– Sung Hak's 2-moons tutorial

– CaloFlow (2106.05285) example

Simulation Based Inference

aka "ABC approx Bayesian computation" & "likelihood-free inference"

(good ref: Cranmer, Brehmer, Louppe "Frontier of SBI".)

Suppose we want to know $\underbrace{p(D|\theta)}_{\text{likelihood}}$ or $\underbrace{p(\theta|D)}_{\text{posterior}}$

$$\underset{\text{Model}}{\theta} \longrightarrow \underset{\text{"forward model"}}{\boxed{\text{Simulation}}} \longrightarrow \text{Data}$$

- Suppose likelihood is unknown (intractable) or slow
  - maybe D is too high dimensional
  - maybe $p(D|\theta)$ only known for summary statistics (eg χ^2 or Gaussian covariance)
  - example Gaussian Lkl w/ high dim'l D → slow matrix inv.

- Idea of SBI:

  can learn p from samples (θ, D)!

  $\theta \sim p(\theta) \leftarrow$ prior

  $D \sim p(D|\theta) \leftarrow$ simulation from intractable likelihood

- Different techniques / approaches to SBI
  - Most direct "neural posterior or likelihood estimation"

    train flow on samples, learn $p(D|\theta)$ or $p(\theta|D)$

    can sample from it quickly

    speed up posterior inference by many orders of magnitude (weeks/months $\rightarrow$ seconds!)

  - Neural ratio estimation

    From samples (θ, x)

    produce marginal product $\theta, x \sim p(\theta)p(x)$ (just randomly pair)

    Binary classifier:

$$\frac{p(\theta, x)}{p(\theta)p(x)} = \frac{p(\theta|x)}{p(\theta)} = \frac{p(x|\theta)}{p(x)}$$

    Given data $x \rightarrow$ get likelihood & posterior up to normalization!

  - pros & cons to both
  - limitations to SBI — only as good as simulation...

Recent example of SBI: (2310.12209)

pulsar timing arrays ← radio astronomy

Nanograv
12.5 yr dataset
obs ~ 2 weeks
irregularly

~10 pulsars × 50 residuals each

t_i ToA → δt_i residuals

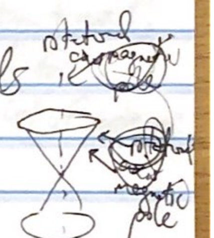
ms pulsar

rotating ≲ 10 ms → very stable periods i.e.

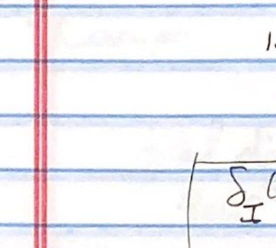
old rapidly spinning neutron stars

maybe X-ray binary

radio pulses from rotating magnetic field



transferring ang. mom.



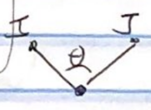
array of ms pulsars → very sensitive to passing gravitational waves

→ "stochastic gravitational wave background"
SGWB

incoherent superposition supermassive BH binaries

$$\langle a_I(f)\, a_J(f')\rangle \sim A_{GW}^2 (f)^{-\gamma_{GW}} \chi_{IJ}\, \delta(f-f')$$

$$\boxed{\delta_I(t) = \int_{-\infty}^{\infty} df\, a_I(f) \cos 2\pi f t + (\sin)}$$



$\gamma_{GW} = 13/3$ predicted

χ_{IJ} = "Hellings Downs curve" fn of θ

$S_0(\delta_I(t_i)\,\delta_J(t_j')) \sim C_{IJ}(t_i - t_j')$

$\downarrow$

"Gaussian process"

Giant gaussian cov

size $N_{pulsar} \times N_{residuals}$ ~ 5000×5000

$t_{iI}\left[C_{IJ}(t_i - t_j)\right]^{-1} t_{jJ}$ → takes long time to invert!

~1 week for one plot of posterior

→ neural posterior inference

→ ~seconds!

Example 2: "optimal cosmological analysis"

2202.05282 Dai & Seljak

Generalizing cosmological inference beyond simple Gaussian CMB

- large scale structure – not Gaussian
- 3d maps (weak lensing, 21cm, ...) — too high dimensional

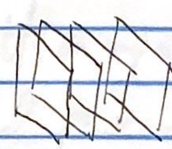
and beyond summary statistics

Goal: learn likelihood

$p(x|y)$

data, cosmo pars

Their example: 4 2D maps of matter overdensities 128^2 pixels each



← Nbody sim FastPM

vary Ω_m, σ_8 uniform prior

other ΛCDM pars fixed to Planck 2015

total dim:

$4 \times 128^2 = 65536$

→ too large for ordinary flow!

→ Impose translation & rotation symmetries
reduce model parameters
want $f: X \to Z$ to be
rot & transl equivariant
then latent space has same symmetries as data

$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \end{pmatrix}$ pixel intensities

translation $i \to i+1$ these should have same PDF
$\vec{x}_i' = \vec{x}_{i-1}$ f map should preserve that automatically

$e^{-\vec{z}^2/2}$ is translation inv't

$\sum_i z_i^2 \quad \sum z_{i-1}^2$

Build f out of pixelwise nonlinearities $\Psi(\vec{x})$
Fourier transform convolution

$$\int d\vec{r}\,' \underbrace{T(\vec{r}-\vec{r}\,')}_{= \int d^3k\, e^{ik|\vec{r}|} T(k)} x(\vec{r}\,')$$

← rot'l inv'a

Ψ & T param by NNs

→ "TRENF"

Impressive results on improving sensitivity to Ω_m, σ_8 vs conventional power-spectrum approach!