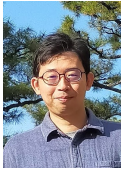


観測的宇宙論への機械学習の導入事例：エミュレーション技術とす ばる望遠鏡への応用



西道啓博

京都大学基礎物理学研究所
takahiro.nishimichi@yukawa.kyoto-
u.ac.jp

宇宙開闢後僅か 38 万年後の姿を捉えた、宇宙マイクロ背景放射 (Cosmic Microwave Background: CMB) は、観測可能な最大スケールにおける宇宙の姿を明らかにした。全天のあらゆる方向から届く、灼熱のビッグバン宇宙の黒体輻射の名残りは、到来方向に依存して $\mathcal{O}(10^{-5})$ の微小な温度の違い (揺らぎ) を示しており、その詳細な分析が **ΛCDM モデル** と呼ばれる現代の標準宇宙モデルの確立に決定的な役割を果たした。ところが、最近になって CMB と近傍宇宙の観測との矛盾が取り沙汰されている。**ハッブルテンション** と呼ばれるこの問題は、ΛCDM モデルの綻びを示しているかもしれない。

CMB 期から遙かな時を経て、微小な種揺らぎは重力により増幅され、やがて形作られる星、銀河、銀河団といった階層的な宇宙の大規模構造。宇宙初期の姿を 2 次元天球面上のスナップショットとして写し出した CMB に対して、大規模構造は複雑な非線形進化を経て作られた複雑なネットワーク状の 3 次元パターンであり、潜在的により多くの情報を有している。時間的にも距離スケールのにも CMB とは相補的な大規模構造は、ΛCDM モデルの妥当性をより厳密に検証する可能性を秘めている。

大規模構造の観測データから宇宙モデルやそこに内包される宇宙論パラメータを導くには、理論予言と観測データの比較に基づく統計推論が必要となる。様々なモデルとパラメータの組み合わせの中で、観測を最もよく再現するものは何かという問題である。これを高精度に行うには精巧な理論予言が必要となる。宇宙の構造形成シミュレーションは、計算コストの高さから長ら

く統計推論への応用が叶わなかった。近年飛躍的に進展したデータ科学的方法論は、文字通り天文学的に大容量のデータを用いて人類が実証可能な最大スケールの現象を解き明かそうという宇宙論においても有用である。上記の問題は、いわゆる「シミュレーションに基づく推論」の範疇にあり、宇宙論、物理学だけに留まらず気象科学、生態学、疫学、分子動力学、工学、経済学などの諸分野に共通する大きなテーマとなっている。

シミュレーションに基づく宇宙論を可能とすべく、我々は大規模シミュレーションデータベースの構築と、**エミュレータ**の開発を目的とした「ダーククエスト計画」を 2015 年より推進している。2018 年には初期のデータベースに基づく、ソフトウェア「ダークエミュレータ」の完成を見た。その後、模擬観測データを用いた種々のテストをクリアした後、このほどすばる望遠鏡が世界最高精度で測定した重力レンズ効果、およびスローン・デジタル・スカイ・サーベイが提供する現存する最大の銀河の 3 次元地図へと応用し、遂に宇宙論的帰結を導くことに成功した。計算コストの大きなシミュレータをエミュレータに置き換えることが、シミュレーションに基づく推論の具体的な実装例として機能することを実証した。

今後ますます増える観測データと、先鋭化するデータ科学的方法論の応用は、宇宙論の景色を大きく変えるかもしれない。すばるが、日本が、世界の宇宙論をリードするには、このような新しい手法が導き出した帰結を、どれだけ説得力を持った形で世に提示できるかが 1 つの鍵になるであろう。今後の展開からますます目が離せない。

—用語解説—

ΛCDM モデル：

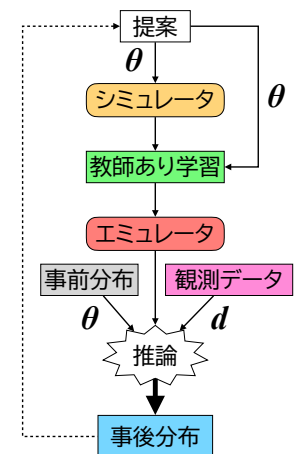
宇宙項 Λ と冷たいダークマター (Cold Dark Matter: CDM) を主成分とする「標準」宇宙モデル。それぞれ、宇宙の加速膨張、大規模構造形成や動力学を決める重力源としての役割を担う。

ハッブルテンション：

現在の宇宙の膨張速度を表すハッブル定数の推定の観測間の不一致。特に宇宙背景放射からの推定には ΛCDM モデルが仮定されているため、このモデルの破れを示唆する可能性がある。

エミュレータ：

計算コストの高い数値シミュレーションを低コストの統計モデルに置き換えること、およびこれを実現するソフトウェア。「シミュレーションに基づく統計推論」を実現する方法論の 1 つ。



エミュレータを中核に据えた統計推論モデル。θ はパラメータ、d はデータを表す。

1. はじめに

ここ数年、宇宙の標準モデルとされてきた Λ CDM モデルの綻びを示唆する観測事実が報告されつつある。正体不明の重力源である冷たいダークマター (Cold Dark Matter: CDM) と、宇宙の加速膨張を説明するアインシュタインの宇宙項 Λ が、現在の宇宙の全エネルギーのおよそ 95% を占めるという、この驚くべき Λ CDM モデルの確立には、宇宙年齢約 38 万年頃、光で満たされたビッグバン宇宙が冷えることで迎えた宇宙の晴れ上がりの結果、はじめて直進できるようになった光子を捕えた、宇宙マイクロ波背景放射 (CMB: Cosmic Microwave Background) の観測が大きな役割を果たした。ところが、CMB から決定した現在の宇宙の膨張率、ハッブル定数 H_0 は、近傍宇宙からの直接の測定値と矛盾していることが報告され、ハッブルテンションと呼ばれている。CMB は現在の宇宙を見ている訳ではないため、そこからハッブル定数を導く際には Λ CDM モデルが仮定されている。従って、ハッブルテンションの最も単純な解釈は、 Λ CDM モデルの破れである。

最新の結果を見ると、プランク衛星による CMB の観測から $H_0 = 67.4 \pm 0.5 \text{ km s}^{-1} \text{ Mpc}^{-1}$ が¹⁾、一方でケフェイド変光星と Ia 型超新星爆発の観測データに基づく距離梯子から導かれた、SH0ES チームの最新の値が $H_0 = 73.04 \pm 1.04 \text{ km s}^{-1} \text{ Mpc}^{-1}$ となっており²⁾、両者は 5σ 程度の違いを示している¹。近傍宇宙の直接測定としては、距離が比較的良好に分かっている別の星を利用したもの⁴⁾や、重力レンズ多重像を持つクエーサーの像の間の到来時間の遅延⁵⁾なども利用可能であり、これらの観測からもプランク衛星の結果よりも速い膨張速度が報告されている。これが本当に Λ CDM を超える宇宙モデルを示唆すると言えるのか、完全にコンセンサスが得られたとは言いがたいが、この矛盾を説明しうる様々なシナリオが検討されている³⁾。

既存の宇宙膨張測定手法に加え、電磁波対応天体を持つ重力波源⁶⁾や、高速電波バーストの分散量度²⁾の赤方偏移依存性⁷⁾など、新たな天体現象を利用した手法が提案・実装され、その有効性が検討されている。現状ではサンプル数が限られる、あるいは系統誤差が完全に理解されていないため、先に挙げたものと比べてハッブル定数の制限としては弱い、今後サンプル数の大きな増加が見込まれ、強力な手段となることが期待される。今後の観測から、 Λ CDM モデルの検証がますます徹底的になされるであろう。

2. 揺らぎの宇宙論

宇宙膨張の精密観測の行き着く先はどこだろうか。現在の膨張速度に関する上記の課題も重要だが、宇宙膨張の歴史を明らかにすることで、加速膨張という大問題に迫るこ

とが大きなゴールの 1 つであろう。一般相対論に未知のダークエネルギーを導入するのか、重力法則自体を修正するのか⁸⁾といった本質的な問いに、将来の観測は答えることはできるだろうか。ダークエネルギーの導入、重力の修正のいずれにせよ、確たる物理的、数学的原理がないまま、大きな自由度を内包する経験則的なモデルを用意すれば、事実上任意の宇宙膨張を説明してしまい、いつまで経ってもこの問いに答えることができないだろう。

そこで、宇宙膨張と同時に、揺らぎの大きさを整合的に説明することを要請すれば、より厳しいテストが可能となる。統計的有意性は 2σ 程度と小さいものの、実は既に揺らぎの大きさに関して、近傍宇宙と CMB の間の食い違いが報告されている (詳しくは本記事後半とあわせて、解説記事⁹⁾を参照されたい)。以下、宇宙の揺らぎの予言、観測、解析にまつわる困難と、解決策について見ていこう。この節で触れる事項については、図 1 にもまとめる。

2.1. 宇宙の揺らぎの理論予言

より詳しい解説は先にも挙げた解説記事⁹⁾に譲るが、宇宙における物質密度の非一様性の進化を簡単にまとめると、インフレーション期に生成された微小な揺らぎを出発点とし、光子を主成分とするビッグバン宇宙における宇宙プラズマの物理 (CMB)、温度が十分に冷えて光子のエネルギー密度が下がり、一転して物質優勢となった宇宙における重力が駆動するによる構造の成長を経て、星やその集合である銀河といった天体の形成に至った、となる。現在の宇宙は物質の密度をダークエネルギーの密度が追いついてしばらく経過した時期に当たる。

最初の段階である揺らぎの生成では、単一のスカラー場による単純なインフレーション模型の一般的な帰結として、振幅がほとんどスケールに依らず、正規分布に従う断熱揺らぎがある。ここで、「断熱」と言っているのは、CMB にせよ複雑な過程を経て形成された銀河が作る大規模構造にせよ、共通する唯一無二の祖先であるスカラー場の量子揺らぎに帰着する、という極めて重要なことを示唆している。様々な観測手段を通じて宇宙の各段階の揺らぎを測定できれば、揺らぎについての一貫したストーリーを描き出せるというわけだ。このおかげで、近傍宇宙の大規模構造の分析から、宇宙の初期条件とその後の進化を支配する物理法則についての知見を引き出すことができる。CMB とは時期も距離スケールも異なる上、原則、空間 3 次元のデータを提供するため、大規模構造の持つ宇宙論的情報の潜在的価値は大きい。特に、宇宙がダークエネルギー優勢に転じたのは宇宙年齢にしておおよそ 100 億年頃であり、初期宇宙を写し出した CMB よりもダークエネルギーに直接的にアプローチできる。

¹ 本稿では誤差は一貫して 68% 信用区間で表記する。

² 電子プラズマを伝搬中の電波は、周波数毎に群速度が変わるため、周波数に応じた到来時刻にばらつきが生じる。このことから電波源と観測者の間の電子の柱密度が分かり、距離の指標として利用できる。

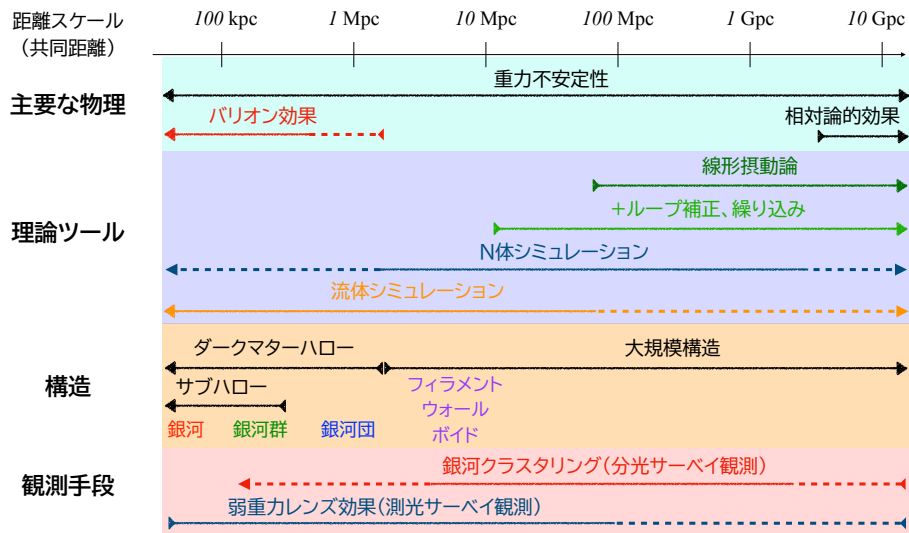


図1 様々な距離スケールにおける宇宙の大規模構造形成に関する物理過程，理論計算手法，対応する構造，観測手段をまとめたもの．銀河スケールから現在の地平線スケールまでを示している．全域にわたり重力不安定性が主な成長源となって，諸階層の構造を形成する．有効なスケールの異なる様々な計算手法が存在し，相補的な役割を担っている．本稿では，観測手段として，銀河クラスティング及び弱重力レンズ効果に着目する．

2.2. 非線形領域を攻略する

宇宙の揺らぎの歴史の中で，時間とともに揺らぎは大きくなるため，線形化した基礎方程式などを使った解析的計算による評価は現在に近くなるほど難しくなる．特に難しいのが，最終段階に当たる天体形成で，重力以外の様々な素過程を考える必要が生じる．以後は質量を持つ CDM と通常の物質（以下バリオンと呼ぶ）の密度揺らぎ $\delta_{\mathbf{k}} = \rho_{\mathbf{k}}/\bar{\rho}$ に注目しよう． $\bar{\rho}$ は宇宙の平均密度， $\mathbf{k}$ は波数ベクトルを指す．時間軸と同時に考える必要があるのが，距離スケール $k \equiv |\mathbf{k}|$ に対する依存性である．以下で，特にスケール依存性を気にしない場合には単に δ と表記する． Λ CDM では，密度揺らぎ $|\delta_{\mathbf{k}}|$ は小スケールほど大きく，より早い時期に非線形段階に至る．逆に，大スケールでは線形解に漸近し，宇宙の初期条件を色濃く残していると考えられる．

非線形性の弱い領域では，摂動展開の高次項を逐次的に計算することで，非線形性を評価する試みが精力的に進められてきた¹⁰⁾．最近では，高次項の漸近的振る舞いを手がかりとした繰り込み（再和）の手法^{11, 12, 13)} や，複雑な小スケールの揺らぎの進化に興味ある大スケールから切り離す有効場の理論^{14, 15, 16)}（詳しくは後述）も開発されている．しかし，揺らぎが $\mathcal{O}(1)$ に届き，いよいよ重力崩壊して天体が形成されるほどの強非線形領域に至ると，数値シミュレーションに頼らざるを得ない．現在の宇宙ではおおよそ 100 Mpc 程度の大スケールまで線形理論が有効であり，高次摂動を加えても数 10 Mpc が関の山である³⁾．

重力が支配的な領域では，質量を担う CDM およびバリオンを代表する「粒子」を多数用意し，それらの間の重力

相互作用を膨張時空の元で解く「宇宙論的 N 体シミュレーション」が有効である．計算資源の大規模化とともに，効率良い計算手法が整備され，今や 10^{12} 体を超える規模の計算がなされる時代に突入した¹⁷⁾．この手法を用いれば，上は Gpc スケールから，下は 10–100 kpc 程度まで正確に追いかけることができる．小スケールの限界付近では，ダークマターが密集し，重力的に束縛された構造であるハローや，その内部構造を調べることができる（図2）．大小のハローは階層を成して存在し，最大級のは銀河団，内包される小構造（サブハロー）は銀河におおよそ対応する．ハローと銀河を統計的に対応づけることで，銀河の諸性質を完全に特定することなく理論予言を与えることが可能で，ハローモデルとして知られている¹⁸⁾．

より小さいスケールでは，ガスの放射冷却，星形成，そして超新星爆発や活動銀河核などの天体現象を通じた周囲へのフィードバックといった重力以外の効果が重要となる．これらは，重力だけのモデルが予言する密度分布を変更するとともに，銀河などの光を放つ天体の形成を駆動する．ところが，これらの素過程の典型的スケールは，宇宙論解析に使われる揺らぎのスケール (Mpc ~ Gpc) からかけ離れており，第一原理的な計算は事実上不可能となっている．考えられる素過程を列挙し，サブグリッドモデルとして流体シミュレーションに実装するのが精一杯で，モデルに含まれる経験則的な関数やパラメータは観測データと合うように調整される．シミュレーションで銀河まで作り，その色や明るさなどまで知りたい場合，このような手法に頼らざるを得ない． N 体シミュレーションと比べて，同数の質量素片あたりに要する計算量が膨大なため，カバーできる体積はせいぜい数 100 Mpc 程度にとどまる¹⁹⁾．

いずれの方法を取るにせよ，我々は特定の観測装置を用

³⁾ 本稿では距離の単位として pc（パーセク； $1\text{ pc} \simeq 3 \times 10^{16}\text{ m} =$ 年周視差が 1 秒角となる距離）を用いる．参考までに天の川銀河の直径が約 50 kpc．近傍の大きな銀河団である，おとめ座銀河団で半径 2 Mpc 程度．

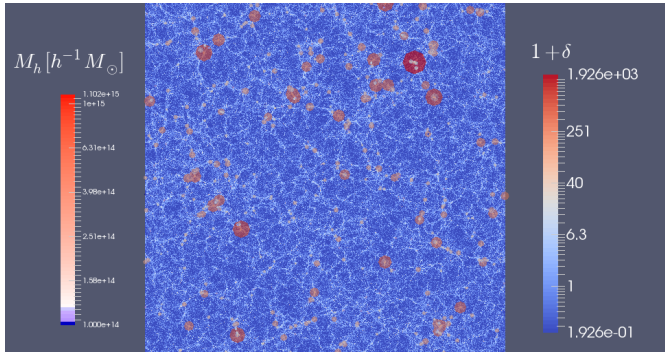


図2 大規模構造中のダークマターハローの分布。背景の色は低密度領域を青系、高密度領域を白～赤系で描画、ハローは質量に応じて異なる大きさの丸印で表示している。大質量のハローはフィラメントの交差点に形成され、付近に中小のハローが群れ集まっている。

いて、明るさや色などについて特定の基準を課すことで定義した銀河サンプル⁴の性質を第一原理的に予言する方法を持ち合わせていない。従って、銀河の空間分布の統計的性質に基づいて宇宙論モデルをテストするとき、その銀河サンプルが持つ特性と宇宙論的情報とを区別することは一般に難しい。注目すべきは、たとえサーベイ領域全体にまたがるような大スケールの揺らぎだけに着目したとしても、銀河の大域的分布を通してこれを測定する以上、銀河形成に関わる極めて局所的で複雑な素過程に左右される点だ。銀河の大域的分布は、質量を持つ物質全体の分布と似通っていると考えられるが、前者は通常物質であるバリオンが極めて高密度に密集した地点にのみ存在する一方で、後者はCDMが支配的である。このため、2つの分布の違いを「銀河バイアス」と呼ぶが、どれだけ大きなスケールで眺めても一般に $\delta_k \neq \delta_{g,k} \equiv n_{g,k}/\bar{n}_g$ であり (n_g は銀河の数密度)、 δ に刻まれた宇宙論的情報の評価を困難にする。本来の揺らぎの大きさと、銀河バイアスとの縮退を解くことのできる(複数の)観測量(の組み合わせ)を作れるかどうか、揺らぎを使った宇宙論テストの鍵となる。

2.3. 近年の揺らぎの測定の潮流

銀河バイアスの困難はどのように回避できるであろうか。ここでは具体的な観測手段に基づいて見ていこう。大域的な宇宙の密度揺らぎを見る最も直感的な方法は、銀河の分光観測が描き出す銀河クラスターだ。多数の銀河に対して天球面上の位置に加え、分光により赤方偏移を測定する。赤方偏移は光が光源から観測者に届くまでの宇宙膨張に起因するので⁵、そこから奥行き方向の距離が決定でき、銀河の3次元分布が得られる(図3右)。こうして得られた密度揺らぎは、前述の銀河バイアスの影響をそのまま受けているため、ダークマターを主成分とする宇宙の物質全体

の揺らぎの大きさと銀河バイアスの特性とを見分けることができない。これは線形理論の枠内では $\delta_g = b_g \delta \propto b_g \sigma_8$ と書ける。揺らぎの大きさは慣習的に σ_8 パラメータで表されることが一般的となっている⁶。宇宙論的に重要な情報である σ_8 は、銀河サンプルの特性である線形バイアスパラメータ b_g と完全に縮退する。

一方で、銀河バイアスに左右されない観測として知られるのが、言わずと知れた一般相対論の重要な帰結の一つである重力レンズ効果である。特に、大域的な宇宙の揺らぎを捕らえるのに適しているのが、銀河の形状の測光観測から得られる弱重力レンズ効果である。これは、銀河の像の歪みの程度が小さく、多重像に分離したり、極端に細長く伸びた構造にならない段階を指す。銀河は固有の形状を持っており、個別の天体に対して弱重力レンズ効果を本来の形状から分離することはできないが、空間的に近接する多数の銀河の像を測定することで、統計的に弱重力レンズ効果を引き出すことができる。現在すばる望遠鏡に搭載された広視野カメラ Hyper Suprime-Cam (HSC) を用いて進められているのがこの効果の測定だ^{20, 21}。重力レンズ効果は、観測された銀河(ソース銀河)と観測者の間に横たわるあらゆる構造によって引き起こされる。光る光らないに関わらず、質量を持っていれば寄与するため、銀河バイアスのような不定性なく密度揺らぎ δ を直接測定できる点が、この効果の大きな魅力である。光の伝搬中、どの位置の構造がどんな効率で寄与するかは、レンズ核関数(図3の $W(z')$; z' は赤方偏移=観測者からの距離)で決まり、おおよそソース銀河と観測者の中央付近の寄与が最も卓越する。 $W(z')$ は大きな広がりを持っているため、様々な時刻(=奥行き方向の距離)の構造を平均したものしか分からないこと⁷、また、銀河の分光観測と違い、天球面上に射影された2次元情報しか得られないことが弱点となっている。

そこで、分光、測光の2つの観測データを組み合わせ、両者の長所を活かした解析はできないだろうか。実は、各々から得られた銀河の空間分布と形状の2つの相互関数を取ったものを考えればよく、これを銀河・銀河レンズ効果と呼ぶ(図3中央)。より遠方にある銀河の形状と、前景にある赤方偏移既知の銀河サンプルの疎密の間の空間相関を見たこの効果は、銀河クラスターリング・弱重力レンズの双方の問題点にアプローチする。まず、銀河バイアスの不定性の問題を抱える銀河クラスターリングの立場から見ると、銀河・銀河レンズ効果は、興味のある銀河サンプルを「レンズ銀河」と見なし、これらの周りの密度場が背景のソース銀河から出た光をどれだけ曲げたかを見たものと言える。重力レンズには全ての質量が寄与するので、この効果の測定を通じて、レンズ銀河の周囲の物質の密度構造(CDMを含む)を

⁴一般に、大域にわたって均一性、信頼性の担保された銀河サンプルを構築することは難しい。誤って星などが混入するのを避け、狙った性質を持つ銀河だけを均一に選ぶよう、様々な基準が課される。

⁵他の原因として銀河の速度に由来するドップラー効果がある(4.2節)。

⁶線形近似のもとで $8h^{-1}\text{Mpc}$ で評価した値。

⁷この際、異なる赤方偏移にある構造は、長さが同じでも見込み角が異なるため、時間方向だけでなく、距離スケールも混ざることにも注意が必要。

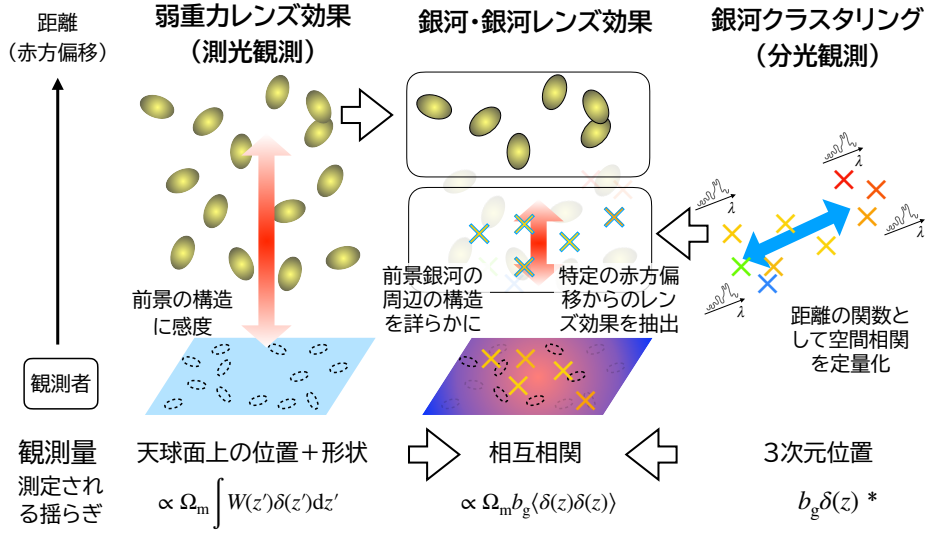


図3 宇宙の大規模構造の測光および分光観測から分かる情報。測光観測から多数の遠方銀河（ソース銀河）の詳細な形状を測定することで、統計的に弱重力レンズ信号を引き出すことができる（左）。これは、レンズ核関数 W に従って光源から観測者の間の様々な位置（＝時刻）の構造を拾い上げる。一方で、分光観測により奥行き方向の位置（赤方偏移）を決定することで、銀河の3次元分布が得られる（右）。こちらには理論的な予言が難しい銀河バイアスの影響が現れる（* これに加えて実際にはドップラー効果が混入する。詳しくは4.2節を参照されたい）。後者をレンズ銀河として利用し、背景にあるソース銀河の形状の統計的性質を見ることで、特定の赤方偏移の構造を調べることができる（中）。こうして得られた複数の観測量を組み合わせることで、銀河バイアスの不定性を除去した頑強な宇宙論テストが可能となる。

知ることができる。一方で、重力レンズ側からすると、ソース銀河と観測者の間の全ての構造の重み付き平均で表現される銀河像の歪みのうち、赤方偏移既知のレンズ銀河の存在する時刻からの寄与だけを引き出したものと見なせる。

この相互相関に加え、クラスターリング、レンジングそれぞれの自己相関を加えた計3種の相関関数の統合解析が現在の主流となっており「 3×2 点解析」と呼ばれる。3つの相関関数の持つ銀河バイアス及び宇宙論パラメータに対する異なる依存性のおかげで、 3×2 点解析は銀河バイアスの不定性に対して頑強な宇宙論テストを可能にする。ここから最も強く制限することが可能な宇宙論パラメータの組み合わせは、観測の深さに加えて解析に用いられる理論モデルにも依存するが、歴史的に弱重力レンズ効果単独の解析に用いられてきた $S_8 = \sigma_8(\Omega_m/0.3)^{0.5}$ を踏襲することが多い⁸。これは、重力レンズ効果は揺らぎ δ ではなく、重力源となる物質の総量 $\rho = \bar{\rho}\delta$ で決まること、さらにレンズ効率 $W(z')$ が宇宙の幾何学に依存することに起因しており、純粋な揺らぎの大きさを表す σ_8 と物質の密度パラメータ Ω_m を組み合わせたものが精度良く測定される。

このような解析で現在すばる HSC をやや先行しているのは、欧州主導の Kilo-Degree Survey（KiDS）と米国主導の Dark Energy Survey（DES）である。2つのグループは既に 3×2 点解析の結果を報告しており⁹、KiDS²²⁾ は

⁸最も強く制限される組み合わせになるよう、冪0.5を最適化したものもしばしば見られる。

⁹KiDSは現在最大の銀河分光掃天観測である Sloan Digital Sky Survey (SDSS) のデータをレンズ銀河として利用しているが、南天を観測するDESは大規模な分光観測とオーバーラップがないため、測光銀河の色情報から見積もった大まかな赤方偏移に基づく銀河サンプルを自前で用意し、レンズ銀河として利用している。

$S_8 = 0.766^{+0.020}_{-0.014}$ 、DES²³⁾ は $S_8 = 0.776^{+0.017}_{-0.017}$ を報告している。これらの値は、プランク衛星の観測から導かれる値、 $S_8 = 0.834 \pm 0.016$ と比べるとおよそ $2 \sim 3\sigma$ 程度低くなっており、 S_8 テンションとも呼ばれている。

3. 統計推論とエミュレータの導入

3.1. 宇宙論パラメータのベイズ推定

観測データと理論モデルを比べて、モデルパラメータに制限をつける際、宇宙論では一般的にベイズの定理に基づいて議論される：

$$P(\theta|d) \propto P(d|\theta)P(\theta) \quad (1)$$

左辺は、観測データ d が与えられたとき、モデルパラメータ θ の従う確率分布で、事後分布と呼ばれる。一方で、右辺は、 θ を指定したときの d の分布（尤度）と、実験前に θ について知りうる情報（事前分布）の積となっており、物理モデルを使って計算することができるのが尤度である。ベイズ推定とは、ベイズの定理を使って事後分布を決めることである。

数値計算により順モデル d_{sim} が作成可能だとしても、唯一の観測データ d_{obs} と比較して事後分布を推定するのは一般に難しい。これは、尤度 $P(d|\theta)$ を評価するのにかかる計算コストに起因するものである。安直には θ を固定して多数回の順モデル d_{sim} を発生し、 d_{sim} の確率分布を推定する操作を、 θ を変えて繰り返す必要がある。式(1)を見ると、尤度の評価は $d = d_{\text{obs}}$ の地点だけで行えば十分なのだが、この方法では d_{sim} が d_{obs} 付近に来るようあらかじめ狙い撃ちできないため、膨大な数の順モデルを要する。 d の次

元が高いほど、観測データ d_{obs} 付近にヒットする確率が低くなり、効率良いサンプリングが難しくなる。 d_{sim} と d_{obs} の比較から統計的推論を行う問題は、「シミュレーションに基づく推論」として分野横断的な一大テーマとなっている²⁴⁾。この問題において計算コストの大半を占める、尤度の評価を迂回するという意味で、「尤度無し推論」も、本来の意味を少し離れて、同義のキーワードとして使われる。

宇宙論がやや特殊なのは、観測量の背後にある物理過程がよく分かっており、その起源が正規分布に帰着する点にある(2.1節参照)。このため、揺らぎの持つ情報を少数の数量に圧縮した「要約統計量」が知られており(大スケール極限では必ず2点相関関数が情報をくまなく取り込んだ「十分」統計量)、なおかつこれを観測領域の形状などに依存せずに無バイアスに推定する方法が古くから開発されている。統計量を経由した推定では、式(1)のデータ d を統計量 $S(d)$ に置き換える。これにより、データの次元を大きく削減することができる。また、多くの場合、統計量の尤度は正規分布に近いことが知られており¹⁰⁾、期待値と共分散行列を特定するだけで十分となっている。通常は、共分散は観測の設定で決まると考え、宇宙論パラメータを固定して生成した多数の模擬データから推定し、宇宙論パラメータ依存性を無視する。この近似の妥当性は問題設定にも依るが、概ね妥当であると考えられている。残っているのは期待値を十分な精度で予言する部分だけだ。こちらについてはもちろん宇宙論パラメータの関数として決定する必要がある。

「エミュレータ」は高コストな数値計算を低コストな統計モデルで置き換えることを指し、シミュレーションに基づく推論を可能にする。工学では代理モデルとして知られていたこの方法論は、宇宙論ではハイトマン (K. Heitmann) らの2006年の論文²⁵⁾あたりから検討され、概念実証レベルの研究を経て、観測データに応用可能なレベルのものが構築されるようになってきた。以下ではこれを実現するための有力な手法の一つであるガウス過程回帰 (Gaussian Process Regression; GPR¹¹⁾) について紹介する。

3.2. 回帰問題

前節の議論に従うと、要約統計量の期待値 $\langle S(d) \rangle$ を宇宙論パラメータ θ の関数として与えることが最後の課題となる。これを数値計算で行おうとした場合、計算コストの大きさのため θ として少数の標本しか選ぶことができない。一般に、連続値を取る未知の関数 $f(x)$ に対し、いくつか抽出した入力 x_i に対して出力 f の測定値 y_i が与えられたとき、元の関数 $f(x)$ を推定する問題を回帰と呼ぶ。入力 x_i 、出力 y_i とも1変数である必要はなく、一般にベクトル量を

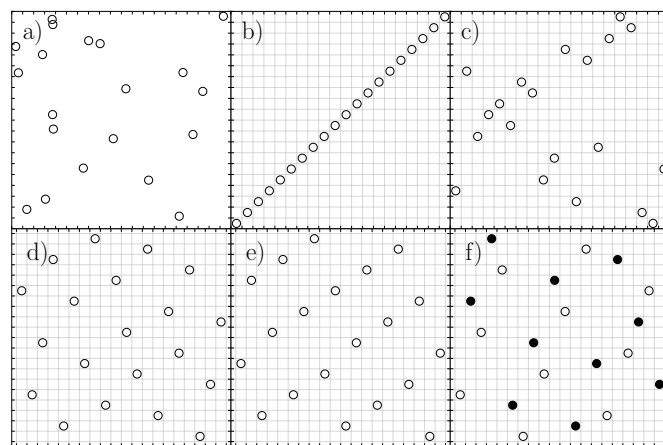


図4 20 サンプル点から成る2次元デザイン。a: 一様分布からのランダム標本, b~f: 様々なラテン超方格 (b: 対角線上のデザイン, c: ランダム標本, d: マキシミン距離設計, e: 非等方な距離計量に対するマキシミン距離設計, f: 2層マキシミン距離設計のスライス LHD)。

考えてよい。入力が多次元の場合、重回帰と呼ぶのが一般的である。エミュレータは、有限個のシミュレーションサンプルから入出力間の対応関係を学習し、任意の宇宙論パラメータに対して所望の統計量を予測することで、回帰問題を担う統計モデルである。

回帰の概念自体は内挿(補間)や外挿と同じなので、機械学習などといった大それたものを持ち出さなくてもよいように思われるかもしれない。単純なものでは線形補間、滑らかさを要求するのであればスプライン補間などの区分多項式の利用が考えられるであろう。ところが、入力次元の高い重回帰の問題では、低次の多項式でも自由に決められる係数の数が膨大になる。さらに、区分多項式を使うために十分に細かい間隔の格子状区域に分けるには、膨大なサンプル点が必要となる。この問題を回避する方法として、ここではまず、現実的な計算時間で実行可能な少数のシミュレーションを、多次元入力空間にどのように配置するのが有効か議論し(3.2.1節)、続いて具体的な回帰モデルとしてGPRを概説する(3.2.2節)。

3.2.1. ラテン超方格サンプリング

この問題を考える上で、最初に重要なのは多次元入力空間をいかに効率よく埋めるかである。その最適解は、目的となる関数 $f(x)$ の性質に応じて異なるが、事前に何も知らない場合¹²⁾に最も無難なものとしてよく採用されるのが、ラテン超方格サンプリング^{26, 27)}である。

ラテン超方格の定義は極めて単純で、超直方体状の入力空間を1次元あたり実行する予定のシミュレーション数と同数の均一な格子に区分した上で、どの行どの列をとっても唯一のサンプル点を置く手法である(図4のbからf)。こ

¹⁰⁾ これは初期条件が正規分布であることとは直接関係なく、統計量の推定に多数の確率変数の平均操作が入ることで、中心極限定理から導かれる。

¹¹⁾ 地球統計学では古くから利用され、考案者の名前を取ってクリギングと呼ばれている。

¹²⁾ 実験・シミュレーションのみを通じて解が得られる設定での最適化問題で、ブラックボックス最適化と呼ばれる。

れにより、どの入力次元に対しても均一にサンプルを取ることが保証される。ただし、この条件だけでは多数の選び方が許されるため (b: 対角線上, c: ランダム), 追加の条件をかけてより望ましいものを選ぶ。空間充填度の均一性の観点からは、最近接のサンプル間の距離を最大化するマキシミン距離設計が1つの解となる。ただし、入力空間に絶対的な距離指標は存在しないため、d (等方な距離計量) や e (横方向に充填度を高めた非等方計量) のように異なるデザインが構築できる。機械学習の応用を考えた場合、訓練データ、検証データに分類し、それぞれのデータについてサンプリングの均一性を要請することも考えられるが、f のように個々のカテゴリに属するサンプルと、全サンプルとを同時に最適化したものを作ることでもできる²⁸⁾。

3.2.2. ガウス過程回帰

シミュレーションを実行し、 (x_i, y_i) の組からなるデータが蓄積したら、いよいよ回帰である。回帰問題は数学的には不良設定問題であり、データ点を再現する無数の解が存在する。ここに、関数 $f(x)$ についてこうあって欲しいという期待を何らかの拘束条件として課すことで、それらしい解を構築すること (正則化) を考える。関数の連続性や平滑さなどがこれに相当する。

GPR はバイズの枠組みにより正則化を自然に実装する。はじめに、関数 f は正規分布に従って確率的に発生する、という仮定を事前情報として与える。「関数が確率的に発生する」という過程はやや想像しにくい、サイコロを振るたびに様々なカーブが描かれるようなイメージで良い。単一の数値に関する通常の正規分布は平均と分散で完全に特定されるが、関数全体が確率的に発生する状況では、平均関数 (通常はゼロに取る) と共分散関数 (核関数とも) がこれらに代わる。図5の “prior” と表示されたパネルは、異なる核関数に従って無作為に発生させた関数の例を示している。次に、実験データを加味した事後分布について考える。これは、データの存在下での条件つき確率を計算することで実現できる (“posterior” と示されたパネル)。GPR が便利なのは、正規分布という素性の良い分布を仮定することで、あらゆる確率計算が解析的に可能であるところにある。ここまでで説明した操作は全て仮定された核関数の元での行列積の形に帰着する。

こうして得られた事後分布の良し悪しは、最初に仮定した核関数の妥当性に左右される。適切な核関数をあらかじめ知る術はないので、通常はパラメトライズした核関数を用意し、実験データと最も整合的であるようパラメータ最適化を行うことでより正確な予言を実現する。図5の上下では、異なる関数形の核関数に対してこれを行った結果を示しており、左半分が最適化前、右半分が最適化後に対応している。この例では、答えの関数が周期性を示している

ことから、最適化後は、核関数に三角関数を入れた下のパネルの方が、データが欠損した領域においても真の値に近い予言を返す。このように関数 f について何らかの知見を持っていれば、適切な核関数の選択という形で実装することができる。

GPR 自体はバイズに基づく古典的な手法である。しかし、その背後にあるデータに自然に適応する形でより良い予測を行う仕組みは、ニューラルネットなどの機械学習の手法と共通している。機械学習においては、ニューラルネットの隠れ層の数などの超パラメータを調整するが、GPR の核関数がちょうどこれに該当し、関数の滑らかさや典型的な振幅を調整する。GPR は、多次元入力空間でも自然に機能する。図6は2つの入力次元を持つ関数に対する回帰問題へのGPの応用例を示している。まず、核関数をでたために選んだ場合、事前分布から抽出した関数には、大きな振幅を持つ細かい波状の構造が見られ (中央上)、実験データを加えて得られた事後分布にもこれを引きずった特徴が見られる (右上)。一方で、核関数をデータ点に対して最適化することで、真の関数の持つダイナミックレンジや滑らかさに近い事前分布が得られ (中央下)、事後分布も真の関数により近いものとなっている (右下)。GPR に要する計算量は、入力次元数ではなく、用意したデータ点の数で決まる。よって、前節で述べた方法で少数のサンプルにより効率良く多次元入力空間を埋めておけば、次元が高い問題設定でも問題なく機能する。GPR の事後分布の平均や分散の推定には、逆行列の推定が必要であるため、データ点が多くなると、そのままでは機能しなくなることだけ注意しておきたい。

4. ついに実現したエミュレータ宇宙論

4.1. ダークエミュレータとすばるへの応用

我々は、ここまで述べた方法を利用して、すばるのデータを最大限に活かし、クリーンな形で宇宙論テストを行うべく、2015年度より「ダーククエスト」シミュレーション計画を推進している²⁹⁾。2018年までに完了した第1期の取り組みでは、 Λ CDM にダークエネルギーの状態方程式 w を入力として加えた、6次元 w CDM モデルを考え、効率的に選んだ101のサンプル点において、 $2,048^3$ 粒子の N 体シミュレーションを実行し、ハローの統計的性質を精度良く計算するダークエミュレータを開発した。この計算には、国立天文台のスーパーコンピュータ「アテルイ」「アテルイII」を1,000コア程度用いて1モデルにつき約2日を要した。ダークエミュレータは、観測可能性を必ずしも保証しない諸量の予言に終始していたこれまでのエミュレータと異なり、ハローを媒介して様々な種類の銀河を取り扱える具体的な枠組みに乗せたことで、実際に得られる様々なデータに応用可能なソフトウェアとなった。これを用いると、スーパーコンピュータで2日間かかる計算量を、ラッ

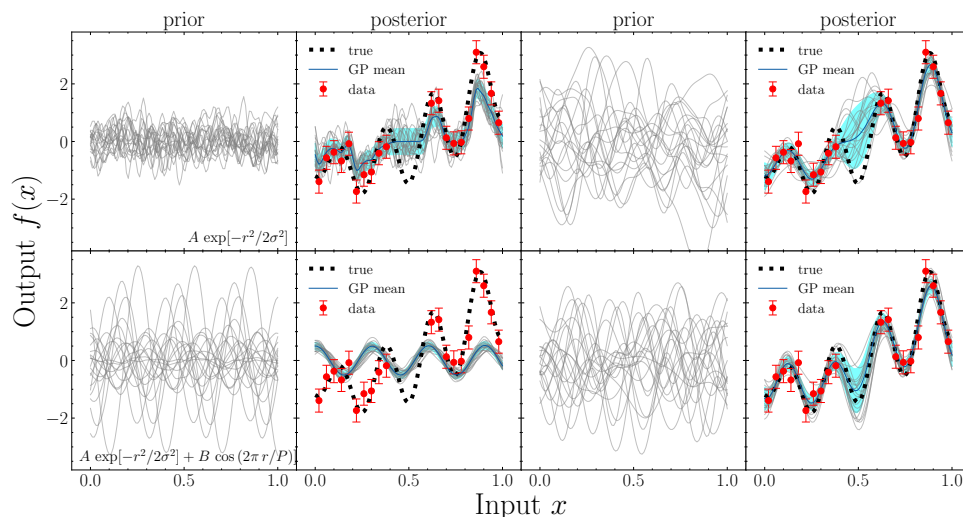


図5 簡単なテスト関数に対するGPRの適用例. 2列目, 4列目に点線で示された関数を, 同じく誤差棒で示された実験データ(各パネルで共通)から推定する問題設定. GPは平均関数(ゼロに取った)及びカーネル関数(上下それぞれ左パネルに書かれた関数を採用)で特徴づけられ, 事前情報として機能する. 1, 3列目はそれぞれのGPから抽出された20の標本関数を示している. 2, 4列目はデータ点による条件付き確率として得られた事後分布. 実線が平均, 影は 1σ で, 細い実線で20のランダム標本を合わせて示している. 左半分は最適化前, 右半分は超パラメータ最適化を行った結果を表す.

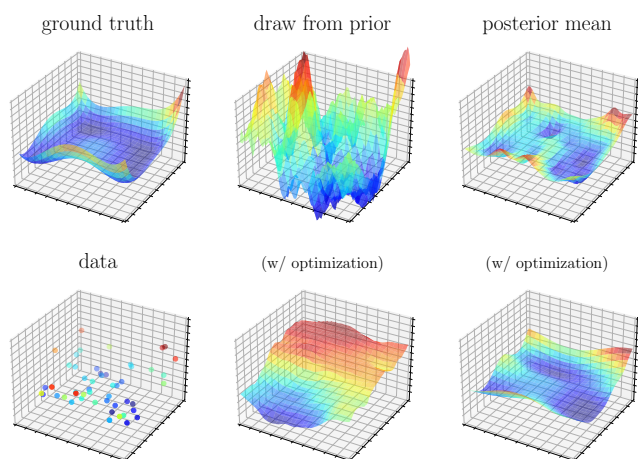


図6 2次元入力に対するGPRの例. 左上に示されたテスト関数からランダムに選ばれた, ノイズ入りのデータ点(左下)に対してGPRを試みる. 事前分布(平均ゼロ, Matérn32カーネルを持つガウス分布)から無作為に抽出した関数を中央に, データ点による条件付き確率として導かれる事後分布の期待値を右に示した. 中央, 右は, いずれも上下で超パラメータ調整なし, ありの場合を表す.

ブトップコンピュータで数秒に圧縮でき, バイズ推定の枠組みに無理なく乗せることができる.

そこで, これを利用した最初の大規模な宇宙論テストとして, すばるHSCの2016年春までに取得された初期データから測定した弱重力レンズ効果と, SDSSから得られた銀河クラスタリングとの統合解析を行った³⁰⁾. 待ちに待ったエミュレータ宇宙論の実現である. 系統誤差の理解の観点から, 3×2 ではなく銀河・銀河レンズ効果と銀河クラスタリングの 2×2 点解析にとどめたが, 制限値 $S_8 = 0.795^{+0.049}_{-0.042}$ を得ることができた. この制限は, 現時点では先に述べた2チームの結果からは見劣りするが, 組み合わせる観測量

と理論モデルの性質の違いから, 我々のチームはやや異なるパラメータの組み合わせ $\sigma_8(\Omega_m/0.3)^{0.17}$ を最もよく制限し, およそ4~5%の範囲に抑えることができています.

先に述べた, KiDSとDESが, それぞれ1,000平方度¹³および, 5,000平方度にもわたる広域の撮像データから結果を出しているのに対し, HSCの初期成果はたった140平方度の観測から来ていることを強調しておきたい. 8.2mの大口径の高い集光力と, 一度に1.77平方度もの広域を撮像できるHSCの広視野, マウナケアの観測に最適大気が可能とするシャープな天体像を活かし, HSCは競合する計画と比べてより遠くの宇宙まで観測し, 立体角あたりに3倍程度の高い数密度でソース銀河の像を取得することができる. これらの観測計画の比較については, 本稿で紹介した解析に先行して行われた, HSCの弱重力レンズの単独解析について紹介した記事³¹⁾も参照されたい. 最終的に1,400平方度を目指すHSCが完遂すれば, よりはっきりと S_8 テンションの真偽に迫ることができるであろう.

4.2. 更なる展開

ダーククエストが構築したデータベースの有効性は, 弱重力レンズ効果だけにとどまらない. 実は, 分光観測による銀河の位置決定には, 個々の銀河の持つ速度に起因するドップラー効果が混入する. 弱い等価原理から, あらゆる構造が同様に重力を受けると期待されるため, その結果生じた銀河の速度は, 銀河バイアスとは無関係に宇宙の構造を測定する可能性をもたらす. この効果についても同データベースからエミュレータを構築し³²⁾, SDSSのパワース

¹³最終的には1,500平方度を予定.

ペクトルの「全部入り解析¹⁴」を実現している³³⁾。SDSS グループによる公式の解析は、パワースペクトルの中でも理論化が比較的容易と思われる一部（バリオン音響振動と赤方偏移歪み）に特化したものであった。我々の解析では、揺らぎの大きさを表すパラメータを $\sigma_8 = 0.786^{+0.036}_{-0.037}$ と約 5% の精度で決定することができた。エミュレータを使うことで、余すことなくデータ分析ができるようになってきた賜物である。同時期に、揺らぎの摂動展開に基づく有効場の理論という全く別の方法を使って、同じ課題に挑戦するグループが現れた。その一つがイバノフ (M. Ivanov) らによる解析で、 $\sigma_8 = 0.721 \pm 0.043$ を報告している³⁴⁾。その後、複数のグループが独自の理論予言を用いて同様な解析を行い、概ね整合的な値を報告している。これらをプランク衛星による値 $\sigma_8 = 0.812 \pm 0.0073$ と比較すると、現状の達成精度では、 Λ CDM モデルの妥当性について決定的なことは言えない。

このように全く異なるアプローチから、整合的な結果が得られている点は大変興味深い。エミュレータ（ハローモデル）では、銀河はハローに作られることを要請し、ハローと銀河の対応関係については多数のパラメータを導入してできるだけ柔軟に表現できるようにするが、有効場の理論では、対称性や保存量などの観点から許されるあらゆる自由度をパラメータとして導入する、より保守的な方法を取っている。究極的には銀河形成の物理を全て特定することで、このような自由度を理論的に制限することが望ましいが、それが叶わない現状で、どの方法にどれだけクレジットを与えるべきか、客観的に判断することは難しい。そこで、詳細を伏せたままシミュレーションにより構築した模擬銀河カタログを分析するプログラム³⁵⁾ を実施し、業界全体が納得する 1 つの尺度を与えられないか模索している。

5. まとめと展望

ダークエミュレータの完成により、宇宙の大規模構造の重力レンズ・クラスタリング統合解析を、非線形領域まで正確に取り扱った N 体シミュレーションに基づく統計推論が可能となった。また、「全部入り」の理論モデルが手に入ったことで、すばる HSC と競合する DES、KiDS の 3 つの計画が測定したレンズシグナルの整合的のテストに、ダークエミュレータが利用されるなど、大きな国際プロジェクト間の連携にも一役買っている。

次なる課題は、エミュレータ宇宙論の結果をどのようにフィードバックし、シミュレーションデータベース増強から宇宙論パラメータ推定まで自律的に行う枠組みの構築にある（リードページの図を参照）。第一弾のエミュレータと、観測データに応用して得られた宇宙論パラメータの事

後分布のもとでは、もはやブラックボックス最適化問題ではなくなった。今後すばる HSC の弱重力レンズデータの統計誤差が低減するにつれ、理論モデルに要求される精度も高くなる。また、HSC の後は同じくすばるに搭載予定の多天体同時分光器 Prime Focus Spectrograph (PFS) を使った大規模分光掃天観測が予定されている。これを用いることで、ニュートリノ質量の絶対値¹⁵など、物理的に重要な課題に迫ることが期待されている。そのような問題にアプローチするには、エミュレータの入力次元の拡張と、高赤方偏移のより軽いハローに形成される銀河への対応が必要となる。

より急進的に機械学習に頼り、本稿で紹介した 2 点相関関数以上の情報を取り出す可能性についても議論が進みつつある。宇宙の密度場という超多自由度の確率場自体を機械学習にかけて、宇宙論パラメータを導く大規模数値実験³⁶⁾ や、最近では何と銀河 1 個の構造のみから宇宙論を推定しようという試み³⁷⁾ まで見られるようになってきた。これらはいずれも数値シミュレーションを使った教師あり学習に基づくものであり、その成否は用いたシミュレーション自体の信頼性にかかっている。完璧な宇宙構造形成シミュレーションは存在せず、解像度や計算領域の有限性に起因する数値誤差、経験則的に実装されたバリオン物理の妥当性など、クリアしなければならない問題は山積みである。なおかつ、学習に十分な数の模擬データを用意するには、実行可能な最大級のシミュレーションと比べ、遥かに小規模なものにとどめざるを得ないのが現実だ。今後、機械学習技術の流入がどこまで進むのか、説得力を持った形でこれまでの解析を凌駕する結果を出せるのか、期待と不安が半々と言ったところではなかろうか。

参考文献

- 1) N. Aghanim, *et al.*, [Planck Collaboration], *Astron. Astrophys.* **641**, A6 (2020).
- 2) A. G. Riess, *et al.*, eprint arXiv:2112.04510 (2021).
- 3) E. Di Valentino, *et al.*, *Class. Quant. Grav.* **38**, 153001 (2021).
- 4) W. L. Freedman, *et al.*, *Astrophys. J.* **882**, 34 (2019).
- 5) K. C. Wong, *et al.*, *Mon. Not. Roy. Astron. Soc.* **498**, 1420 (2020).
- 6) B. P. Abbott, *et al.*, *Nature* **551**, 85 (2017).
- 7) S. Hagstotz, R. Reischke and R. Lilow, *Mon. Not. Roy. Astron. Soc.* in press (2022).
- 8) 辻川信二, 日本物理学会誌 **69**, 444 (2014).
- 9) 高橋龍一, 日本物理学会誌 **71**, 598 (2016).
- 10) F. Bernardeau, *et al.*, *Phys. Rept.* **367**, 1 (2002).
- 11) M. Crocce and R. Scoccimarro, *Phys. Rev. D* **73**, 063519 (2006).
- 12) T. Matsubara, *Phys. Rev. D* **77**, 063530 (2008).
- 13) A. Taruya and T. Hiramatsu, *Astrophys. J.* **674**, 617 (2008).
- 14) D. Baumann, *et al.*, *J. Cos. Astropart. Phys.* **7**, 051 (2012).
- 15) J. J. M. Carrasco, M. P. Hertzberg and L. Senatore, *J. High Energy Phys.* **9**, 82 (2012).

¹⁴ これまでに解析されていた効果に加えて、パワースペクトルの広い波数領域にわたる形状から来る情報を引き出す。より専門的には、ここから物質と輻射 (光子) のエネルギー密度が同じとなる時期が分かり、これは現在の物質の密度の測定と等価。

¹⁵ ニュートリノはほぼ光速で飛び回るため、重力による構造形成を抑制する。これを手がかりに、ニュートリノ質量が全質量に占める割合が測定でき、別途分かっている数密度から質量が推定できる。

- 16) J. J. M. Carrasco, *et al.*, J. Cosmol. Astropart. Phys. **7**, 057 (2014).
- 17) T. Ishiyama, *et al.*, Mon. Not. Roy. Astron. Soc. **506**, 4210 (2021)
- 18) A. Cooray and R. Sheth, Phys. Rept. **372**, 1 (2002).
- 19) D. Nelson, *et al.*, Computational Astrophysics and Cosmology **6**, 2 (2019)
- 20) 高田昌広, 日本物理学会誌 **69**, 140 (2014).
- 21) 宮崎聡, 日本物理学会誌 **69**, 149 (2014).
- 22) C. Heymans, *et al.*, Astron. Astrophys. **646**, A140 (2021).
- 23) T. M. C. Abbott, *et al.*, [DES Collaboration], eprint arXiv:2105.13549 (2021).
- 24) K. Cranmer, J. Brehmer and G. Louppe, Proc. Natl. Acad. Sci. U S A **117**, 30055 (2020).
- 25) K. Heitmann, *et al.*, Astrophys. J. **646**, L1 (2006).
- 26) V. Eglajs and P. Audze, Problems of Dynamics and Strengths **35**, 104 (1977).
- 27) M. D. McKay, R. J. Beckman and W. J. Conover, Technometrics **21**, 239 (1979).
- 28) S. Ba, *et al.*, Technometrics **57**, 479 (2015)
- 29) T. Nishimichi, *et al.*, Astrophys. J. **884**, 29 (2019).
- 30) H. Miyatake, *et al.*, e-print, arXiv:2111.02419 (2021).
- 31) 日影千秋, 日本物理学会誌 **75**, 422 (2020).
- 32) Y. Kobayashi, *et al.*, Phys. Rev. D **102**, 063504 (2020).
- 33) Y. Kobayashi, *et al.*, e-print, arXiv:2110.06969 (2021).
- 34) M. M. Ivanov, M. Simonović and M. Zaldarriaga, J. Cos. Astropart. Phys. **5**, 42 (2020).
- 35) T. Nishimichi, *et al.*, Phys. Rev. D **102**, 123541 (2020).
- 36) A. Mathuriya, *et al.* e-print, arXiv:1808.04728 (2018).
- 37) F. Villaescusa-Navarro, *et al.*, e-print, arXiv:2201.02202 (2022).

著者紹介または非会員著者の紹介

西道啓博氏： 専門は大規模数値計算に基づく宇宙の大規模構造の理論研究。近年は機械学習の導入や観測プロジェクトとの連携に注力している。

(2022 年 4 月 28 日原稿受付)

Machine learning in observational cosmology: application of emulation to Subaru observations

Takahiro Nishimichi

abstract: Emulation is a technique that replaces costly simulators with much cheaper statistical models in Bayesian inference problems. We have developed DARKEMULATOR that predicts basic statistical quantities in the cosmological structure formation. This code is then applied to real data from Subaru Hyper Suprime-Cam as well as Sloan Digital Sky Survey. The combination of these two data sets, the former probes the weak gravitational lensing effect while the latter probes the three dimensional galaxy distribution on large scales, offers a unique opportunity to break the degeneracy between cosmology and galaxy physics. We report competitive bounds on the parameter S_8 , the amplitude of the cosmological fluctuations at present. This new methodology would serve as a realistic solution for simulation-based inference, which can reveal unexplored information on small scales where nonlinearity is significant.