布

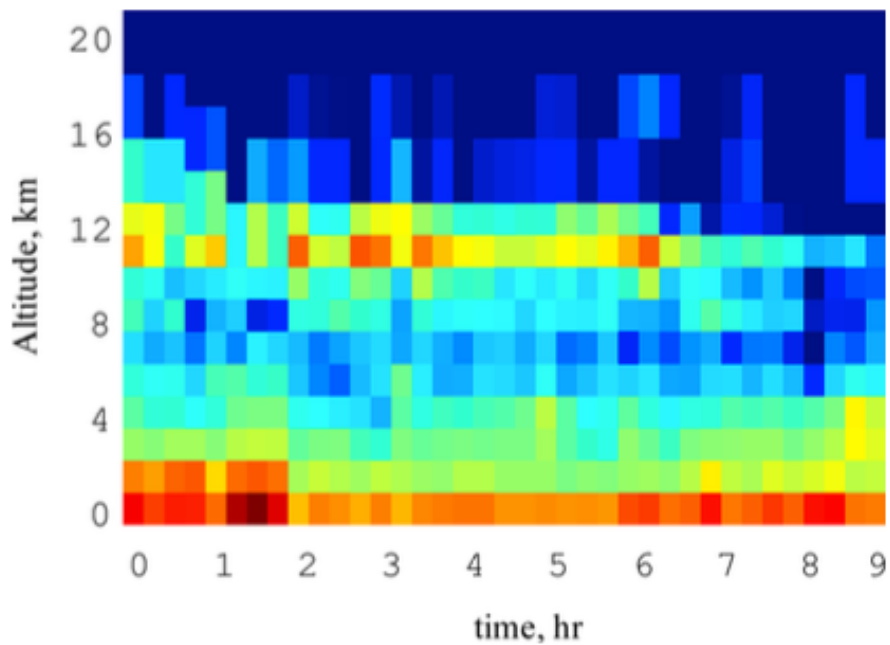


図2.3 大気揺らぎの時刻別高さ分布 [7]

図2.4は気温の高度分布と大気層の区分の模式図である。この図から、対流圏境界の高度がおおよそ高度 10 km 付近であることが確認できる。

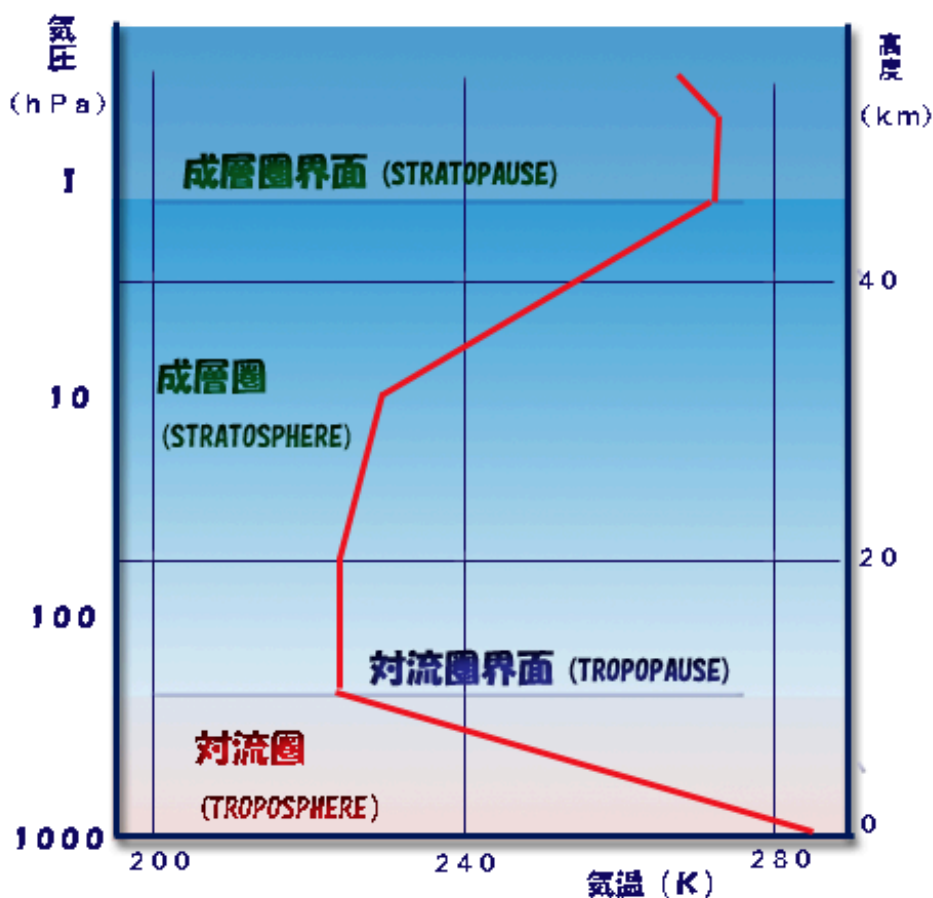


図2.4 気温の高度分布と大気層の区分の模式図
[\[https://www.jma.go.jp/jma/kishou/known/whitep/1-1-1.html\]](https://www.jma.go.jp/jma/kishou/known/whitep/1-1-1.html)

2.5 Tip-Tilt 成分

波面補正は、Tip-Tilt 成分と、より周波数の高い波面成分に分けて行われる。上記のように、補償光学に用いるガイド星は自然ガイド星とレーザーガイド星の二種類を用いる。下図のように、レーザーガイド星 (LGS) は望遠鏡の側に付いている装置から発射されたレーザーを用いて生成される。レーザーの往復により大気揺らぎで生じる Tip-Tilt 成分は相殺されるので大気揺らぎの Tip-Tilt 成分を測定することができず、Tip-Tilt 成分を測定するためには LGS が必要となる。

また、大気ゆらぎの強度と空間周波数の関係を示したものが図2.6である。大気ゆらぎが Kolmogorov 乱流と呼ばれるモデルで表されるとき、 $-11/3$ のべきで落ちていくことが知られており、この図から、空間周波数が小さいほど、ゆらぎの強度が大きくなり、波面のゆらぎの支配的な成分になっていることが示唆される。

波面の tiptilt 成分を知るために、天体を測定したときの、検出器上での明るさ重心の位置測定

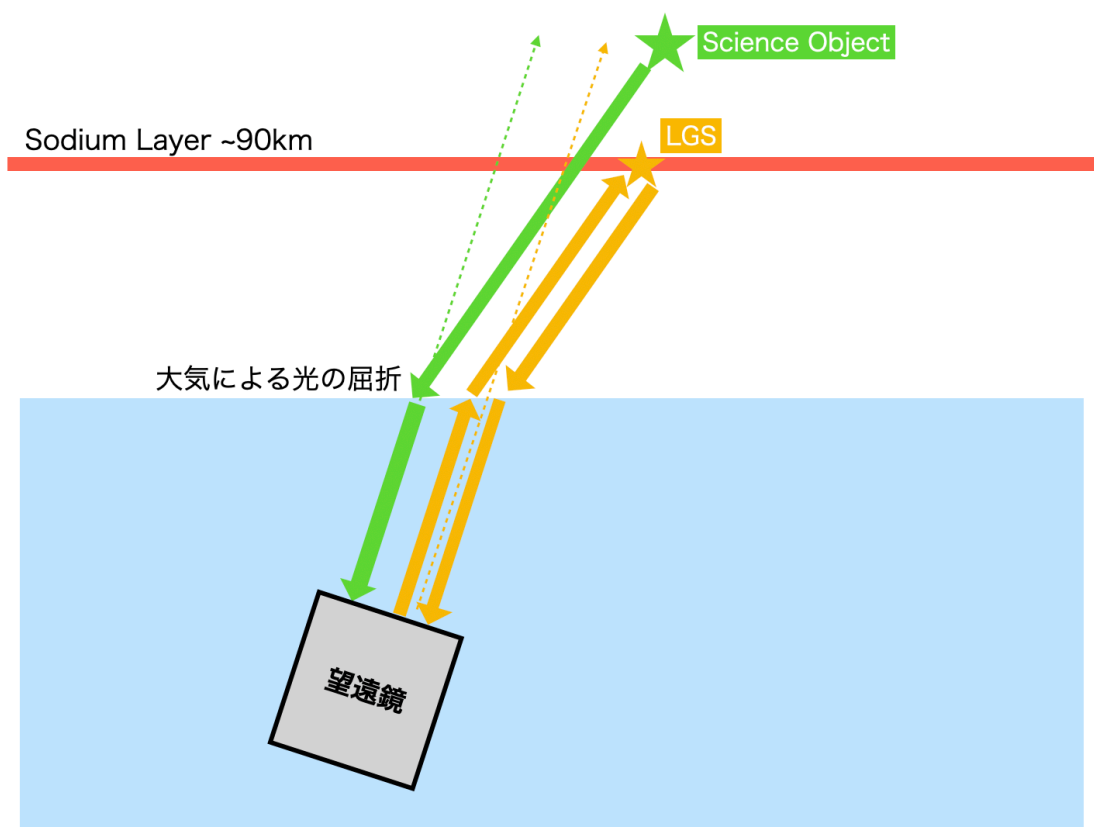


図2.5 Tip-Tilt 成分

が行われる。ここで、tip-tilt 成分の補正も同様に補償光学系のループを回すことで、補正が行われている。すると、原理的に temporal error が発生してしまう。

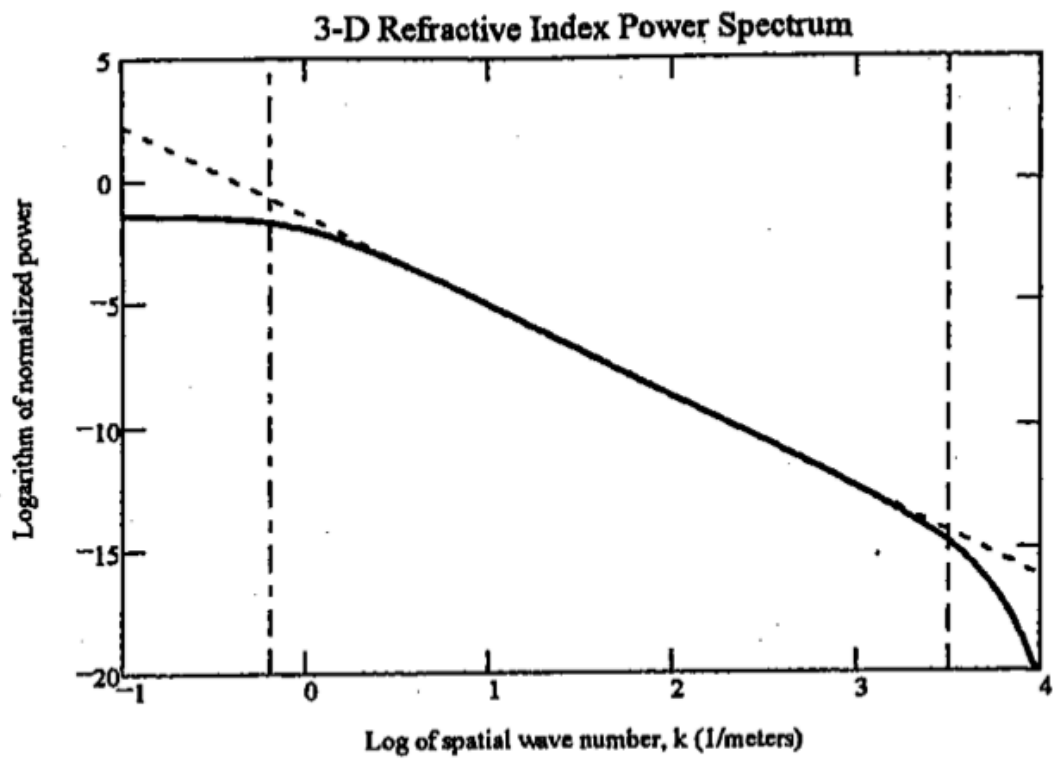


図2.6 大気ゆらぎの強度と空間周波数の関係

第3章

時系列解析

時系列解析とは、時間軸に沿って変化している現象の解析のことをいい、過去の自身のデータから未来のデータを予測する為に用いられる手法である。言い換えれば、過去のデータから、その変動の傾向・周期・不規則さなどを統計的もしくは確率的な手法を用いて記述、モデリングおよび予測する手法である。近年この手法は、マーケティング分野や製造業、セキュリティ分野、ロケットの軌道推定、ロボットの制御といった様々な分野で広く利用されている。本論文はこの時系列解析を大気揺らぎの分析に用いて、その予測精度を評価する。

時系列解析の手法について、以下 [6] に従って説明する。

3.1 時系列データ

前述の通り、補償光学における temporal error の改善のため、将来の観測値を予測することを考える。そのためには存在しないものと過去の値との自己相関を評価する必要がある。これはデータに何らかの構造を仮定する必要がある、時系列解析では確率過程を導入する。

具体的には時系列データを、ある確率法則を背景に形成されている実現値であるとみなす。つまり時系列データ $\{u_t\}_{t=1}^T$ はある確率変数列 $\{u_t\}_{t=-\infty}^{\infty}$ の実現値の一つであると考え。この確率変数列を確率過程といい、時系列解析では、確率過程の構造を時系列モデルという。

また、得られた観測値をモデルが最も再現しやすくなるようにパラメータを推定することを最尤法と言い、最小二乗法では推定が難しい構造を持ったモデルの推定に使われることが多くある。

時系列データで用いられる代表的な統計量を以下に示す。

- 期待値
- 平均
- 分散
- 標準偏差
- 自己共分散 (自己相関)

これらの統計量は時系列解析の中でも重要な役割を果たす。例えば、時系列データ u_t の平均的な値や変動幅の予測というのは将来の u_t の期待値と分散 (標準偏差) の評価をしている事になる。また、自己相関は時系列解析特有の統計量であり、時系列データを扱う際には非常に重要な役割を果たす。

3.2 時系列モデル

上述の様に、時系列解析の目的は過去のデータを基に未来の動きを予測する手法である。ここでは、こういったアルゴリズムで予測を行なっているのかを、具体的な数式で示す。

3.2.1 AR モデル

ある時点の出力が過去の出力の線形結合として得られる場合、これを表すモデルを自己回帰モデル (AR model; Autoregressive Model) と呼ぶ。

ある時点 t の観測値 y_t について、確率変数を用いて次の様に定式化できる。

$$y_t = c + \phi_1 y_{t-1} + \epsilon_t, \quad \epsilon_t \sim W.N.(\sigma^2) \quad (3.1)$$

ただし、 $W.N.(\sigma)$ はホワイトノイズであり、分散 σ に従う正規分布である。上式は、1 時点前の自分自身を説明変数とした (単) 回帰モデルと考えることができるため、AR モデルと呼ばれる。なお、AR モデルでは考慮するラグ数 (式中の p) を次数と呼ぶ。式 (上) では、被説明変数と説明変数はラグ 1 の関係であるため、1 次の AR モデルと呼び、AR(1) モデルと表記する。

AR(1) 過程の性質を詳しく見ていく。切片 c と自己回帰係数 ϕ_1 が既知の場合には y_1 も過去の情報であるため、 y_t に新たな情報を与える要素は ϵ_t のみである。なお、 ϵ_t はホワイトノイズを過程しており自己相関がなく、過去の $\epsilon_{t-1}, \epsilon_{t-2}, \dots$ は ϵ_t に影響を及ぼさない。したがって、AR(1) モデルにおいて、 y_t の構成要素は以下のように分解できる。

- 過去の情報から決まる項 : $c + \phi_1 y_{t-1}$
- 過去とは無関係に確率的に与えられる項 : ϵ_t

また、AR(1) モデルにおいて、現在時点 p 時点さかのぼった y_{t-p} と y_t の関係性を考える。

$$y_t = c + \phi_1 y_{t-1} + \epsilon_t \quad (3.2)$$

$$= (1 + \phi_1)c + \phi_1^2 y_{t-2} + \epsilon_t + \phi_1 \epsilon_{t-1} \quad (3.3)$$

$$\dots \quad (3.4)$$

$$= \left(\sum_{k=0}^{p-1} \phi_1^k \right) c + \phi_1^p y_{t-p} + \sum_{k=0}^{p-1} \phi_1^k \epsilon_{t-k} \quad (3.5)$$

これより、 y_{t-p} は p 時点前のデータ y_t に ϕ_1^p を乗じた影響を受けていることが分かる。また、AR(1) モデルの平均 $E[y_t]$ 、分散 $Var[y_t]$ 、自己相関係数 ρ_k は次のようになる。

- $E[y_t] = \frac{1}{1-\phi_1}c$
- $Var[y_t] = \frac{\sigma^2}{1-\phi_1^2}$
- $\rho_k = \phi_1^k$

以下に、一般化した AR(p) モデルの性質をまとめた。AR(p) モデルは次のように定式化される。

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \epsilon_t, \epsilon_t \sim W.N.(\sigma^2) \quad (3.6)$$

- 平均 $\mu = E[y_t] = \frac{c}{1-\sum_{i=1}^p \phi_i}$
- 分散 $\gamma_0 = Var[y_t] = \frac{\sigma^2}{1-\sum_{i=1}^p \phi_i \rho_i}$
- 自己相関係数 $\rho_k = \sum_{i=1}^p \phi_i \rho_{k-i}, k \geq 1$
- AR 過程の自己相関は指数関数的に減衰する

3.2.2 MA モデル

ある時点の出力が過去および現在のホワイトノイズの線形和で表される場合、これを移動平均モデル (MA model; moving average model) と呼ぶ。従って、MA モデルは AR モデルとは異なり、ある一定の値の周りをランダムに動くモデルとなる。1 次の MA 過程 (MA(1) 過程) を数式で表すと以下になる。

$$y_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1}, \epsilon_t \sim W.N.(\sigma^2) \quad (3.7)$$

上式をみると、 y_t のモデルと y_{t-1} モデルが ϵ_{t-1} という共通項を持つので、 y_t と y_{t-1} の間に相関が生じる。つまり、MA モデルは1 次の自己相関を持つモデルとなっている。また、上式をみると θ_1 が1 次自己相関の強さを決めていることもわかる。 θ_1 が大きければ1 次の自己相関が大きくなるためグラフは滑らかになり、 θ_1 が負の場合は AR モデルと同様に出力値が反転する。加えてホワイトノイズの期待値は時点によらず0 という性質から、系列が μ の周りを変動することもわかる。したがって、MA(1) 過程の期待値は μ であることが予想されるが、以下の様に確認できる。

$$E[y_t] = E[\mu + \epsilon_t + \theta_1 \epsilon_{t-1}] \quad (3.8)$$

$$= E[\mu] + E[\epsilon_t] + E[\theta_1 \epsilon_{t-1}] \quad (3.9)$$

$$= \mu \quad (3.10)$$

$$(3.11)$$

最後の等号はホワイトノイズの期待値が0 であるという性質より成り立つ。

また MA(1) 過程では、攪乱項の分散よりも系列の分散が大きくなる、という性質がある。この性質は以下の様に確認することができる。

$$\gamma_0 = \text{Var}[y_t] \quad (3.12)$$

$$= \text{Var}[\mu + \epsilon_t + \theta_1 \epsilon_{t-1}] \quad (3.13)$$

$$= \text{Var}[\epsilon_t] + \theta_1^2 \text{Var}[\epsilon_{t-1}] + 2\theta_1 \text{Cov}[\epsilon_t, \epsilon_{t-1}] \quad (3.14)$$

$$= (1 + \theta_1^2) \sigma^2 \quad (3.15)$$

$$(3.16)$$

最後の等号はホワイトノイズの自己共分散は0、時点によらず分散は一定という性質より成り立つ。したがって、MA(1) 過程の分散は $\theta_1^2 \sigma^2$ の分だけ、攪乱項の分散よりも大きくなる。

次に、MA(1) 過程の自己共分散および自己相関を計算する。1 次自己共分散はホワイトノイズの性質を利用して次式で表現できる。

$$\gamma_1 = Cov[y_t, y_{t-1}] \quad (3.17)$$

$$= Cov[\mu + \epsilon_t + \theta_1 \epsilon_{t-1}, [\mu + \epsilon_{t-1} + \theta_1 \epsilon_{t-2}]] \quad (3.18)$$

$$= Cov[\epsilon_t, \epsilon_{t-1}] + Cov[\epsilon_t, \theta_1 \epsilon_{t-2}] + Cov[\theta_1 \epsilon_{t-1}, \epsilon_{t-1}] + Cov[\theta_1 \epsilon_{t-1}, \theta_1 \epsilon_{t-2}] \quad (3.19)$$

$$= \theta_1 Cov[\epsilon_{t-1}, \epsilon_{t-1}] \quad (3.20)$$

$$= \theta_1 \sigma^2 \quad (3.21)$$

$$(3.22)$$

したがって、MA(1) 過程の1次自己相関は

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\theta_1}{1 + \theta_1^2} \quad (3.23)$$

となる。この式から $\theta_1 = \pm 1$ のときに、MA(1) 過程の自己相関の絶対値が最大値 $1/2$ をとることがわかる。つまり、MA(1) 過程では1次自己相関の絶対値が $1/2$ より大きな過程はモデル化できないことを意味する。さらに、MA(1) 過程の自己共分散の性質についてみていこう。1次自己共分散の計算と同様に、2次以降の自己共分散を計算する。 $k \geq 2$ とすると、

$$\gamma_k = Cov[y_t, y_{t-k}] \quad (3.24)$$

$$= Cov[\mu + \epsilon_t + \theta_1 \epsilon_{t-1}, \mu + \epsilon_{t-k} + \theta_1 \epsilon_{t-k-1}] \quad (3.25)$$

$$= 0 \quad (3.26)$$

この結果は、MA(1) 過程の2次以降の自己相関が0になることを意味している。逆にいえば、MA(1) 過程は1次自己相関をモデル化することはできるが、2次以降の自己相関を記述することはできない。

一般的に q 次移動平均過程 (MA(q) 過程) は次で定義される。

$$y_t = \mu + \epsilon_t + \sum_{i=1}^q \theta_i \epsilon_{t-i}, \quad \epsilon_t \sim W.N.(\sigma^2) \quad (3.27)$$

以上より MA(1) 過程の性質は次の様にまとめられる。

- $E[y_t] = \mu$
- $\gamma_0 = Var[y_t]$
-

$$\rho_k = \begin{cases} \frac{\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2}, & 1 \leq k \leq q \\ 0, & k \geq q + 1 \end{cases}$$

- MA 過程は常に定常

これらの性質のうち重要なのは MA 過程は常に定常であることと、MA(q) 過程の $q+1$ 次以降の自己相関が0になることを意味している。また、MA(q) 過程は観測できないホワイトノイズの

線形和で表されるので、モデルの解釈が難しいという問題が挙げられる。加えて、同様の理由により、モデルの推定や予測が複雑になるという問題もある。

3.2.3 ARMA モデル

自己回帰移動平均モデル (ARMA model; autoregressive moving average model) は AR モデルと MA モデルの両者の性質を含んだモデルであり、AR モデルに残差の移動平均を加えたモデルである。ARMA モデルで扱う時系列は確定的でも確率的でも良いが、特に入力 ϵ_t がホワイトノイズであるとき、ARMA モデルの出力は自己回帰移動平均 (ARMA) 過程と呼ばれる。

ARMA 過程は自己回帰項と平均移動項を含んだ過程である。(p,q) 次 ARMA 過程 (ARMA(p,q) 過程) は次の様に定義できる。

$$y_t = c + \epsilon_t + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{i=1}^q \theta_i \epsilon_{t-i}, \quad \epsilon_t \sim W.N.(\sigma^2) \quad (3.28)$$

ここで、 p と ϕ_i はそれぞれ自己回帰の次数および自己回帰係数であり、 q と θ_i は移動平均の次数および平均移動係数である。ARMA(p,q) 過程は、AR 過程と MA 過程の性質を併せもっており、両過程の性質のうち強いほうが ARMA 過程の性質として現れる。例えば、定常性を考えると、MA 過程は常に定常であるが、AR 過程は常に定常であるとは限らない。この場合、AR 過程の性質が ARMA 過程の性質として現れ、ARMA 過程は定常になるとは限らない。次に、ARMA(p,q) 過程の性質を示す。なお、この性質は ARMA(p,q) 過程が定常であることを前提とする。

- ARMA モデルの期待値は AR モデルにおける期待値と等しくなる (ホワイトノイズの期待値が0であるため)

$$\mu = E[y_t] = \frac{c}{1 - \sum_{i=1}^p \phi_i} \quad (3.29)$$

- ARMA 過程の自己相関は次数に関して指数的に減衰する

3.2.4 ARIMA モデル

前述した ARMA モデルは定常過程にしか対応していなかったが、自己回帰和分移動平均モデル (ARIMA model; autoregressive integrated moving average model) は ARMA モデルを非定常過程に対応させたものである。データの差分をとることで、トレンドを除去し定常過程に変換する操作が加わっている。

ARIMA モデルは ARMA モデルに加えて、データ間の差分 d を定義する。この差分をとる系列は和分 (Integrated) と呼ばれ、 $I(d)$ で表される。AR、MA 過程と合わせて、ARIMA (自己回帰和分移動平均) 過程、もしくは ARIMA(p,q,d) と呼ばれる。

和分過程と ARIMA 過程の定義を以下に示す。

- 和分過程 : $d-1$ 階差分をとった系列は非定常過程であるが、 d 階差分をとった系列が定常過程に従うものは、 d 次和分過程もしくは $I(d)$ 過程と呼ばれる。また、 $I(d)$ 過程は定常過程で定義される。
- ARIMA 過程 : d 階差分をとった系列が定常かつ反転可能な ARMA(p,q) 過程に従う過程は次数 (p,d,q) の ARIMA 過程、もしくは ARIMA(p,q,d) 過程と呼ばれる。

ARIMA(p,d,q) 過程は次のように表現される。

$$\Delta^d y_t = \sum_{i=1}^p \phi_i \Delta^d y_{t-i} + \Delta^d \epsilon_t + \sum_{j=1}^q \theta_j \Delta^d \epsilon_{t-j} \quad (3.30)$$

ここで、

$$\Delta^1 y_t = y_t - y_{t-1} \quad (3.31)$$

$$\Delta^2 y_t = \Delta^1 y_t - \Delta^1 y_{t-1} = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) \quad (3.32)$$

$$= y_t - 2y_{t-1} + y_{t-2} \quad (3.33)$$

$$\Delta^d y_t = \Delta^{d-1} y_t - \Delta^{d-1} y_{t-1} \quad (3.34)$$

である。

最後に ARIMA 過程の性質を示す。

- 平均 : $\mu = E[y_t] = \text{constant}$
- 分散 : $\gamma_0 = \text{Var}[y_t] = E[(y_t - \mu)^2] = \text{constant}$
- 自己共分散 : $\gamma_k = \text{Cov}[y_t, y_{t+k}] = E[(y_t - \mu)(y_{t+k} - \mu)] = \text{constant}$

ARIMA モデルを適用するかどうかの判定は、トレンドなどの平均値揺動が存在するかどうかによる。1次階差、2次階差を計算した結果、平均値揺動が解消されているかどうかを検討し、トレンドが見られなくなったところで d の値を決定すればよい。ここで、平均値揺動とは、時系列データの平均値が時刻と共に変化することを表す。

3.3 統計的モデリング

時系列解析などの統計解析においては、偶然をとまなう現象はある確率分布に従う確率変数の実現値とみなされる。統計的モデルや時系列モデルは、この確率分布をデータに基づいて推定したものである。この章では、真の確率分布とモデルの確率分布の近さを評価する基準を導入し、その基準に基づくパラメータ推定やモデルの選択など、統計的モデルの構成のための統一的な接近法を示す。

3.3.1 最尤法

最尤法は、最小二乗法とともに確率モデルのパラメーターを推定するのによく用いられる方法である。最尤法は、特に最小二乗法では推定が難しい複雑な構造をもったモデルの推定に用いられることが多い。最小二乗法ではモデルで説明できない部分が最小になるようにパラメータを推定したが、最尤法では得られた観測値をモデルが最も再現しやすくなるようにパラメーターを推定する。これにより、最尤法は最小二乗法では解くことができない複雑なモデルのパラメーターを推定することができる。(後述)

3.3.2 赤池情報量基準

時系列データの将来を予測するには、ある時系列モデル(例えば ARIMA モデルなど)を仮定し、手持ちのデータから正しいモデルを選択する必要がある。ARIMA モデルを過程した場合、その次数 (p,d,q) を決める必要があるのだが、これは赤池情報量基準 (AIC; Akaike Information Criterion) によって決定する [10]。情報量基準とは、最尤法の推定結果をもとに最適なモデルを選択する客観的な基準である。この情報量基準が最小になるモデルが最適なモデルとなる。次数が高い (p,q が大きい) モデルは既知のデータへの当てはまりがよくなるが、うまく予測ができなくなってしまう。逆に次数の低いモデルはデータへの当てはまりが悪くなるが、未知のデータに当てはまることがある。AIC は、端的に言えば、当てはまりの良さとモデルの単純さ(次数の低さ)の間をとった情報量基準といえる。

大気ゆらぎの時系列データに対して、図3.1に示すように (p,d,q) の値をグリッドサーチして決定した。

```
info_df.sort_values(by="AIC").head()
```

	p	d	q	AIC
30	3	0	2	1542.39
31	3	0	3	1544.3
32	3	0	4	1545.01
51	5	0	3	1545.99
42	4	0	4	1546.02

図3.1 AIC のグリッドサーチの一例

3.4 Box-Jenkins 法

Box-Jenkins 法とは、モデル作成・予測手順を次の三段階で行う方法である。

1. 自己相関係数や自己偏相関係係数をみて自己回帰の次数決定やデータの定常化を行いモデルを特定する。
2. 特定したモデルのパラメータを平均平方誤差が最小になるよう推定する。
3. モデルを予測に使用する。

3.5 Python による時系列データ解析

Python は 1991 年にオランダ人のガイド・ヴァンロッサムというプログラマーによって開発されたプログラミング言語で、学習も比較的容易である。その特徴として、コードを簡潔に書けること、専門的なライブラリが豊富にあることなどが挙げられる。

本研究では、Python のライブラリである statsmodels を使用して、時系列分析の基本である Box-Jenkins 法を用いた実装を行った。なお、以下のバージョンを用いて解析を行った。

- Python 3.7.1
- Matplotlib 3.1.1
- Numpy 1.18.5
- Pandas 0.25.0
- StatsModels 0.13.0

第4章

制御工学

制御 (control) とは、「ある目的に適合するように、対象となるものに所用の操作を加えること」と定義される。補償光学系を設計する際、天体からの光の物理学や大気ゆらぎなどの自然科学的な理論だけではなく、波面補正を担う可変形鏡に付属しているアクチュエータや波面の傾きを補正する Tip-Tilt Mirror といった、ハードウェア側の特性を知ることが必要不可欠である。

以下、制御工学の概要について文献 [13] にしたがって説明する。

4.1 制御系の分類

制御系 (control system) はシステムの構成による分類として、次のように分類される。

4.1.1 フィードフォワード制御 (feedforward control)

図4.1に示すように、システムに外乱が入力されたならば、システムは外乱が出力に影響を及ぼす前に先回りして打ち消すように動作を行うようにする制御系をいう。したがって、外乱と出力の因果関係を把握できていないと制御が難しい。

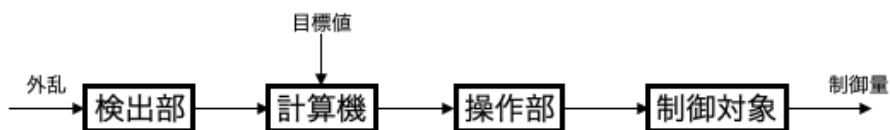


図4.1 フィードフォワード制御の構成例

4.1.2 フィードバック制御 (feedback control)

制御量をフィードバックさせ、目標値の値と比較させ、一致するように訂正動作を行う制御系をいう。図4.2に示すように一と+でフィードバックさせるものがあり、-の場合を負帰還 (negative

feedback)、+の場合を正帰還 (positive feedback) という。一般的には負帰還が使われることが多い。

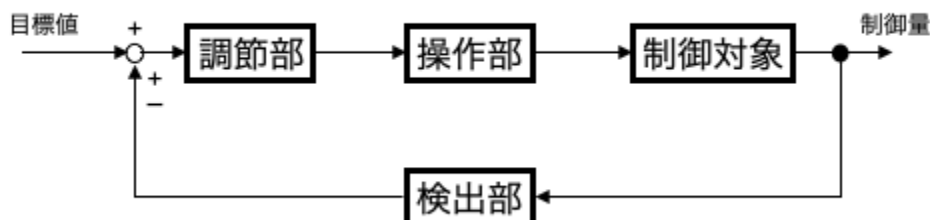


図4.2 フィードバック制御の構成例

フィードバック制御のような閉じたループ構成をもつシステムを「閉ループ制御系」、フィードフォワード制御のようにフィードバック経路をもたない制御系を「開ループ制御系」という。

4.2 Laplace 変換

$t \geq 0$ で与えられる時間関数 $f(t)$ に対して、区間 $[0, \infty]$ での積分が

$$F(s) = \mathcal{L}[f(t)] = \int_0^{\infty} f(t)e^{-st} dt \quad (4.1)$$

で定義されるとき、 $F(s)$ 、あるいは $\mathcal{L}[f(t)]$ を $f(t)$ のラプラス変換 (Laplace transform)、そして複素数である s をラプラス演算子 (Laplace operational calculus) という。

4.3 過渡応答

システムを知るには、規則的に変化する入力を要素やシステムに与え、その時の反応から特性を調べる。定常状態から、再び定常状態に落ち着くまでの時間的経過の反応を過渡応答 (transient response) という。この応答は要素やシステムの動特性を評価する方法に用いる。一般的に用いる規則的な入力信号として、次の3種類がある。

- ステップ応答
- インパルス応答
- ランプ入力

4.4 伝達関数

システムの入力側に信号（原因）を作用させると、その信号の影響を受けてシステムの状態を変化させる。このときの原因となる信号を入力信号（input signal）という。結果として、変化する状態を信号として取り出すのだが、それを出力信号という。信号の関係を複素領域で関数形式として表現するものが伝達関数である。図4.3に示すように、時間領域での入力 $u(t)$ 、出力 $y(t)$ の制御系を考える。この時間領域での入力と出力をラプラス変換し、 $U(s)$ 、 $Y(s)$ とする。そこで、入力 $U(s)$ に対する出力 $Y(s)$ の比をとった関数を $G(s)$ 伝達関数と定義できる。

$$G(s) = \frac{Y(s)}{U(s)} \quad (4.2)$$

ただし、ラプラス変換を施す際は、すべての初期値を 0 にする。したがって、ある要素の伝達関数が分かっているならば、入力 $u(t)$ による出力 $y(t)$ をラプラス逆変換によって示すことができる。

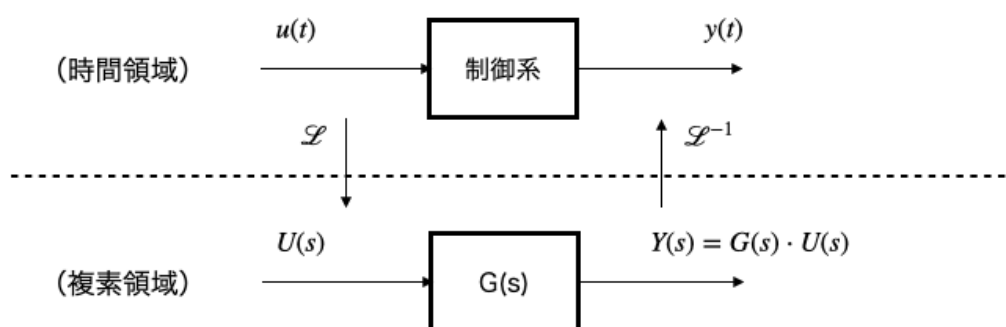


図4.3 伝達関数

一般的には、ある制御器の伝達関数を次の二次遅れ系と言われる形で仮定し、ステップ応答を与えてその制御器の過渡応答を測定し、各係数を求める。

$$G(s) = \frac{K\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (4.3)$$

$G(s)$ をもつ制御系にステップ入力を加え、ラプラス逆変換を施すと、時間領域での出力 $y(t)$ は次のような曲線になる。

$$y(t) = K \left\{ 1 - \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}} \sin \left(\omega_n \sqrt{1-\zeta^2} t + \tan^{-1} \frac{\sqrt{1-\zeta^2}}{\zeta} \right) \right\} \quad (4.4)$$

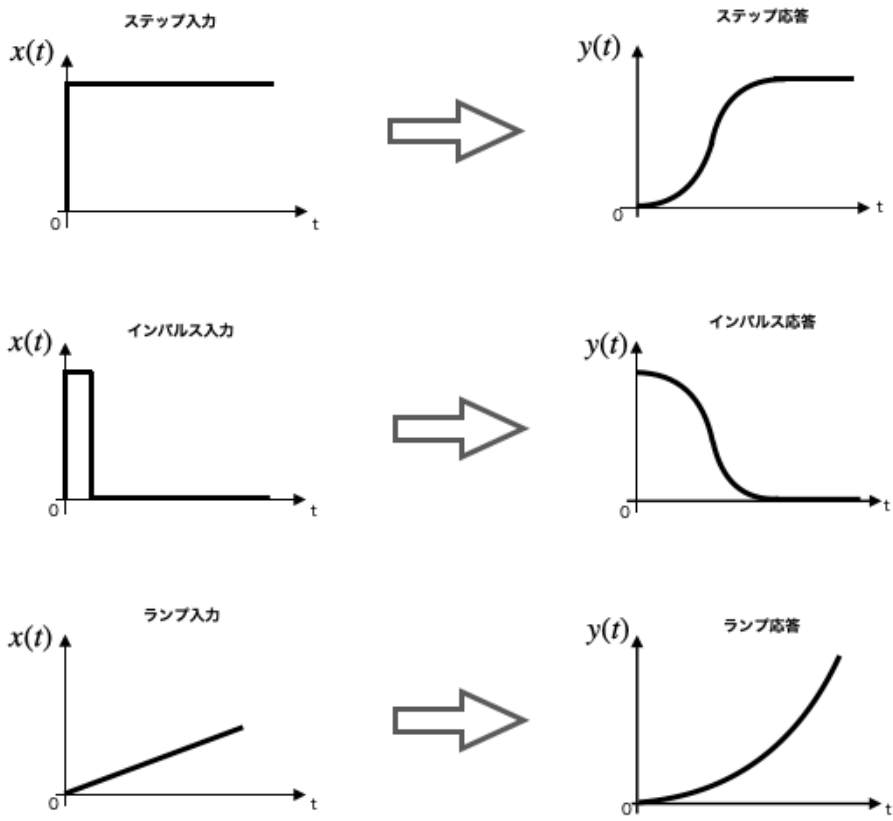


図4.4 入力信号に対する応答の関係

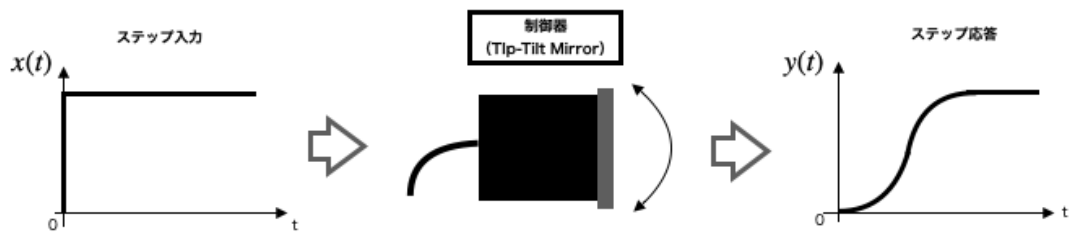


図4.5 入力信号に対する応答の関係 2

4.5 セットイング

Tip-Tilt Mirror は piezo 素子と呼ばれる材質でできており、信号として電圧を加えることで変形する。しかし、電圧を大きくした場合と小さくした場合の変位曲線は同じ軌跡を通らないという性質を持っている (ヒステリシス)。ここで、Servo system を使うことで、ヒステリシスを修正することができる (signal として位置情報を送る)。一方、このシステムを使うと、Tip-Tilt Mirror を所定の位置に動かすのに時間がかかってしまう。

本研究では Servo off、さらにステップ入力を用いて Tip-Tilt Mirror の応答を調べた。

第5章

データセット

本研究では、時系列データとして、東北大学 51cm 望遠鏡を用いて行った観測データを用いた。

5.1 東北大学 50cm 望遠鏡を用いた観測

観測は 2019 年 2 月 1 日 23 時頃、2 月 3 日 23 時 30 分頃に行なった。観測天体には Sirius を用いており、以下、それぞれを Sirius1, Sirius2 と表記する。露光時間は 11[msec] でデータを取得した。観測時の様子を図5.1に示す。また、図中の赤丸は PI 社の Tip-Tilt Mirror[S330-2SL] である。

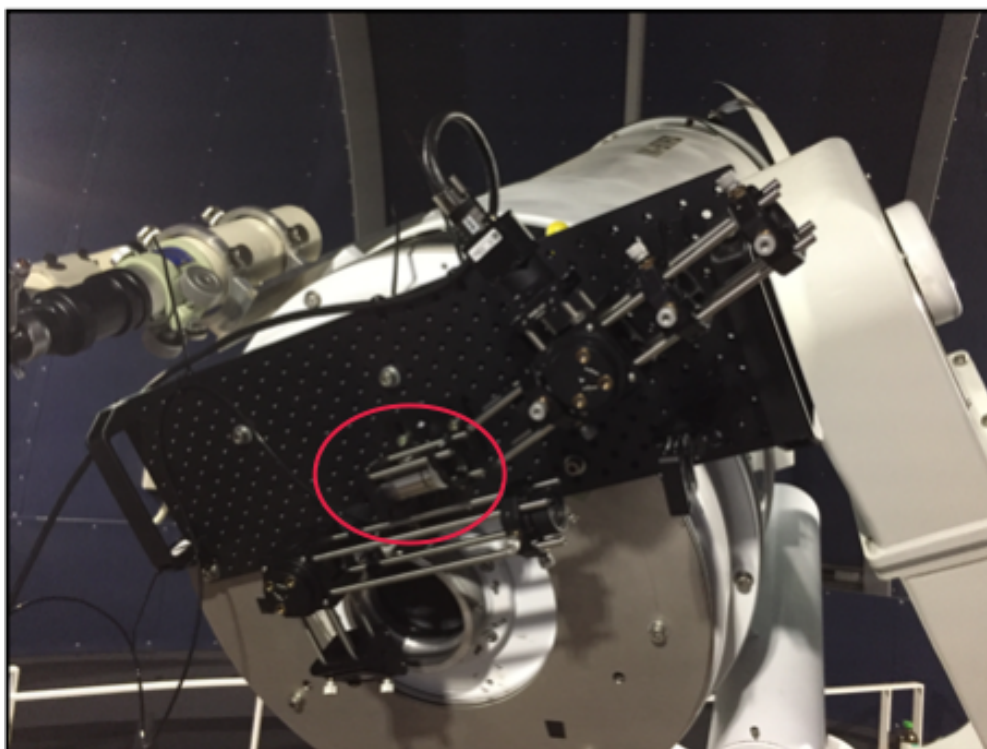


図5.1 東北大学 50cm 望遠鏡

波面の Tip-Tilt 成分として、天体を測定したときの、検出器上での明るさ重心の位置測定を行った。したがって時系列データは、 x_0, x_1, x_2, \dots または y_0, y_1, y_2, \dots として得られる。

Tip-Tilt 成分の補正も同様に補償光学系のループを回すことで、補正が行われている。すると、原理的に temporal error が発生してしまう。そこで、観測した時系列データから、明るさ重心が次に来るであろう位置を正確に予測できれば、予め Tip-Tilt 成分の動きを打ち消すように制御器を動かせる。そうすることで temporal error が小さくなり、それにより波面全体の補正精度の向上につながる。

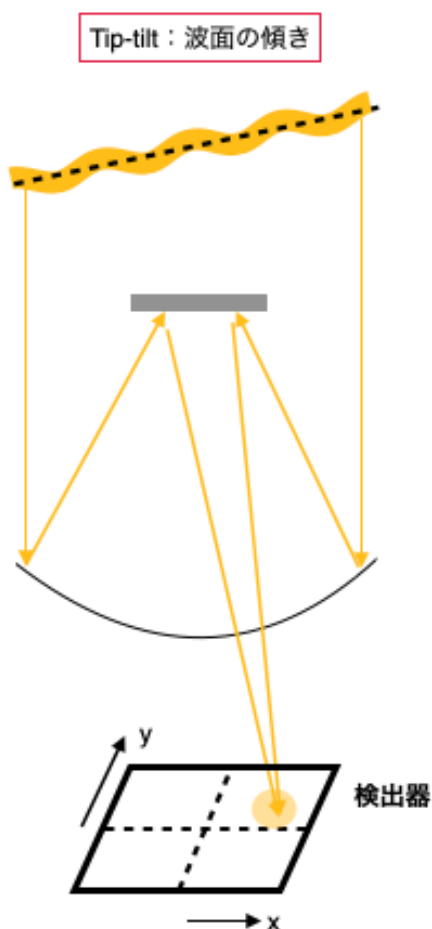


図5.2 波面の Tip-Tilt 成分

5.2 MATLAB

予測したデータを用いて大気ゆらぎを直接補正することを考える。そのシミュレーション手法として、MATLAB と呼ばれる数値解析ソフトウェアを用いてシミュレーションを行った。図5.3が、

従来の補償光学の概念図である。対して、右下の図は予測データを用いた新手法の概念図で、大気ゆらぎ (raw_data) を直接補正できるよう設計した。

シミュレーションとして、先ほどの大気ゆらぎを入れて、従来の補償光学と、新手法で得られた結果を、出力された残差を用いて評価した。

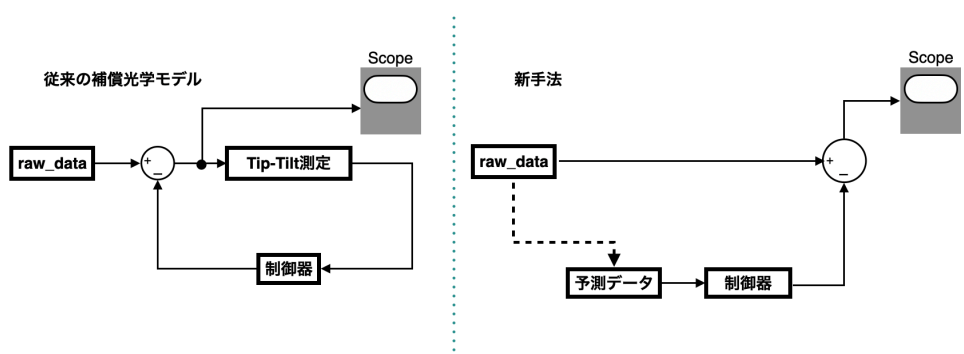


図5.3 MATLAB シミュレーション

第6章

結果

この章では、実際に取得したデータを用いて新しく提案する手法を実践した過程および結果について述べる。

6.1 コレログラム

時系列データにおいて、時間差が k である 2 つの時点の値の相関関数 r_k は k の値により異なるが、これを k の関数と考えたものを自己相関関数という。 k をいろいろ変えてできる r_k をコレログラム (correlogram) という。コレログラムを見ることで、データの周期性がわかる。 k が 0 のとき $r_k = 1$ であり、一般には k が大きくなると $|r_k|$ は 0 に近づいていく。すべての k について r_k を知るにより、その時系列においてどのような時間差をもつ 2 つの時点の値が関係が強い、あるいは弱いかわることができる。(付録 (A.1)) ここでは、大気ゆらぎのデータと、それに対応するコレログラムを示す。ただし、縦軸は自己相関係数、横軸はラグ数 (元データからずらした量) である。観測で得られた時系列データの時間幅は 11 ms であり、(横軸のラグ数) \times 11 ms の値だけ離れた時間との相関を示すグラフである。コレログラムの左から 1 番目はラグ 0 の自己相関係数を表していて、(ラグ 0) = (元データ) そのものなので自己相関係数は 1 になっている。

二つの異なる日付のデータを用いてコレログラムを作成したが、それぞれの日付に関して異なる特徴が見られた。Sirius1 の方では、緩やかに自己相関が減少するようなコレログラムが得られたが、Sirius2 では振動しながら減衰していくような特徴が見られる。時系列解析モデルとして本研究では ARIMA モデルを適用したが、ARIMA モデルをさらに拡張したモデルとして SARIMA モデルという時系列モデルが知られている。具体的には、時系列方向に ARIMA モデルを使い、さらに周期方向にも ARIMA モデルを使っているモデルで、季節変動がある非定常データに対し ARIMA モデルよりもデータにフィットする場合がある。Sirius2 のデータにおいて SARIMA モデル適用すれば、予測精度が向上する可能性がある。しかし、データに周期性が見られない Sirius1 では SARIMA モデルを適用する意味を見出せず、手法の統一性やモデル構築に要する時間をできるだけ短くするという観点において、SARIMA モデルの必要性は議論の余地がある。

また、それぞれのデータ取得において異なるのは日付だけであるが、周期性の有無のような違いが見られた。手法を確立させるためには、さまざまな日付、観測地のデータを取得する必要がある。

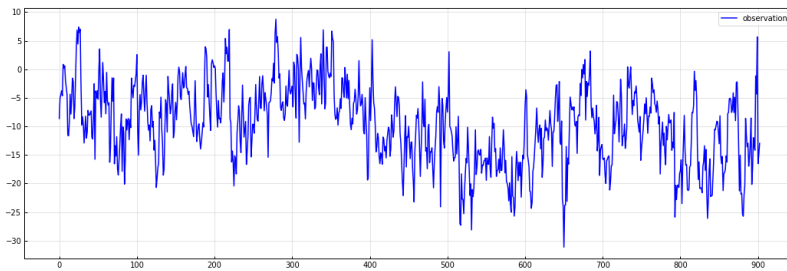


図6.1 大気ゆらぎの時系列データ (Sirius1,x 座標 [pixel])

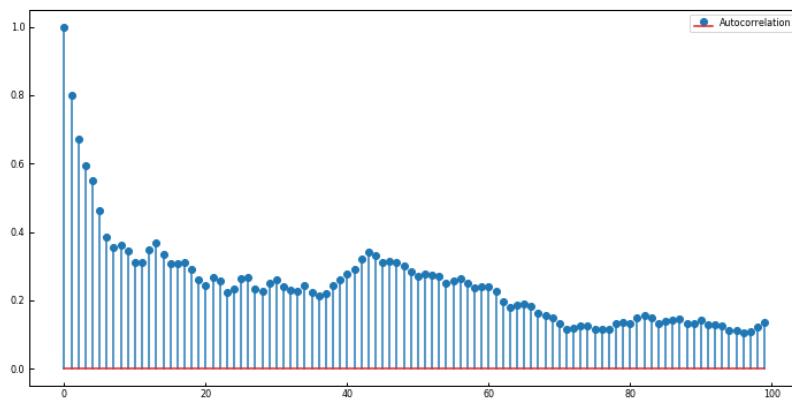


図6.2 大気ゆらぎのコレログラム

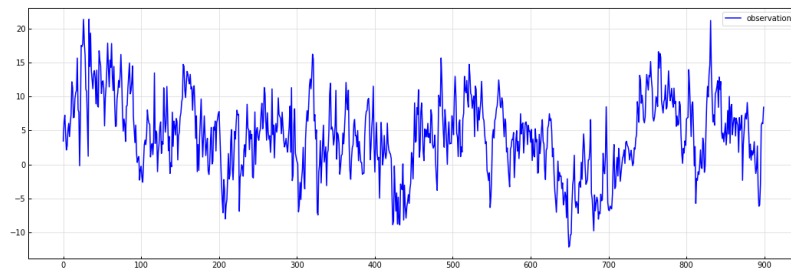


図6.3 大気ゆらぎの時系列データ (Sirius1,y 座標 [pixel])

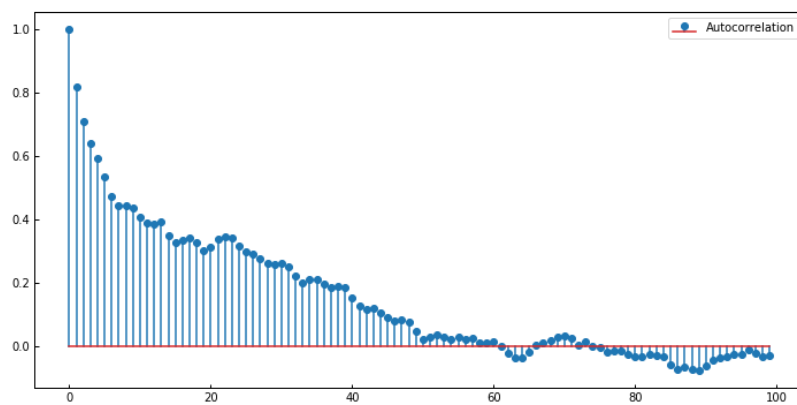


図6.4 大気ゆらぎのコレログラム

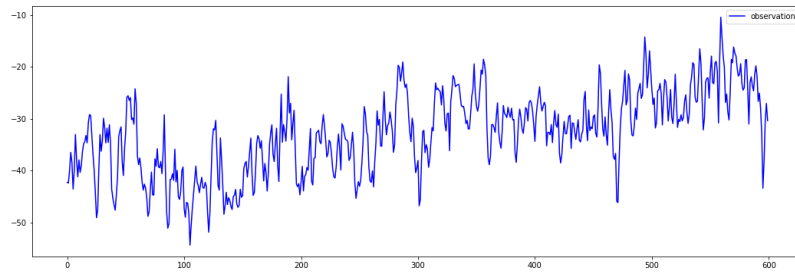


図6.5 大気ゆらぎの時系列データ (Sirius2,x 座標 [pixel])

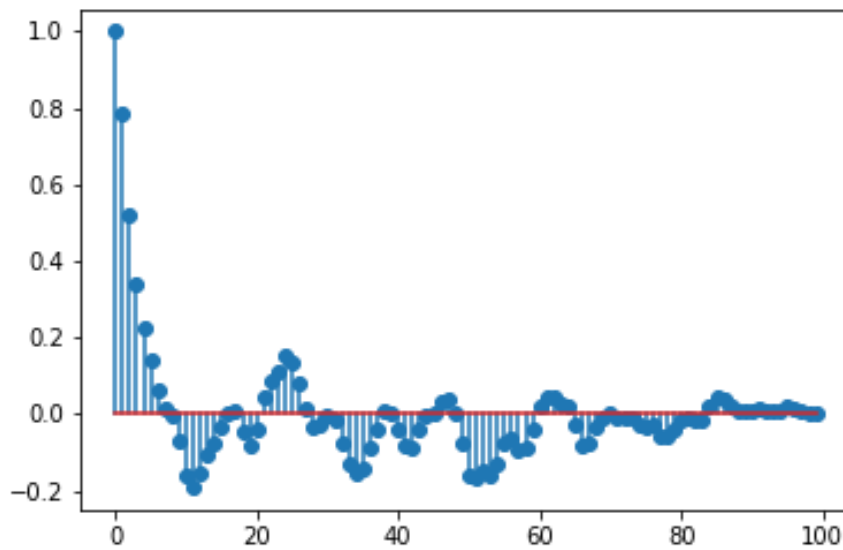


図6.6 大気ゆらぎのコレログラム

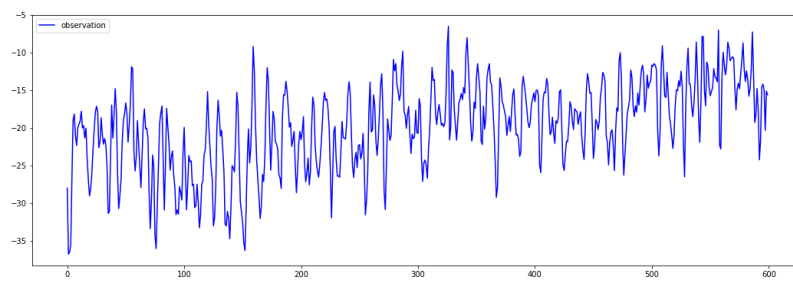


図6.7 大気ゆらぎの時系列データ (Sirius2,y 座標 [pixel])

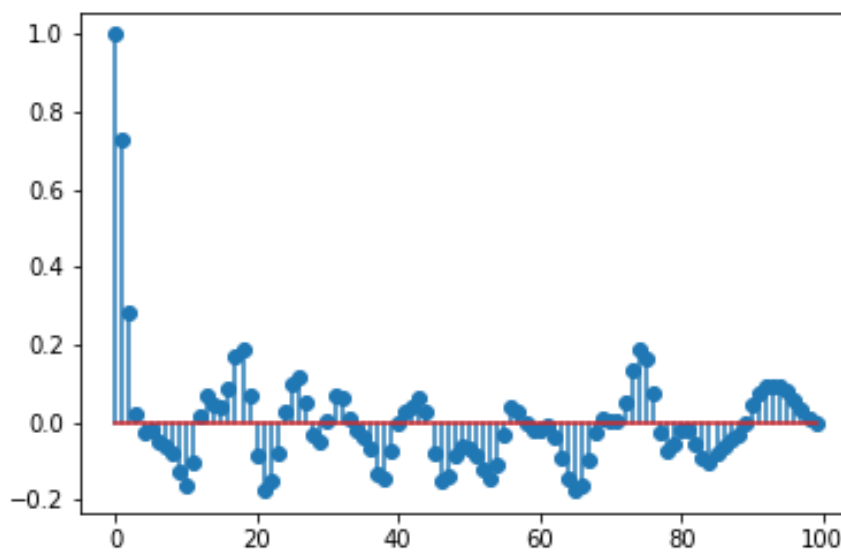


図6.8 大気ゆらぎのコレログラム

比較として、ランダムデータのコレログラムを示す。大気ゆらぎのコレログラムは、時点が近い点と明らかに相関が見られ、右肩下がりの傾向が見られるのに対し、ランダムデータのコレログラムには相関や周期性が見られないことがわかる。

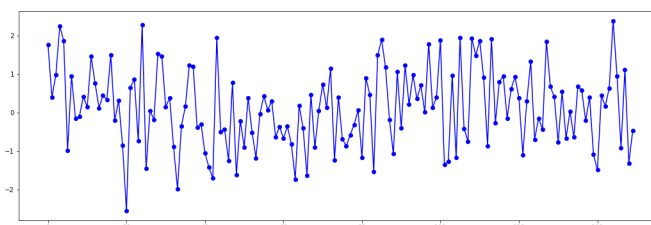


図6.9 ランダムデータ

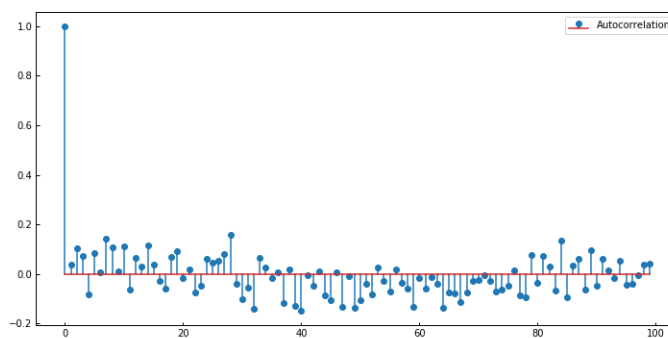


図6.10 ランダムデータのコレログラム

6.2 予測

図6.11は行った手法のイメージ図である。データの最初の300点、3.3秒分に当たるデータをモデルの次数(p,d,q)を決めるための入力データとし、そして決定したモデルを使って、入力データをずらしながら、最尤推定を行い、回帰係数 ϕ, θ の値を求め、一点先を予測する、という解析を繰り返し行った。

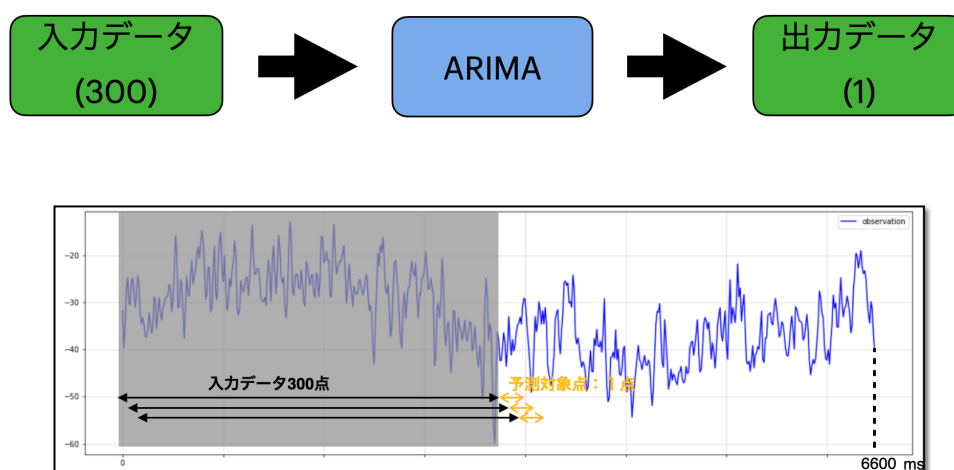


図6.11

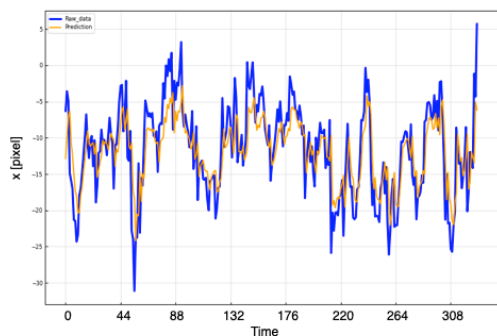


図6.12 Sirius1;x 座標シミュレーション

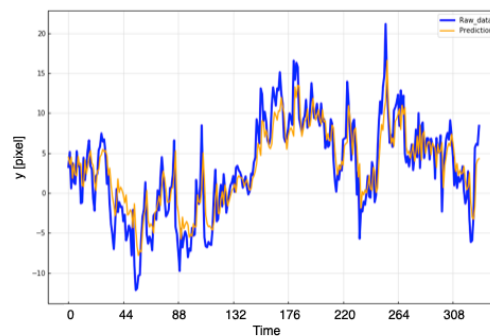


図6.13 Sirius1;y 座標シミュレーション

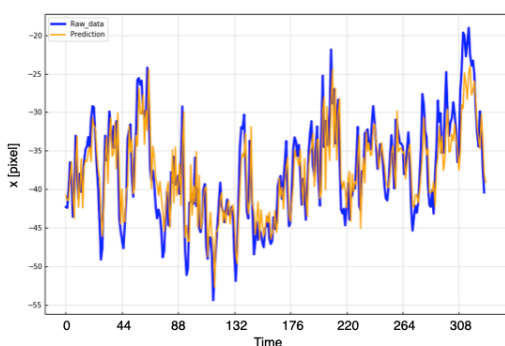


図6.14 Sirius2;x 座標シミュレーション

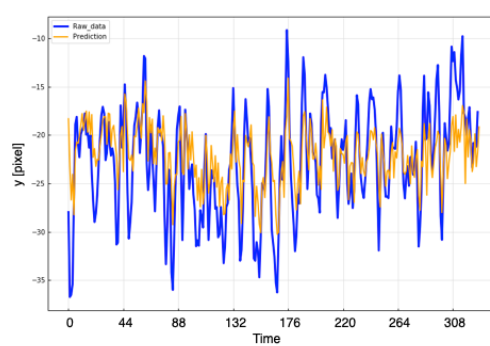


図6.15 Sirius2;y 座標シミュレーション

大気ゆらぎの時系列データに対して ARIMA モデルフィッティングを行い、ゆらぎの将来を予測した結果を図6.12～図6.15に示す。これらの図から、ゆらぎが急激に変化しスパイクのようになっている箇所では正確に予測しきれていないが、大まかなトレンドは追うことができている。

データの将来を予測するにあたって、予測精度がもっとも良くなるモデルを事前に知ることは不可能である。本研究では、AIC が最小になるようにモデルの次数を決定した。この次数決定のプロセスを省き、回帰係数をあらかじめ適当に決定することで、計算時間を短くすることができる。例えば、AIC が最小になるような次数をグリッドサーチした際、 $(p,d,q)=(3,0,2),(1,0,1)$ …のような結果が得られたとしよう。この場合、既存のデータに対し適切なモデルの一つとして $(p,d,q)=(3,0,2)$ の ARIMA モデルであるが、フィッティングの計算時間、予測の正確性を考慮した結果 $(p,d,q)=(1,0,1)$ が適切な場合もある。補償光学に実装するという観点において、モデルの構築方法に関して AIC を用いる必要性はないことから、予測精度とモデルの次数において更なる解析を行う必要がある。

6.3 過渡応答

図6.16は、Tip-Tilt Mirror にステップ応答を入力し、その過渡応答を調べたものである。これに、式4.4を最小二乗法を用いてフィッティング (gnuplot fit) し、各係数を求めた。

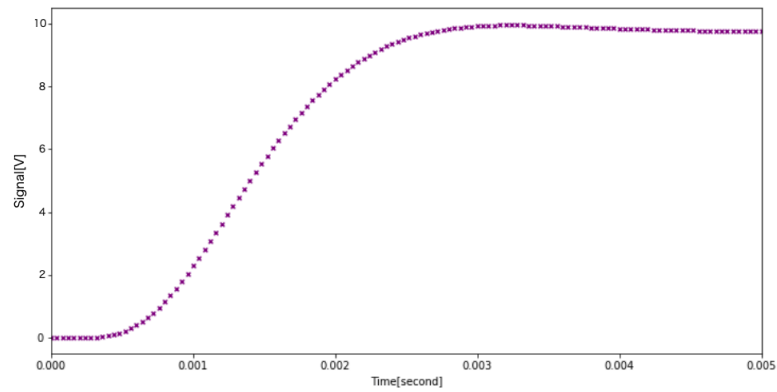


図6.16 Tip-Tilt Mirror の過渡応答

- $K = 9.74673$
- $\zeta = 0.704116$
- $\omega_n = 1259.56[\text{rad/s}]$

この結果から、複素領域上での伝達関数が次のように求められた。

$$G(s) = \frac{1.546 \times 10^6}{s^2 + 1774 \times s + 1.586 \times 10^6} \quad (6.1)$$

6.4 MATLAB を用いたシミュレーション

まず、従来の AO システム系で MATLAB を用いたシミュレーションの結果を示す (図6.17、図6.18、図6.19、図6.20)。紫色が大気ゆらぎ、赤色が Tip-Tilt Mirror の駆動、水色が補正後の明るさ重心の残差である。縦軸が pixel 値、横軸は時間 [ms] である。大気ゆらぎのデータが pixel 値でいう所の 0 付近からずれているのは望遠鏡の追尾機能が原因であると考えられる。

いずれのシミュレーションでも、水色のゆらぎが 0 付近に集まっていることが確認できる。ゆらぎが大きく変動している箇所では残差が大きく振れているが、これは従来の AO システム系が 1 時点前の値をフィードバックしていることが原因ではないかと考えている。

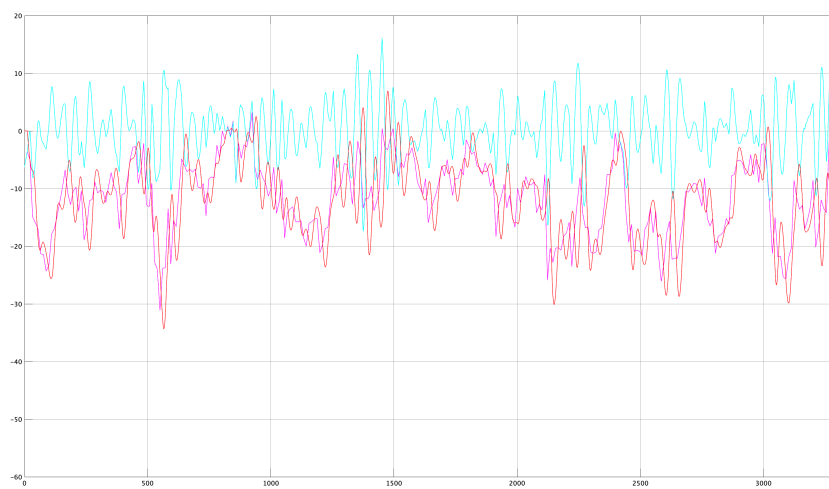


図6.17 Sirius1,x 座標

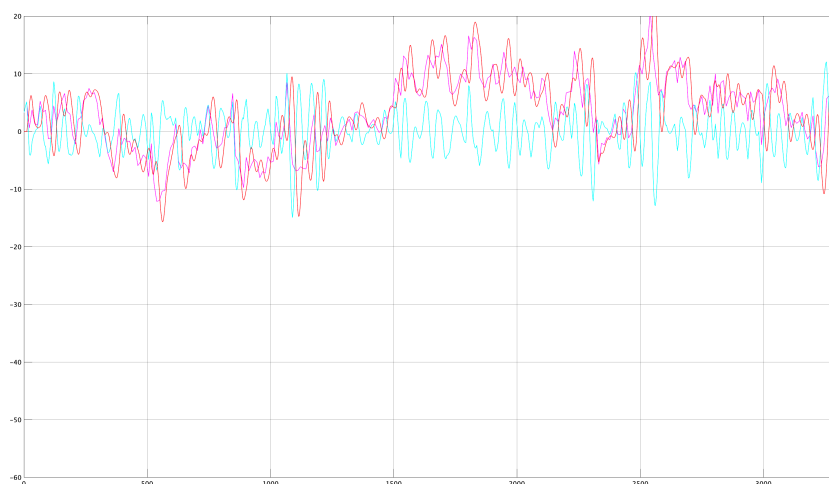


図6.18 Sirius1,y 座標

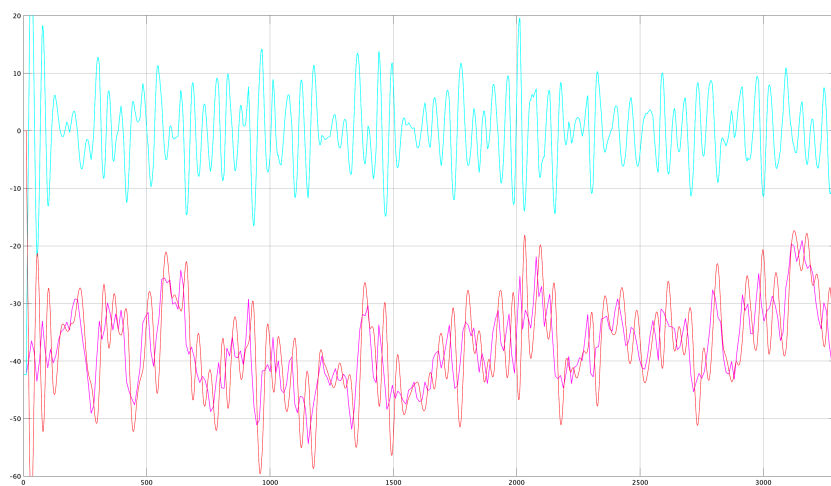


图6.19 Sirius2,x 座標

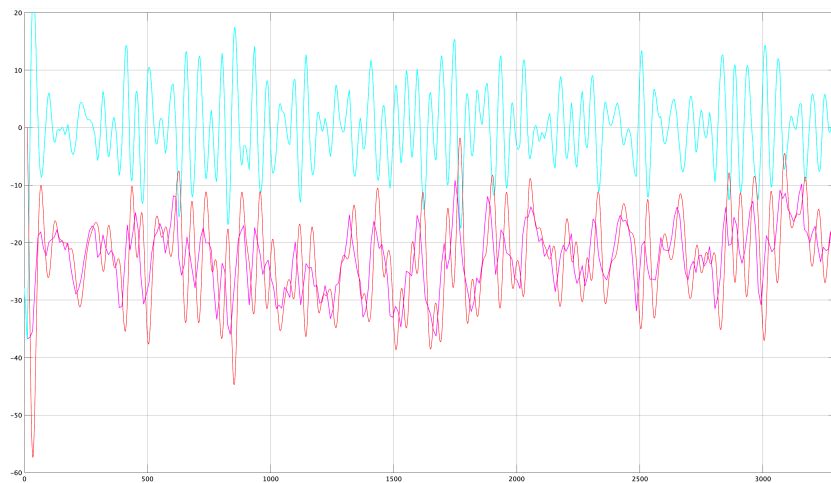


图6.20 Sirius2,y 座標

続いて、新手法を用いて行ったシミュレーションの結果を示す(図6.21、図6.22、図6.23、図6.24)。ただし、図の緑線は Tip-Tilt Mirror に送られた信号を表している。

ここで、従来と同様、補正結果が0付近に集まっており、さらに、従来の AO 補正法よりも、変動が抑えられているのが確認できた。

しかし新手法は開ループ制御光学系を想定しており、観測時間が進むにつれ残差が蓄積し、安定した制御ができない可能性もある。将来的な実装を考えるにあたり、本研究で想定した光学系に限らず、閉ループ制御と組み合わせたような光学系を考案すべきである。

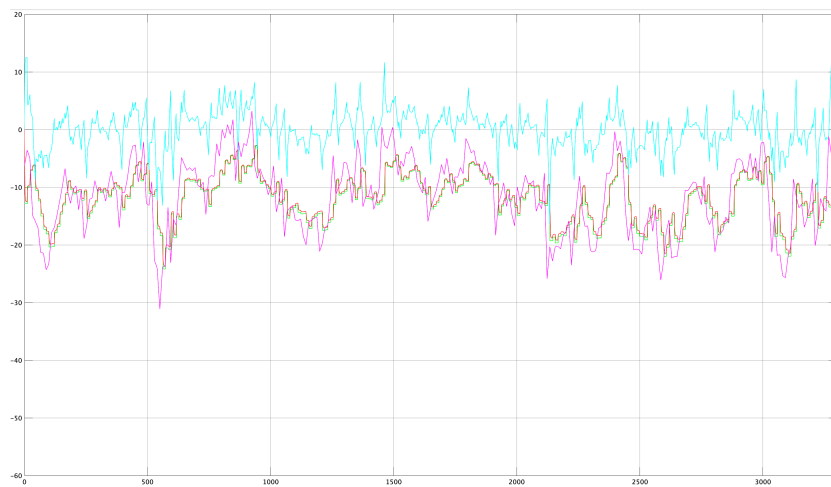


図6.21 Sirius1,x 座標

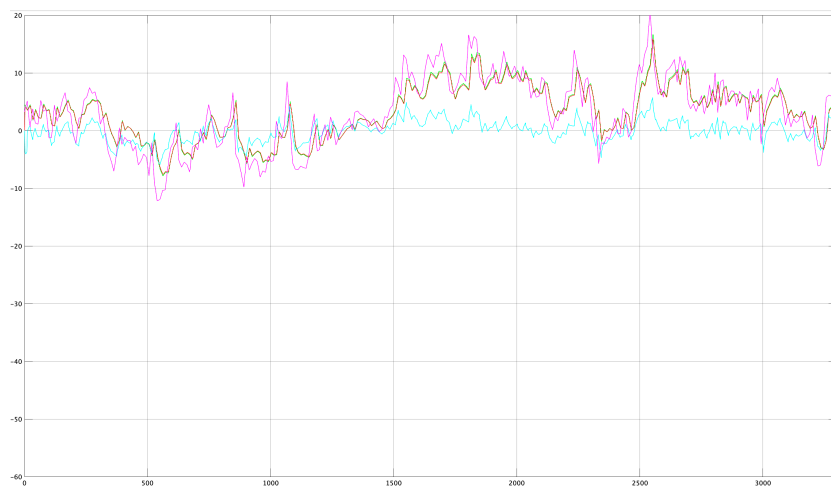


図6.22 Sirius1,y 座標

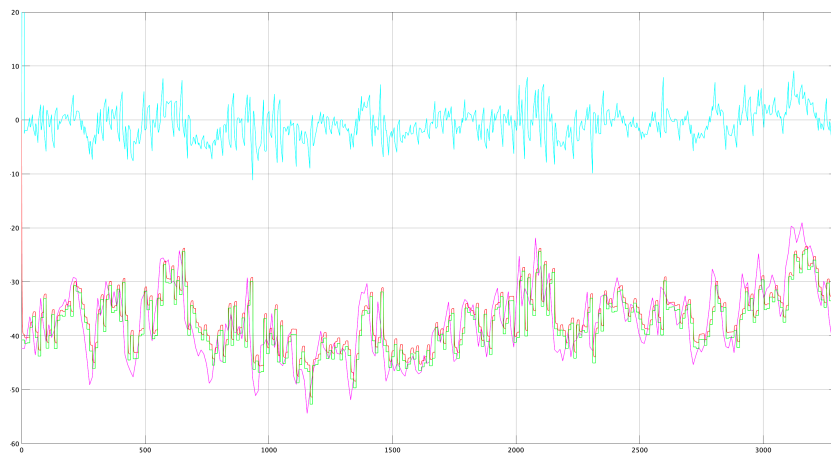


图6.23 Sirius2,x 座標

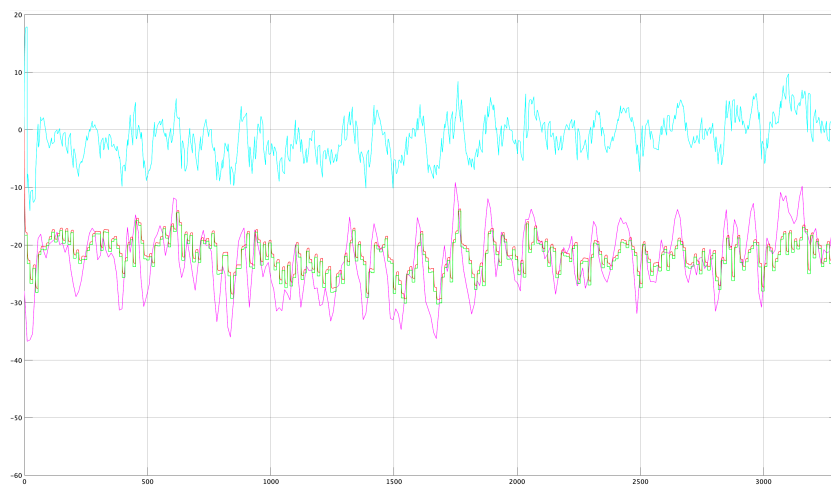


图6.24 Sirius2,y 座標

6.5 周波数解析

続いて、得られた結果をフーリエ変換し、どの周波数帯でどれくらい補正が効いているのかを評価した（図6.25、図6.26、図6.27、図6.28）

青線が大气ゆらぎ (raw_data)、黄線が従来の AO モデル (ao)、赤色が今回試した新手法 (arima) の結果である。

周波数が低い領域では従来モデルでも新手法でも補正ができています。しかし、高周波数側では新手法において補正精度が悪くなっている。これは、ゆらぎ予測において、低周波数の変動は容易で、高周波数の変動はノイズの成分が支配的になるため予測精度が悪いことを示唆している。

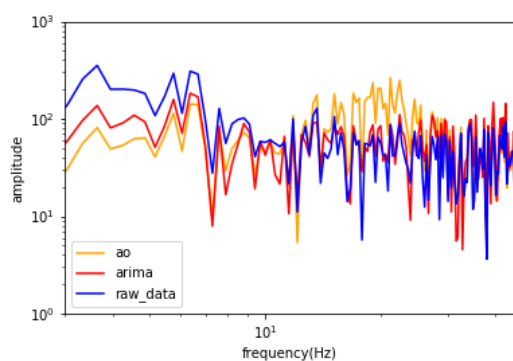


図6.25 Sirius1,x 座標

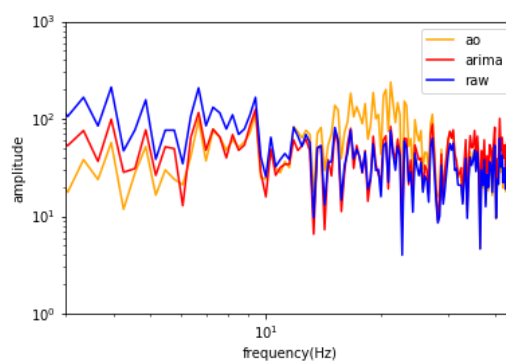


図6.26 Sirius1,y 座標

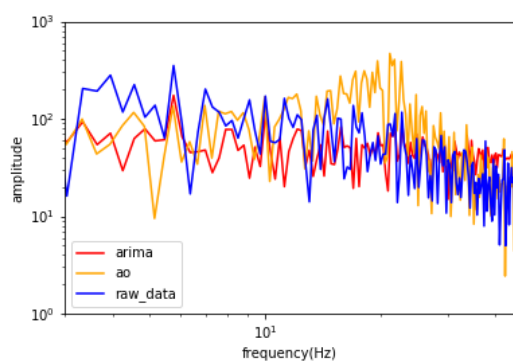


図6.27 Sirius2,x 座標

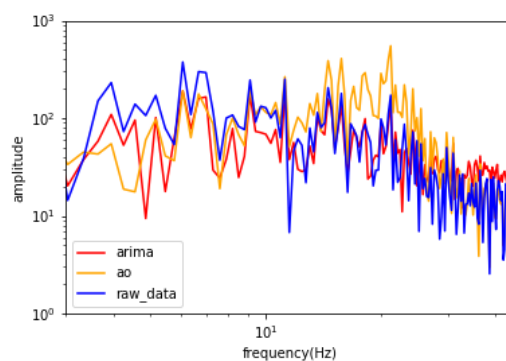


図6.28 Sirius2,y 座標

第 7 章

議論

7.1 評価

7.1.1 分散

今回の新手法と、従来の AO のシミュレーションとを評価する際に、pixel 値の分散を計算し、補正精度を評価した。それを下の表にまとめる。

- $0.19''/\text{pixel}$
- $\text{seeing} \sim 2.35\sigma$

分散の結果から、今回試案した大気ゆらぎ補正法は、従来の AO と比較して、分散において良い改善が見られた。また使用したデータは時刻のみが異なっている。直感的には気候状況、風速などが予測の精度に影響を与えると考えられ、これが新手法の精度にどのように影響が出るか調べる必要がある。

表7.1 Sirius1,x 座標

	分散 [pixel^2]	seeing[arcsec]
従来	26.129	2.282
新手法	17.502	1.868

表7.2 Sirius1,y 座標

	分散 [pixel^2]	seeing[arcsec]
従来	15.795	1.774
新手法	9.547	1.380

表7.3 Sirius2,x 座標

	分散 [$pixel^2$]	seeing[arcsec]
従来	52.988	3.250
新手法	10.300	1.433

表7.4 Sirius2,y 座標

	分散 [$pixel^2$]	seeing[arcsec]
従来	49.95	3.250
新手法	14.638	1.708

7.1.2 パラメータの推定

本研究では、与えられた時系列データをフィッティングする際に、AIC を用いた。統計学的には、あくまで一つの手法であり、さまざまな情報量基準が提案されている。例えば、モデルにフィットすることを最優先にしてパラメータを推定することを考えると、必然的に次数は大きくなり、1 時点先を予測した値はより平均に近づくように決定される。

また今回用いたデータは、データの時間間隔が 11ms であり、原理的にこれ以上の高周波数成分を検出できない。より高精度な補正を行うには、さらに細かく時間幅を取らなければならないが、本研究ではパラメータ決定にかかる計算時間は 11ms よりも十分短いと仮定しているが、この部分との兼ね合いも考慮すべきである。

7.1.3 VAR モデル

ベクトル自己回帰モデル (VAR; vector autoregressive model) は AR モデルをベクトルに一般化したものである。ただし、AR モデルは自分自身のラグのみを含んで推定するのに対し、VAR モデルはモデル内の他の変数のラグも含んで推定する。AR モデルと同様に VAR モデルの推定/検定にあたっては、データが定常であることが前提となる。

VAR(p) モデルは \mathbf{y}_t を定数ベクトル \mathbf{c} と自身の過去 p 時点の値 $\mathbf{y}_t, \dots, \mathbf{y}_{t-p}$ に回帰したモデルであり、以下で表される。

$$\mathbf{y}_t = \mathbf{c} + \sum_{i=1}^p \Phi_i \mathbf{y}_{t-i} + \epsilon_t, \epsilon_t \sim W.N.(\sigma^2) \quad (7.1)$$

ここで、 \mathbf{c} は $n \times 1$ の定数ベクトルであり、 Φ_i は $n \times n$ の係数行列であり、 Γ_k は $n \times n$ の分散共分散行列である。VAR モデルの期待値は、

$$\mu = E[\mathbf{y}_t] = \left(\mathbf{I}_n - \sum_{i=1}^p \Phi_i \right)^{-1} \mathbf{c} \quad (7.2)$$

であり、自己共分散 Γ_k は

$$\Gamma_k = \sum_{i=1}^p \Phi_i \Gamma_{k-i} \quad (7.3)$$

で求めることができる。

7.1.4 グレンジャー因果

データだけから、時系列データに対する因果性の有無を判断できる概念がグレンジャー因果 (Granger Causality) である。別の言い方をすると、グレンジャー因果はある変数 (群) が他の変数 (群) の予測の向上に役立つかどうかを判定する手法である。

グレンジャー因果を次のように定義する。現在と過去の x_t の値飲みに基づいた将来の予測と、現在と過去の x_t と y_t の値に基づいた将来の x の予測を比較して、後者の MSE (平均二乗誤差; mean squared error) の方が小さくなる場合、 y から x へのグレンジャー因果が存在する、とする。

例えば、2 変量 VAR(2) モデルを用いてグレンジャー因果性の説明する。2 変量 VAR(2) モデルを以下のように表す。

$$y_{1t} = c_1 + \phi_{11}^{(1)} y_{1,t-1} + \phi_{12}^{(1)} y_{2,t-1} + \phi_{11}^{(2)} y_{1,t-2} + \phi_{12}^{(2)} y_{2,t-2} + \epsilon_{1t} \quad (7.4)$$

$$y_{2t} = c_2 + \phi_{21}^{(1)} y_{1,t-1} + \phi_{22}^{(1)} y_{2,t-1} + \phi_{21}^{(2)} y_{1,t-2} + \phi_{22}^{(2)} y_{2,t-2} + \epsilon_{2t} \quad (7.5)$$

y_{2t} から y_{1t} へのグレンジャー因果性が存在しないということは、 $\phi_{12}^{(1)} = \phi_{12}^{(2)} = 0$ ということと同値になる。また、一般的に、 y_{2t} から y_{1t} へのグレンジャー因果性が存在しないということは、VAR の y_1 の式において y_2 に関連する係数が全て 0 になることと同値である。したがって、VAR の枠組みでは、F 検定を用いてグレンジャー因果性を検定することができる。2 変量 VAR(2) モデルの場合の具体的な手順は次のようになる。グレンジャー因果性を検定するには、帰無仮説 $H_0 : \phi_{12}^{(1)} = \phi_{12}^{(2)} = 0$ を検定すればよい。式7.4を推定し、その残差平方和を SSR_1 とする。次に、制約 $\phi_{12}^{(1)} = \phi_{12}^{(2)} = 0$ を課したモデル

$$y_{1t} = c_1 + \phi_{11}^{(1)} y_{1,t-1} + \phi_{11}^{(2)} y_{1,t-2} + \epsilon_{1t} \quad (7.6)$$

を推定し、その残差平方和を SSR_0 とする。このとき、F 検定量は

$$F \equiv \frac{(SSR_0 - SSR_1)/2}{SSR_1/(T-5)} \quad (7.7)$$

で定義される。ここで、 T は標本数である。 $2F$ は漸近的に $\chi^2(2)$ に従うことが知られている。したがって、 $2F$ の値を $\chi^2(2)$ の 95% 点と比較して $2F$ の方が大きければ y_{2t} から y_{1t} へのグレンジャー因果性が存在しないという帰無仮説を棄却し、 y_{2t} は y_{1t} の将来の予測に有用であると判断できる。より一般的な n 変量 VAR(p) モデルにおいて、ある変数 (群) から y_{kt} へのグレンジャー因果性が存在するかどうかの検定の手順を以下にまとめる。

- VAR モデルにおける y_{kt} のモデルを推定し、その残差平方和を SSR_0 とする
- VAR モデルにおける y_{kt} のモデルに制約を課したモデルを推定し、その残差平方和を SSR_1 とする
- F 検定量を以下で計算する

$$F \equiv \frac{(SSR_0 - SSR_1)/r}{SSR_1/(T - np - 1)} \quad (7.8)$$

ここで、 r はグレンジャー因果性検定に必要な制約の数である

- rF を $\chi^2(r)$ の 95% 点と比較して $2F$ の方が大きければ y_{2t} から y_{1t} へのグレンジャー因果性が存在し、小さければグレンジャー因果性は存在しないと結論づける。

また、グレンジャー因果性の性質として

- 因果性が存在する必要条件であるが、十分条件でない
- 真の因果性の方向とグレンジャー因果性の方向が一致するとは限らない
- 定性的概念であり、関係の強さが測れない

があげられる。因果関係の方向がはっきりしている場合を除いては、ある変数が他の変数の予測に有用かどうかという観点で解釈するのが良い。

本研究で用いたデータについてグレンジャー因果性の存在を検定し、まとめたのが以下の表になる。

天体	向き	因果性
Sirius1	x→y	なし
	y→x	なし
Sirius2	x→y	あり
	y→x	あり

結果から、グレンジャー因果性は二つのデータに対して、正反対の結果となった。Sirius1 のデータに対し VAR モデルを適用しても、予測の改善はみられないということになる。しかし、データの時間感覚を短くすることや、SHWFS の各波面のデータのような、時間的空間的な距離が近いデータの場合にどのようなようになるか、詳しい解析と議論が必要である。

7.2 機械学習

近年注目されている機械学習や深層学習（ディープラーニング）のアルゴリズムを用いた未来予測の手法に再帰型ニューラルネットワーク（Recurrent Neural Network：以下、RNN）と呼ばれるものがある。

ニューラルネットワークとは、人間の脳内にある神経細胞（ニューロン）とそのつながり、つまり神経回路網を数式的なモデルで表現したものである。RNNは、ニューラルネットワークを拡張して時系列データを扱えるようにしたものである。ARIMAモデルと共通する考え方として、ある時刻における値は、以前の時刻の値の変化の延長上として表現できるということである。

ARIMAモデルとRNNの最大の違いは、非線形なモデルも組み込める点にある。大気ゆらぎは、様々なランダム成分が要素としてあるが、ある乱流モデルにしたがって空間的、時間的に変化することが知られている。コレログラムからも分かる通り、近い時点との相関が見られることから、RNNを用いれば、何らかの形で大気ゆらぎを物理的性質を学習し、その性質から大気ゆらぎの将来をより正確に予測できるのではないかと考えている。

第 8 章

結論

近年、超大型望遠鏡の建設により、天文学における補償光学の重要性は増しつつある。本研究は、Babcock が唱えた補償光学システムに対し、その構成の根本的な変更を提案したものである。また、我々の生活に身近な天気予報は、大気ゆらぎのマクロな予測である。本研究では、大気ゆらぎデータに相関を見出し、経済・マーケティングの分野で多く用いられる ARIMA モデルに代表される自己回帰型のデータ解析手法を用いて大気ゆらぎの予測を行った。

従来の AO に対して、新手法ではゆらぎの分散を減少させることができた。これが波面補正にも応用できるかは議論の余地があるが、Tip-Tilt 成分の補正のみでも、波面全体での補正精度の向上に重要である。今後の補償光学の分野の更なる発展のために、本研究のような時系列データ解析を用いた補償光学の新たなモデルを提案したい。

謝辞

まずはじめに、本論文の作成にあたり丁寧に指導して下さった秋山正幸教授に感謝申し上げます。お忙しい中私の訪問に嫌な顔一つせず、議論に付き合ってくださいましたこと、御礼申し上げます。

また、秋山研究室のメンバーをはじめ天文学専攻のみなさまには観測ゼミやミーティングで大変お世話になりました。そこでの議論や質問によって、私自身の理解が深まったところや、分かっていた所がなかった所に気づくことができました。

大学院生活は、研究面でも生活面でも未熟でした。思うように研究が進まず、精神的に辛い場面もありました。さらに、春ごろには日本で COVID-19 が発生し、就職活動にもアクシデントが生じ、将来への不安から時にも、先輩や同期の方々をはじめとする多くの方々が支えてくださいました。

天文学の中で、装置系の分野はマジョリティではないのかもしれませんが、私は補償光学という分野を通して、この分野だからこそさまざまな経験を積むことができたと思います。

改めて、秋山研究室のメンバーをはじめ、天文学専攻の研究室の先輩、研究員の方や辛いときに励ましてくれた同期のみなさまに、重ねてお礼申し上げます。

付録 A

付録

A.1 コレログラム

自己相関とは、時系列データに対し、そのデータをあるタイムステップ数ずらしたデータとの相関のことをいう。ずらしたタイムステップ数をラグという。自己相関は過去の履歴に対してどの程度の相関を持つかを示している。時系列データ \mathbf{U} が

$$\mathbf{U} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_t \end{pmatrix} \quad (\text{A.1})$$

として得られたとき、 h 次の自己相関係数 r_k (k だけタイムステップがずれたデータとの自己相関係数) は、 \bar{u} をデータの平均として次で与えられる。

$$r_k = \frac{\sum_{t=k+1}^T (u_t - \bar{u})(u_{t-k} - \bar{u})}{\sum_{t=1}^T (u_t - \bar{u})^2} \quad (\text{A.2})$$

ラグ k との自己相関係数の推移を確認するために、横軸にラグ k をとり、縦軸に自己相関係数の値をプロットしたものをコレログラムという。

A.2 統計的モデリング

A.2.1 確率分布と統計的モデル

Y が確率変数 (random variable) のとき、全ての実数 $y \in R$ に対して $Y \leq y$ となる確率 $Prob(Y \leq y)$ が定められる。これを y の関数とみなして

$$G(y) = Prob(Y \leq y) \quad (\text{A.3})$$

と表すとき、 $G(y)$ を確率変数 Y の分布関数 (probability distribution function) と呼ぶ。時系列解析で用いられる確率変数は通常、連続型と呼ばれるもので、その分布関数は $g(t) \leq 0$ ($-\infty < t < \infty$) を満たす関数の積分によって

$$G(y) = \int_{-\infty}^y g(t) dt \quad (\text{A.4})$$

と表現できる。このとき $g(x)$ を密度関数 (density, probability density function) という。代表的な密度関数として、正規分布 (ガウス分布)、指数分布、 χ^2 分布などがある。

統計解析では、データ y_1, \dots, y_N が得られたとき、これをある確率変数 Y の実現値とみなす。確率変数を特徴付ける密度関数 $g(y)$ を真のモデル (true model) と呼ぶ。

通常真のモデルは未知であるので、与えられたデータからそのデータを生成する密度関数を推定することになる。データから推定された密度関数は統計的モデルと呼ばれ、 $f(y)$ と表される。

A.2.2 KL 情報量とエントロピー最大化原理

現実のデータを生成する真のモデルを $g(y)$ 、それを近似した統計的モデルを $f(y)$ と表すことにする。統計的モデリングでは $g(y)$ になるべく「近い」 $f(y)$ を求めることが主要な目的になるが、そのためにはモデル $f(y)$ の良さを客観的に評価するための基準が必要となる。その基準として、KL 情報量 (Kullback-Leibler information)

$$I(g; f) = E_Y \log \left\{ \frac{g(Y)}{f(Y)} \right\} = \int_{-\infty}^{\infty} \log \left\{ \frac{g(y)}{f(y)} \right\} g(y) dy \quad (\text{A.5})$$

を用いる。この KL 情報量は

1. $I(g; f) \leq 0$
2. $I(g; f) = 0 \Leftrightarrow g(y) = f(y)$

という性質を持っている。また、KL 情報量の符号を反転した量 $B(g; f) = -I(g; f)$ は一般化されたエントロピーと呼ばれ、想定した分布 $f(y)$ から n 個の実現値をとったとき、その相対度数分布が真の分布 $g(y)$ と等しくなる確率の $\frac{1}{n}$ を近似的に与える。従って、KL 情報量が小さいほど確率

分布 f は g に近いと考えることができる。統計的モデルにおいて $B(g; f) = -I(g; f)$ を最大とするようにモデルを構築しようとするのがエントロピー最大化原理である。

A.2.3 KL 情報量の推定と対数尤度

統計解析が行われる場面では、真の分布は未知であり KL 情報量を計算できないため、KL 情報量が実際の統計的モデルの評価に使われることはほとんどない。 $g(y)$ から観測されたデータ y_1, \dots, y_N が与えられている場合を想定し。モデル $f(y)$ の KL 情報量を推定する鳳凰を考える。

エントロピー最大化原理に従ってもっとも良いモデルを求めるためには $B(g; f)$ を最大、すなわち $I(g; f)$ を最小にするモデルを求めればよい。KL 情報量は

$$I(g; f) = E_Y \log g(Y) - E_Y \log f(Y) \quad (\text{A.6})$$

と二つの項に分解できるが、第 1 項はモデル $f(y)$ には依存しない一定の値を取るため、第 2 項を最大にするモデルが良いモデルになる。この右辺第 2 項は平均対数尤度と呼ばれる量で、密度関数をもつ連続型のモデルの場合には

$$E_Y \log f(Y) = \int \log f(y) g(y) dy \quad (\text{A.7})$$

と書ける。

しかし、この平均対数尤度も $g(y)$ が未知の場合には直接計算できないが、データ y_n が密度関数 $g(t)$ に従って生成されることから、大数の法則よりデータ数が $N \rightarrow \infty$ のとき

$$\frac{1}{N} \sum_{n=1}^N \log f(y_n) \rightarrow E_Y \log f(y) \quad (\text{A.8})$$

が成り立つ。従って、本来の評価基準である KL 情報量の代わりに対数尤度 $\sum_{n=1}^N \log f(y_n)$ を最大にするようなモデルを選べば、近似的にエントロピーを最大にすることができる。式 A.8 の左辺の N 倍は観測値が独立に得られた場合の対数尤度 (log-likelihood) と呼ばれ、

$$l = \sum_{n=1}^N \log f(y_n) \quad (\text{A.9})$$

と表される。また、その指数をとった

$$L = \prod_{n=1}^N f(y_n) \quad (\text{A.10})$$

は尤度と呼ばれる。

時系列解析のモデルでは通常、観測値が独立に得られるという仮定は成り立たない。式 A.9、A.10 を一般的に表す場合、 y_1, \dots, y_N の同時分布を用いて

$$L = f(y_1, \dots, y_N) \quad (\text{A.11})$$

$$l = \log L = \log f(y_1, \dots, y_N) \quad (\text{A.12})$$

と書ける。

A.2.4 最尤法によるパラメータ推定

モデルが θ をパラメータとするパラメトリックモデルで $f(y) = f(y|\theta)$ の形をしている場合には対数尤度 l はパラメータ θ の関数と考えることができる。 θ を明示的に示し、

$$l(\theta) = \begin{cases} \sum_{n=1}^N \log f(y_n|\theta) & \text{独立の場合} \\ \log f(y_1, \dots, y_N|\theta) & \text{一般の場合} \end{cases} \quad (\text{A.13})$$

を θ の対数尤度関数と呼ぶ。

対数尤度関数 $l(\theta)$ は θ で表すモデルのよさを評価した量なので、 $l(\theta)$ を最大にする θ を選ぶことによってパラメトリックモデル $f(y) = f(y|\theta)$ の最適な値を求めることができる。このように尤度または対数尤度を最大化することによりパラメータを推定する方法を最尤法 (maximum likelihood method) という。また、最尤法で推定されたパラメータを $\hat{\theta}$ と表し、最尤推定値と呼ぶ。

A.2.5 状態空間モデル

y を l 変量の時系列とする。この時系列を表現するようなモデルを状態空間モデル (state-space model) という。

$$x_n = F_n x_{n-1} + G_n v_n \quad (\text{システムモデル}) \quad (\text{A.14})$$

$$y_n = H_n x_n + w_n \quad (\text{観測モデル}) \quad (\text{A.15})$$

ここで、 x_n は直接観測できない k 次元ベクトルで、状態と呼ばれる。 v_n はシステムノイズあるいは状態ノイズと呼ばれ、平均ベクトル 0、分散共分散行列 Q_n に従う m 次元のホワイトノイズである。一方、 w_n は観測ノイズで、分散共分散行列 R_n に従うホワイトノイズ、 F_n 、 G_n 、 H_n はそれぞれ $k \times k$ 、 $k \times m$ 、 $l \times k$ の行列である。時系列解析で用いられる線形モデルの多くは、この状態空間モデルで表現される。

ここで、簡単な時系列モデルに対応する状態空間モデルの例を示す。時系列 y_n は AR モデル

$$y_n = \sum_{i=1}^m \phi_i y_{n-i} + \epsilon_t, \quad v_n \sim W.N.(\sigma^2) \quad (\text{A.16})$$

にしたがっているものとする。このとき、状態ベクトルを $x_n = (y_n, y_{n-1}, \dots, y_{n-m+1})^T$ と定義すると、 x_n と x_{n-1} の間には式 A.14 が成り立つことが確かめることができる。ただし、 F と G はそ

それぞれ

$$F = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_m \\ 1 & & & \\ & \ddots & & \\ & & 1 & 0 \end{pmatrix}, \quad G = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (\text{A.17})$$

によって定まる $m \times m$ 、 m 次元ベクトルである。

A.2.6 カルマンフィルタ

状態空間モデルに関連した重要な問題は、時系列 y_n の観測地に基づいて状態 x_n の推定を行うことである。以下では観測値 $Y_j = \{y_1, \dots, y_j\}$ に基づいて時刻 n における状態 x_n の推定を行うことを考える。特に、 $j < n$ の場合は観測区間より先の将来の状態を推定する問題で、予測 (prediction, forecast) と呼ばれる。

この状態推定を行うためには、観測値 Y_j が与えられた下での状態 x_n の条件付き分布 $p(x_n|Y_j)$ を求めればよい。ところが式A.14、A.15の状態空間モデルは線形モデルであり、しかも v_n 、 w_n および x_0 がすべて正規分布に従うので、これらの条件付き分布は正規分布になる。したがって、状態空間モデルの状態推定問題を考えるためには、条件付き分布を規定する平均ベクトルと分散共分散行列だけを求めればよい。一般に、観測値 y_1, \dots, y_j が与えられた下での状態 x_1, \dots, x_n の条件付き同時分布を求めるためには莫大な計算量を要するが、状態空間モデルに対しては、逐次的な計算アルゴリズムによって状態 x_n の条件付き周辺分布をきわめて効率的に計算できる。これがカルマンフィルタ (Kalman filter) と呼ばれるアルゴリズムである。

以下では、状態 x_n の条件付き平均と分散共分散行列を

$$x_{n|j} \equiv E(x_n|Y_j) \quad (\text{A.18})$$

$$V_{n|j} \equiv E[(x_n - x_{n|j})(x_n - x_{n|j})^T] \quad (\text{A.19})$$

と表す。ただし、カルマンフィルタで直接取り扱うのは、 $j = n-1$ の場合 (1 時点予測) と $j = n$ の場合 (フィルタ) である。以下のアルゴリズムに従って 1 時点先予測とフィルタを交互に繰り返すことにより、これらを順次求めることができる。

・ 1 時点予測

$$x_{n|n-1} = F_n x_{n-1|n-1} \quad (\text{A.20})$$

$$V_{n|n-1} = F_n V_{n-1|n-1} F_n^T + G_n Q_n G_n^T \quad (\text{A.21})$$

・ フィルタ

$$K_n = V_{n|n-1} H_n^T (H_n V_{n|n-1} H_n^T + R_n)^{-1} \quad (\text{A.22})$$

$$x_{n|n} = x_{n|n-1} + K_n (y_n - H_n x_{n|n-1}) \quad (\text{A.23})$$

$$V_{n|n} = V_{n|n-1} - K_n H_n V_{n|n-1} \quad (\text{A.24})$$

1 時点先予測のアルゴリズムでは、 x_n の予測値 (平均) ベクトル $x_{n|n-1}$ は $x_{n-1|n-1}$ に推移行列 F_n をかければよい。また、その分散共分散行列 $V_{n|n-1}$ は 2 項からなり、第 1 項は F_n による変換の影響、第 2 項はシステムノイズ v_n の影響を表す。

フィルタリングのアルゴリズムではまずカルマンゲインと呼ばれる K_n が求められる。また $y_n - H_n x_{n|n-1}$ は予測誤差、 $H_n V_{n|n-1} H_n^T + R_n$ はその分散共分散行列である。このとき、 x_n のフィルタの平均ベクトルは予測ベクトル $x_{n|n-1}$ と予測誤差にカルマンゲインをかけたものとして求められる。 $V_{n|n}$ に関しては、右辺第 2 項は観測値からの情報によって x_n の状態推定の精度が改善された部分を表している。

A.2.7 長期予測

カルマンフィルタのアルゴリズムを繰り返し用いることにより、長期の予測を行うことを考える。つまり、 $Y_n = \{y_1, \dots, y_n\}$ に基づいて j 点先の状態 x_{n+j} ($j > 1$) を推定する。

まず、カルマンフィルタにより x_{n+1} の 1 時点先予測の平均 $x_{n+1|n}$ および分散共分散行列 $V_{n+1|n}$ が求められる。ここで観測値 y_n は得られないので、形式的に $Y_{n+1} = Y_n$ が成り立つものとみなして計算を続けられればよいことがわかる。この場合、 $x_{n+1|n+1} = x_{n+1|n}$, $V_{n+1|n+1} = V_{n+1|n}$ が成り立つ。したがってカルマンフィルタの $n+1$ 時点に対する 1 時点先予測のアルゴリズムから

$$x_{n+2|n} = F_{n+2} x_{n+1|n} \quad (\text{A.25})$$

$$V_{n+2|n} = F_{n+2} V_{n+1|n} F_{n+2}^T + G_{n+2} Q_{n+2} G_{n+2}^T \quad (\text{A.26})$$

が得られる。これは 2 時点先予測のためには y_{n+1} に対するカルマンフィルタのアルゴリズムを省略し、予測ステップ飲みを実行すればよいことを示している。

一般に、 Y_n に基づいて j 時点までの長期予測は、 $Y_n = Y_{n+1} = \dots = Y_{n+j}$ が成り立つことから、予測ステップを k 回繰り返せばよいことがわかる。

・長期予測

$i = 1, \dots, j$ について

$$x_{n+i|n} = F_{n+i} x_{n+i-1|n} \quad (\text{A.27})$$

$$V_{n+i|n} = F_{n+i} V_{n+i-1|n} F_{n+i}^T + G_{n+i} Q_{n+i} G_{n+i}^T \quad (\text{A.28})$$

以上の方法で求めた状態 x_n の予測を用いて、時系列の予測が直ちに実現できる。すなわち、状態 x_n と時系列 y_n の関係が観測モデル A.15 により与えられているので、 Y_n が与えられたときの y_{n+j} の平均と分散共分散行列をそれぞれ $y_{n+j|n} \equiv E(y_{n+j}|Y_n)$, $d_{n+j|n} \equiv \text{Cov}(y_{n+j}|Y_n)$ とすると

$$\begin{aligned} y_{n+j|n} &= E(H_n + j x_{n+j} + w_{n+j}|Y_n) \\ &= H_{n+j} x_{n+j|n} \end{aligned} \quad (\text{A.29})$$

$$\begin{aligned}
d_{n+j|n} &= Cov(H_{n+j}x_{n+j} + w_{n+j}|Y_n) \\
&= H_{n+j}Cov(x_{n+j}|Y_n)H_{n+j}^T + H_{n+j}Cov(x_{n+j}, w_{n+j}|Y_n) \\
&\quad + Cov(w_{n+j}, x_{n+j}|Y_n)H_{n+j}^T + Cov(y_{n+j}|Y_n) \\
&= H_{n+j}V_{n+j|n}H_{n+j}^T + R_{n+j}
\end{aligned} \tag{A.30}$$

が得られる。このように時系列の観測値 Y_n に基づく y_{n+j} の予測分布は平均 $y_{n+j|n}$ 、分散共分散行列 $d_{n+j|n}$ の正規分布となり、それらはA.29、A.30により求められる。このとき、 y_{n+j} の予測値は $y_{n+j|n}$ 、標準誤差は $(d_{n+j|n})^{1/2}$ で与えられる。

A.2.8 尤度の表現とパラメータ推定

パラメータ θ によって規定される時系列モデルがあり、その状態空間モデルが与えられているものとする。時系列データ $\{y_1, \dots, y_N\}$ が得られたときこの時系列モデルによって定まる y_1, \dots, y_N の N 次元同時密度関数を $f_N(y_1, \dots, y_N|\theta)$ とすると、モデルの尤度は

$$L(\theta) = f_N(y_1, \dots, y_N|\theta) \tag{A.31}$$

と定義される。ここで、 $n = N, N-1, \dots, 2$ について

$$f_n(y_1, \dots, y_n|\theta) = f_{n-1}(y_1, \dots, y_{n-1}|\theta)g_n(y_n|y_1, \dots, y_{n-1}, \theta) \tag{A.32}$$

が成り立つことを繰り返し適用すると、尤度は条件付き密度関数の積により

$$L(\theta) = \prod_{n=1}^N g_n(y_n|y_1, \dots, y_{n-1}, \theta) = \prod_{n=1}^N g_n(y_n|Y_{n-1}, \theta) \tag{A.33}$$

対数尤度は

$$l(\theta) = \log L(\theta) = \sum_{n=1}^N g_n(y_n|Y_{n-1}, \theta) \tag{A.34}$$

と書ける。ただし $g_n(y_n|Y_{n-1}, \theta)$ は観測値 Y_{n-1} が与えられたときの y_n の予測分布で、平均 $y_{n|n-1}$ 分散共分散行列 $d_{n|n-1}$ の正規分布となるので

$$g_n(y_n|Y_{n-1}, \theta) = \left(\frac{1}{\sqrt{2\pi}} \right)^l |d_{n|n-1}|^{-1/2} \exp \left\{ -\frac{1}{2}(y_n - y_{n|n-1})^T d_{n|n-1}^{-1} (y_n - y_{n|n-1}) \right\} \tag{A.35}$$

と表すことができる。したがって、これを式A.34に代入することでモデルの対数尤度は次のように求められる。

$$l(\theta) = -\frac{1}{2} \left\{ lN \log 2\pi + \sum_{n=1}^N \log |d_{n|n-1}| + \sum_{n=1}^N (y_n - y_{n|n-1})^T d_{n|n-1}^{-1} (y_n - y_{n|n-1}) \right\} \tag{A.36}$$

AR モデルや ARMA モデルは線形ガウス型の状態空間モデルの形で表現できる。このような時系列モデルの最尤推定値は式A.36を目的関数として、これを最大とするパラメータを求めればよい。

参考文献

- [1] H.W.Babcock:”The possibility of compensating astronomical seeing”, Publ. Aston.Soc.Pac. 65, 229-236(1953)
- [2] T Sean Ross. Limitations and applicability of the mar´echal approximation. Applied optics, 48(10):1812-1818, 2009.
- [3] Franois Rigaut ,Benoit Neichel.Multiconjugate Adaptive Optics for Astronomy
- [4] Franois Roddier. V the effects of atmospheric turbulence in optical astronomy. In Progress in optics, volume 19, pages 281-376. Elsevier, 1981.
- [5] John W Hardy. Adaptive optics for astronomical telescopes, volume 16. Oxford University Press on Demand, 1998.
- [6] Alexander Mikhailovich Obukhov. Structure of temperature field in turbulent flow. Technical report, AIR FORCE SYSTEMS COMMAND WRIGHT-PATTERSON AFB OH FOREIGN TECHNOLOGY DIVISION, 1970.
- [7] Andres Guesalaga, Benoit Neichel, Angela Cortes, Clementine Bechet and Dani Guzman.Using the C_N^2 and wind profiler method with wide-field laser-guide-stars adaptive optics to quantify the frozen-flow decay, 2014
- [8] シリーズ現代の天文学 15、日本評論社
- [9] Andrei Nikolaevitch Kolmogorov. Local turbulence structure in incompressible fluids at very high reynolds numbers. In Dokl. Akad. Nauk SSSR, volume 30, 1941.
- [10] Akaike H.Information theory and an extension of the maximum likelihood principle, in:

- B.N. Petrov and F. Csaki, eds., 2nd
Internat. Syrup. on Information Theory (Akademia Kiado, Budapest) pp. 267-281. (1973)
- [11] A 100-pc Scale, Fast and Dense Outflow in Narrow-Line Seyfert 1 Galaxy IRAS 04576+0912
- [12] Direct Imaging of Multiple Planets Orbiting the Star HR 8799(2008)
- [13] 加藤隆, 制御工学テキスト, 日本理工出版会 (1997)
- [14] Python による制御工学入門, 南裕樹, オーム社 (2019)
- [15] 島田直希, 時系列解析, 共立出版 (2019)
- [16] 北川源四郎, 時系列解析入門, 岩波書店 (2005)
- [17] Yoshito Ono, Masayuki Akiyama, Yosuke Minowa, Etsuko Mieda, Koki Tero, Hajime Ogane, Kaoru Oomoto, Yuta Iizuka, Shin Oya, Tomoyasu Yamamuro, Daiki Sakurai, and Kazuma Mitsuda, ULTIMATE-START: LTAO Experiment at Subaru
- [18] Richard Davies and Markus Kasper. Adaptive Optics for Astronomy, *Annu. Rev. Astron. Astrophys.* 2012. 50:305-51

