1403/03/10
2024/05/30

بسم الله الرحمن الرحیم

Data modeling : Table of Contents
(Optimization)

Main Subjects that are Covered by my lecture are:

× 0: Relevant Refs.

Most relevant References: Books

1) Modern Cosmology, S. Donelson.
2) The Cosmic Microwave Background, R. Burrer
3) Physical foundation of Cosmology, V. Mukhanov
4) Cosmology, S. Weinberg
5) Neutrino Cosmology, J. lesgourgues et.al
6) Galaxy Formation and Evolution, H. Mo, F.V. den Bosch and S. White
7) Statistics of Galaxy Distribution, V. Martinez, E, Saar
8) Cosmological Physics, J.A. Peacock
9) Topological Complexity of Smooth Random Functions, Adler, Robert, Taylor, Jonathan E., Springer, 2009.
10) Geometry, Topology and Physics, Mikio Nakahara, 1990.
11) Analysis and Data-Based Reconstruction of Complex Nonlinear Dynamical Systems, Using the Methods of Stochastic Processes, M. Reza Rahimi Tabar, Springer, 2019
12) Zomorodian, Afra. "Topological data analysis." Advances in applied and computational topology 70 (2012): 1-39.
13) Edelsbrunner, Herbert, and John Harer. Computational topology: an introduction. American Mathematical Soc., 2010.

بسم الله الرحمن الرحیم

دانشگاه شهید بهشتی
دانشکده فیزیک

Review on Data Analysis Methods & Cosmological Simulations

Seyed Mohammad Sadegh Movahed
Physics department, Shahid Beheshti University
Computational and Cosmological group (CCG-SBU)
www.smovahed.ir

Workshop on Computational Cosmology:
From Theory to Observation
1-2 August 2023

...al Cosmology Group
... Beheshti University

1

COMPLEXLAB

Most relevant References: Papers and Lectures

1) Matsubara, Takahiko. "Statistics of smoothed cosmic fields in perturbation theory. I. Formulation and useful formulae in second-order perturbation theory." The Astrophysical Journal 584.1 (2003).
2) Vafaei Sadr, A., and S. M. S. Movahed. "Clustering of local extrema in Planck CMB maps." Monthly Notices of the Royal Astronomical Society 503.1 (2021): 815-829.
3) Masoomy, H., et al. "Persistent homology of fractional Gaussian noise." Physical Review E 104.3 (2021): 034116.
4) Masoomy, H., S. Tajik, and S. M. S. Movahed. "Homology groups of embedded fractional Brownian motion." Physical Review E 106.6 (2022): 064115.
5) Lesgourges, J. "Cosmological perturbations." Searching for New Physics at Small and Large Scales: TASI 2012. 2013. 29-97.
6) Mostaghel, Behrang, Hossein Moshafi, and S. M. S. Movahed. "Non-minimal derivative coupling scalar field and bulk viscous dark energy." The European Physical Journal C 77 (2017): 1-22.
7) Pranav, Pratyush, et al. "Topology and geometry of Gaussian random fields I: on Betti numbers, Euler characteristic, and Minkowski functionals." Monthly Notices of the Royal Astronomical Society 485.3 (2019): 4167-4208.
8) Pranav, Pratyush. "Topology and geometry of Gaussian random fields II: on critical points, excursion sets, and persistent homology." arXiv preprint arXiv:2109.08721 (2021).
9) Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, Astro- phys. J., 304, 15
10) Bond, J. R., et al., The Astrophysical Journal 379 (1991): 440-460.

Most relevant References: Previous activities

Part A: Previous workshops
1) http://facultymembers.sbu.ac.ir/movahed/index.php/talks-a-presentations/106-data-modeling-workshop
2) http://facultymembers.sbu.ac.ir/movahed/index.php/talks-a-presentations/131-workshop-on-observational-data-analysis-in-cosmology-kurdistan-1398
3) http://facultymembers.sbu.ac.ir/movahed/index.php/talks-a-presentations/144-school-and-workshop-on-statistical-analysis-of-cosmic-fields-1400
4) http://facultymembers.sbu.ac.ir/movahed/index.php/talks-a-presentations/154-school-and-workshop-on-topological-based-data-analysis-1401

Part B: Some lectures:
1) http://facultymembers.sbu.ac.ir/movahed/index.php/courses/142-optimization-and-computational-approaches-fall-2021
2) http://facultymembers.sbu.ac.ir/movahed/index.php/courses/132-advanced-course-on-computational-physics
3) http://facultymembers.sbu.ac.ir/movahed/index.php/courses/139-stochasic-processes

Part C: Some of my talks:
1) http://facultymembers.sbu.ac.ir/movahed/index.php/talks-a-presentations/102-my-talk-at-cosmology-meeting-ipm-96
2) http://facultymembers.sbu.ac.ir/movahed/index.php/talks-a-presentations/148-my-talk-at-sharif-group-meeting-1401-2022
3) http://facultymembers.sbu.ac.ir/movahed/index.php/talks-a-presentations/145-my-talk-at-conference-on-gravity-and-cosmology-1400
4) http://facultymembers.sbu.ac.ir/movahed/index.php/talks-a-presentations/130-data-science

☆ Data modeling workshop at 1397

☆ Data Modeling workshop at Neyshabour 1398

1 General View (Motivations and Scientific Methodology ...)

X 2 Different Classification of "DM"

☆ Model-Based Analysis

☆ Data-Based Analysis

3 Different types of DM approaches.

☆ Bayesian approach

☆ frequentism approach

4 Various Parts of a typical DM approach

5 The Mathematical Foundation of Bayesian "DM" approach.

6. Bayesian Model averaging (BMA)

7. Analytical algorithm for "DM" [Optimization]

8. Computational algorithm for "DM"

9. Fisher forecast (Fisher Information Matrix)

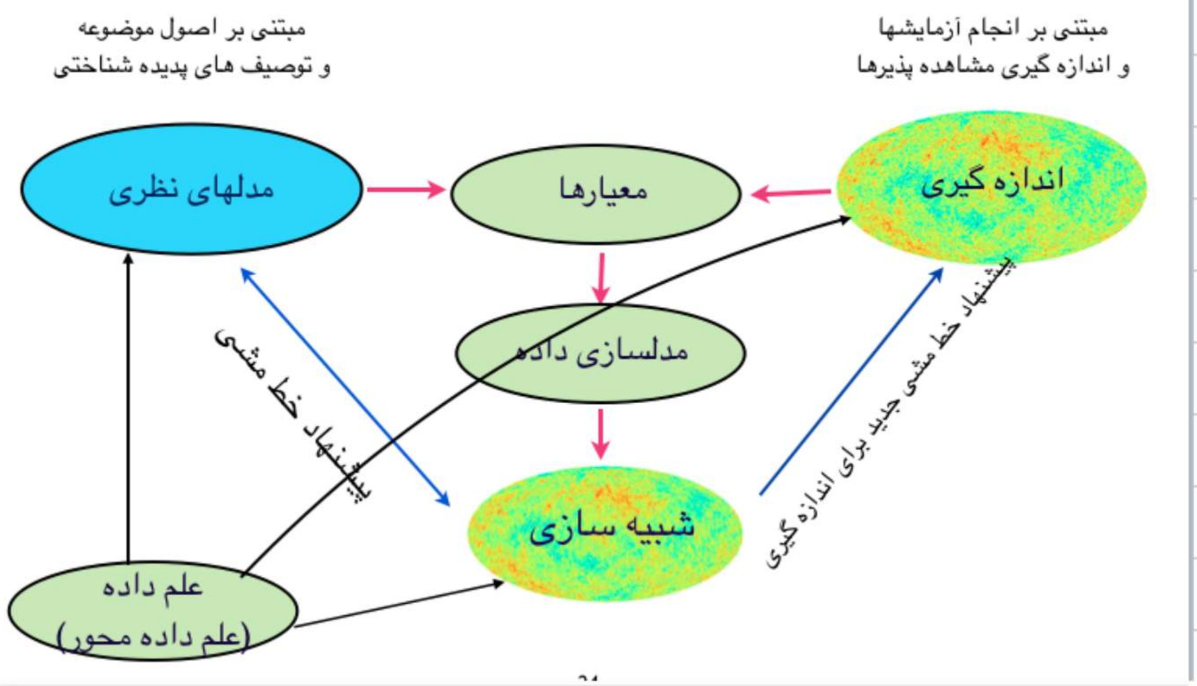
10. Goodness of fit and Confidence Interval for Model's free Parameters.

11. Simulation Based Inference (SBI)

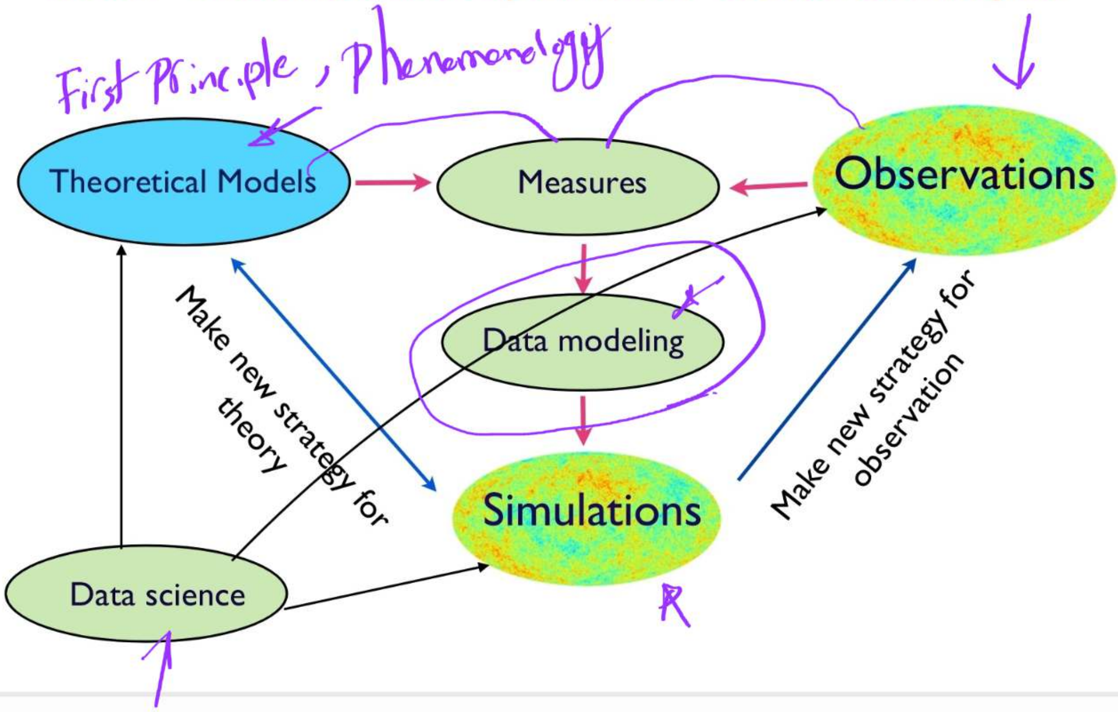
---

1. General View:

چرخه روش شناسی علمی

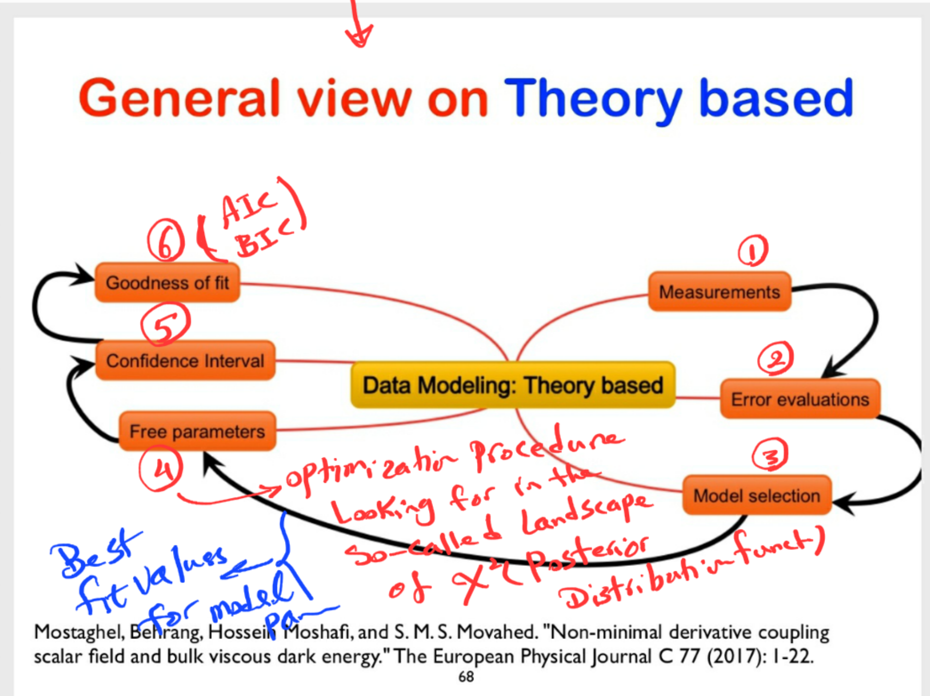


Scientific Methodology
Model - Observation & Experiments - Statistical analysis



② Different Classifications of Data Modeling.

- [ ] Model-Based (Theory-Based approach)
- [ ] Data-Based



Ex 1: Suppose that in an Experimental Set up

We measured Internal Energy of a System.

$$\{E\}$$

from a theoretical point of view we have a

theory for internal Energy of a N-Body System

with intraction $\mathcal{H} = \mathcal{H}_0 + \mathcal{U}$

Free Part $\left(\frac{p^2}{2m}\right)$

Pair intraction

$$\mathcal{U}(r) = \epsilon\left[\left(\frac{\alpha}{r}\right)^a - \left(\frac{\sigma}{r}\right)^b\right] \quad , \quad r = |\vec{r}_i - \vec{r}_j|$$

$\{\theta\} : \{\epsilon, \alpha, a, \sigma, b\}$: Model's free Parameters.

3-D for N-Body system

Un-Weighted TPCF

$$E = \frac{3K_B NT}{2} + \frac{n}{2}\int d^3r\, \underline{u(r)}\, (\tilde{\psi}(r)+1)$$

Theory.

T	E	Error → standard deviation
T_1	E_1	σ_{E_1}
T_2	E_2	σ_{E_2}
T_3	E_3	σ_{E_3}
⋮		
T_p	E_p	σ_{E_p}

$E_{Th}(\{\theta\})$

almost a Well-Defined relation between Model free Parameter and measures from Experiment

We are looking for $\{\theta_{Best}\}$. in such that

$$\boxed{E_{Theory} \longrightarrow E_{experiment}}$$

$\{\theta\}$ — $\boxed{\begin{array}{l}\{\theta_{Best}\} = ? \\ \{\sigma_{\theta}\} = ?\end{array}}$ *

Model (Theory)-Based approach.

Ex2: $T_{\mu\nu} = (\rho + \tilde{p}) u_\mu u_\nu + \tilde{p}\, g_{\mu\nu}$

$\tilde{p} = p + \Pi$ ← Induced by viscosity term

$\nabla_\mu T^\mu_\nu = 0$

Π: (Induced term for Pressure from Bulk viscosity.)

$(\dot{\rho}) + 3H(\rho + p) - 9H^2 \xi = 0$

Viscosity term

$$
\begin{cases}
H^2 = \dfrac{8\pi G}{3} \rho \left(\dfrac{5-3q}{2} \right) - \dfrac{K}{a^2} \\[2ex]
\dfrac{\ddot{a}}{a} = -\dfrac{4\pi}{3} \left(\dfrac{5-3q}{2} \right) G(\rho + 3p)
\end{cases}
$$

$$
\begin{cases}
H^2 = \dfrac{8\pi G}{3} \left(\dfrac{5-3q}{2} \right) (\rho_r + \rho_m + \rho_{DE}) \\
\dot{\rho}_r + 4H\rho_r = 0 \\
\dot{\rho}_m + 3H\rho_m = 0 \\
\dot{\rho}_{DE} + 3H(1 + w_{DE})\rho_{DE} = 9H^2 \xi
\end{cases}
$$

Suppose (Ansatz)

$\xi = \eta_0 H$

$$\frac{H^2}{H_0^2} = \left(\frac{5-3q}{2} \right) \left[\Omega_r (1+z)^4 + \Omega_m (1+z)^3 + \Omega_{DE} \left(1 - \frac{9\eta}{2\sqrt{\Omega_D}} \ln(1+z) \right. \right.$$

SNIa

$$\underbrace{d_L(z)} = \underbrace{(1+z)\int_0^z dz' \frac{1}{H(z)}}$$

$$\{\theta\} = \underbrace{\{\Omega_m, \Omega_r, \Omega_{DE}, \omega_{DE}, \xi, H_0\}}_{\text{Free paramtr}} \leftarrow \boxed{\text{Desired free paramtr}}$$

observatn SNIa.

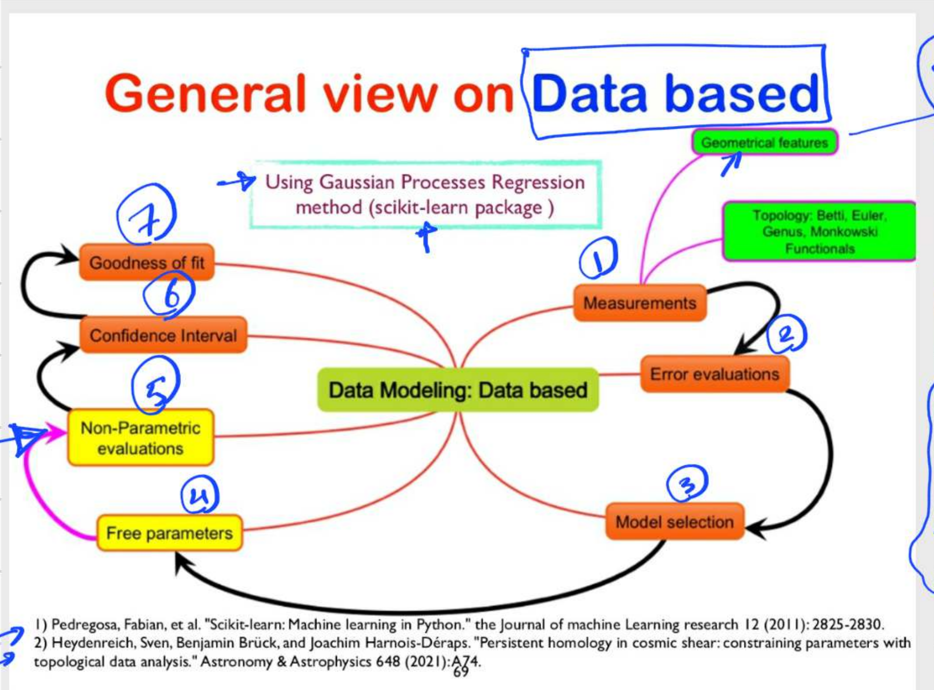
z	Flux	σ
z_1	F_1	σ_{F_1}
z_2	F_2	σ_{F_2}
	$\vdots$	
z_p	F_p	σ_{F_p}

Well Defined Relation

$$\text{Flux} \to d_L \leftarrow$$

$$\{\theta\} = \{\underbrace{\theta_{Best}}\} \to \text{Flux}\Big|_{\text{observat}} \longleftrightarrow F_{\text{theory}}$$

There is No Well Defined Relatn between observations and desired free param



$\langle n_{cr} \rangle$

$\langle n_{pk} \rangle$

$\langle n_{tr} \rangle$

$\langle n_{cmd} \rangle$

$\downarrow$

$\underbrace{\langle n_{cmd}\, n_{cmd} \rangle}$

$\langle \ldots \rangle$

Betti numbers are highly non-linear functions to cosmological parameters

Table A.1. Cosmological parameters of the cosmo-SLICS wCDM simulations. Other fixed parameters are $\Omega_b = 0.0447$ and $n_s = 0.969$.

	Ω_m	S_8	h	w_0	σ_8	Ω_{cdm}
00	0.3282	0.6984	0.6766	-1.2376	0.6677	0.2809
01	0.1019	0.7826	0.7104	-1.6154	1.3428	0.0546
02	0.2536	0.6133	0.6238	-1.7698	0.667	0.2063
03	0.1734	0.7284	0.6584	-0.5223	0.9581	0.1261
04	0.3759	0.8986	0.6034	-0.9741	0.8028	0.3286
05	0.4758	0.7618	0.7459	-1.3046	0.6049	0.4285
06	0.1458	0.768	0.8031	-1.4498	1.1017	0.0985
07	0.3099	0.7861	0.694	-1.8784	0.7734	0.2626
08	0.4815	0.6804	0.6374	-0.7737	0.5371	0.4342
09	0.3425	0.7054	0.8006	-1.501	0.6602	0.2952
10	0.5482	0.6375	0.7645	-1.9127	0.4716	0.5009
11	0.2898	0.7218	0.6505	-0.6649	0.7344	0.2425
12	0.4247	0.7511	0.6819	-1.1986	0.6313	0.3774
13	0.3979	0.8476	0.7833	-1.1088	0.736	0.3506
14	0.1691	0.8618	0.789	-1.6903	1.1479	0.1218
15	0.1255	0.6131	0.7567	-0.9878	0.9479	0.0782
16	0.5148	0.8178	0.6691	-1.3812	0.6243	0.4675
17	0.1928	0.8862	0.6285	-0.8564	1.1055	0.1455
18	0.2784	0.65	0.7151	-1.0673	0.6747	0.2311
19	0.2106	0.8759	0.7388	-0.5667	1.0454	0.1633
20	0.443	0.8356	0.6161	-1.7037	0.6876	0.3957
21	0.4062	0.662	0.8129	-1.9866	0.5689	0.3589
22	0.2294	0.8226	0.7706	-0.8602	0.9407	0.1821
23	0.5095	0.7366	0.6988	-0.7164	0.5652	0.4622
24	0.3652	0.6574	0.7271	-1.5414	0.5958	0.3179
fid	0.2905	0.8231	0.6898	-1.0	0.8364	0.2432

increasing information

Euler characteristic

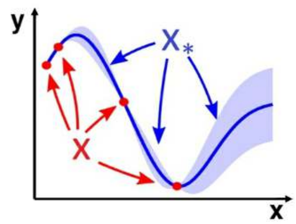
LSS

Betti numbers → Topological Invariant

persistent Betti numbers

Using Gaussian Process Regression method (scikit-learn package)

Pedregosa, Fabian, et al. "Scikit-learn: Machine learning in Python." the Journal of machine Learning research 12 (2011): 2825-2830.



Gaussian Processes Regression

Wang, Jie, and Offroad Robotics. "An Intuitive Tutorial to Gaussian Processes Regression." stat 1050 (2021): 2.