Research notes

All here is public

1 Statistics

1.1 Tensions in Cosmology

Define the posterior probability of parameters θ within model $\mathcal{M}$ given data d as (Bayes' theorem)

$$\mathcal{P}(\theta) := P(\theta|d,\mathcal{M}) = \frac{P(\theta|\mathcal{M})P(d|\theta,\mathcal{M})}{P(d|\mathcal{M})} \tag{1}$$

The following notation will simplify this expression

$$\begin{aligned} \Pi(\theta) &= P(\theta|\mathcal{M}) \text{ , Prior probability} \\ \mathcal{L}(\theta) &= P(d|\theta,\mathcal{M}) \text{ , the likelihood} \\ \mathcal{E} &= P(d|\mathcal{M}) \text{ , the evidence} \end{aligned}$$

$$\mathcal{P}(\theta) := P(\theta|d,\mathcal{M}) = \frac{\Pi(\theta)\mathcal{L}}{\mathcal{E}}$$

These are (maybe?) understandable as the prior probability is the parameters given the model, the likelihood as the data given some parameters of the model, and the evidence as the data give a model. From now on, the model $\mathcal{M}$ will be implied.

Consider now two data sets d_1, d_2. They have a joint likelihood

$$\mathcal{L}(\theta) = P(d_1, d_2|\theta)$$

This **quantifies the likelihood of d_1 and d_2 coming from the same set of parameters of a given model.** Denote now $\mathcal{L}_1, \mathcal{P}_1$ as the marginalized likelihood and Posterior over d_2. (**Is there a procedure/algorithm to find the marginalized probability? Seems generally non-trivial**).

Now duplicate the parameter set to θ_1, θ_2 and perscibe a new joint likelihood $\mathcal{L}(\theta_1, \theta_2) = P(d_1, d_2|\theta_1, \theta_2)$. **This is generally a choice, but how can one make a good choice?** The choice however is not unique.

As such, the constraints are imposed:

1. $\mathcal{L}(\theta_1 = \theta, \theta_2 = \theta) = \mathcal{L}(\theta)$, or in plain english, the joint likelihood if $\theta_1 = \theta = \theta_2$ needs to coincide with the likelihood of θ
2. $P(d_1|\theta_1, \theta_2) = P(d_1|\theta_1)$ once marginalized over d_2. Marginalizing the likelihood over one of the data sets removes the dependancy on the corresponding parameter sets.

This ensures the datasets are conditionally independent,

$$P(d_1, d_2|\theta) = P(d_1|\theta)P(d_2|\theta)$$

Proof? Which means we can choose

$$\mathcal{L}(\theta_1, \theta_2) = \mathcal{L}_1(\theta_1)\mathcal{L}_2(\theta_2)$$

If we further assume the prior distribution can be factorized, the joint posterior is then

$$\mathcal{P}(\theta_1, \theta_2) = \mathcal{L}(\theta_1, \theta_2)\Pi(\theta_1)\Pi(\theta_2)$$

Defining $\Delta\theta = \theta_1 - \theta_2$, the parameter difference posterior is given by

$$\mathcal{P}(\Delta\theta) = \int_{V_\Pi} \mathcal{P}(\theta, \theta - \Delta\theta) d\theta$$

1.2 Quantifying Results

Given some probability P of a parameter shift, the following formula can give you the number of standard deviations if the probability shift comes from a gaussian distribution

$$n_\sigma = \sqrt{2}\mathrm{Erf}^{-1}(P)$$

I have a notebook using two unit gaussian priors separated by a distance a. This example can be computed analytically.

$$\begin{aligned}
\mathcal{P}(\Delta\theta) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\theta^2/2} e^{-(\theta-\Delta\theta)^2/2} d\theta \\
&= \frac{1}{2\pi}\cdot\sqrt{\pi}e^{-(\Delta\theta)^2/4} \\
&= \frac{1}{\sqrt{4\pi}}e^{-(\Delta\theta)^2/4}
\end{aligned}$$

The parameter difference posterior is a gaussian with standard deviation $\sqrt{2}$. The separation is fixed by a, hence the shift is $\mathcal{P}(a)$. Hence the shift probability is

$$\Delta = \int_{-a}^{a} e^{-(\Delta\theta)^2/4} d\Delta\theta$$

Lets use the example $a = 2$. Then $n_\sigma = 2/\sqrt{2} = \sqrt{2}$. Using this we can work backwards to find Δ from a z-table to find $\Delta = 0.9207 - 0.0793 = 0.8414$.

2 Computing

2.1 Definitions

Definition 1. *A sequence $X_1, \ldots, X_n$ of random elements is a Markov Chain If the conditional distribution X_{n+1} depends only on X_n. The set in which X_i take values is called the state space of the chain.*

2.2 Parameter Difference Distribution

As discussed above, the parameter difference distribution $\mathcal{P}(\Delta\theta)$ is the convolution of the two posteriors.

2.3 Normalizing Flows

The method of normalizing flows (MAF) implemented here uses Masked Autoencoders (MADE) to construct the flow. Suppose we have an input to the flow x_i. The output of the map is $y_i = \mu(x_{1:i-1}) + \sigma(x_{1:i-1})x_i$. The μ and σ are found using neural networks which recieve masked inputs $x_{1:i-1} = (x_1, \ldots, x_{i-1}, 0, \ldots, 0)$. Since the input only depends on the first $i-1$ inputs, the normalizing flow is *autoregressive* and the Jacobian is triangular.

The implementation in tensorflow uses *bijectors* which implements a local diffeomorphism between a manifold M and a target manifold N (which are our parameter spaces), i.e. $\phi : M \to N$ such that ϕ is differentiable and injective. In tensorflow it has three operations, Forward, Inverse, and log_deg_jacobian, which are exactly the three we want. By constructing a bijector for each masked input, the full normalizing map can be constructed.

`https://si.biostat.washington.edu/sites/default/files/modules/Geyer-Introduction%20to%20markov%20chain%20Monte%20Carlo_0.pdf`

`https://towardsdatascience.com/monte-carlo-markov-chain-mcmc-explained-94e3a6c8de11`

`https://www.sheffield.ac.uk/polopoly_fs/1.60510!/file/MCMC.pdf`

---

3 Fields

3.1 Multifield Dark Energy

So I don't keep having to look at this

Definition 2. *Given a semi-riemannian manifold M with metric g, the christoffel symbols are given by*

$$\nabla_a\partial_b = \Gamma^c_{ab}\partial_c$$

$$\partial_a g_{bc} + \partial_b g_{ca} - \partial_c g_{ab} = 2g_{dc}\Gamma^d_{ab}$$

Consider a metric of the form $\mathrm{diag}(-1, a^2(t))$, so its determinant is $-a^6$. The action is

$$S = \int d^4x\sqrt{-g}\left[\frac{1}{2}M_p^2 R - \frac{1}{2}\gamma_{ab}\partial_\mu\phi^a\partial^\mu\phi^b - V(\phi) + \mathcal{L}_m\right]$$

I only want to describe a homogeneous background, so the field is only a function of time. The action becomes

$$S = \int d^4x\sqrt{-g}\left[\frac{1}{2}M_p^2 R - \frac{1}{2}\gamma_{ab}\dot{\phi}^a\dot{\phi}^b - V(\phi) + \mathcal{L}_m\right]$$

Varying the field gives

$$\delta S = \int d^4x\sqrt{-g}\left[-\frac{1}{2}\partial_a(\gamma_{bc})\dot{\phi}^b\dot{\phi}^c - \frac{1}{2}\gamma_{ab}\frac{d}{dt}(\delta\phi^a)\dot{\phi}^b - \partial_a V\delta\phi^a\right]$$

$$\delta S = -\frac{1}{2}(\delta\phi^a)(\sqrt{-g}\gamma_{ab}\dot{\phi}^b + \int d^4x\sqrt{-g}\left[-\frac{1}{2}\gamma(\nabla_a\partial_b, \partial_c)\dot{\phi}^b, \dot{\phi}^c - \frac{1}{2}\gamma(\partial_b, \nabla_a\partial_c)\dot{\phi}^b\dot{\phi}^c + 3\frac{\dot{a}}{a}\gamma_{ab}\delta\phi^a\dot{\phi}^b + \gamma_{ab}\delta\phi^a\ddot{\phi}^b - V_a\delta\phi^a\right]$$

$$\delta S = -\int d^4x\sqrt{-g}\left[\frac{1}{2}\Gamma^d_{ab}\gamma_{dc}\dot{\phi}^b\dot{\phi}^c + \frac{1}{2}\Gamma^d_{ac}\gamma_{bd}\dot{\phi}^b\dot{\phi}^c + 3H\gamma_{ab}\dot{\phi}^b + \gamma_{ab}\ddot{\phi}^b + V_a\right]$$

Multiply everything by γ^{aa}. Now lets do some reshuffling of the indices.

$$\gamma^{aa}\Gamma^d_{ab}\gamma_{dc}\dot{\phi}^b\dot{\phi}^c$$

$$a \leftrightarrow d$$

$$\gamma^{ad}\Gamma^a_{db}\gamma_{ac}\dot{\phi}^b\dot{\phi}^c$$

$$b \leftrightarrow d$$

$$\gamma^{ad}\Gamma^a_{bd}\gamma_{ac}\dot{\phi}^b\dot{\phi}^c$$

$$\gamma^{ab}\Gamma^a_{bd}\gamma_{ac}\dot{\phi}^d\dot{\phi}^c$$

$$c \leftrightarrow d$$

$$\gamma^{ab}\Gamma^a_{bc}\gamma_{ad}\dot{\phi}^d\dot{\phi}^c = \Gamma^a_{bc}\dot{\phi}^b\dot{\phi}^c$$

and

$$\gamma^{aa}\Gamma^d_{ac}\gamma_{bd}\dot{\phi}^b\dot{\phi}^c$$

$$a \leftrightarrow d$$

$$\gamma^{ad}\Gamma^a_{dc}\gamma_{ba}\dot{\phi}^b\dot{\phi}^c$$

$$b \leftrightarrow d$$

$$\gamma^{ab}\Gamma^a_{bc}\gamma_{da}\dot{\phi}^d\dot{\phi}^c = \Gamma^a_{bc}\dot{\phi}^b\dot{\phi}^c$$

Thus the equation of motion is found,

$$\ddot{\phi}^a + \Gamma^a_{bc}\dot{\phi}^b\dot{\phi}^c + 3H\dot{\phi}^a + V^a = 0$$

$$D_t\dot{\phi}^a + 3H\dot{\phi}^a + V^a = 0$$

The riemann curvature is

$$R^\alpha_{\beta\gamma\delta} = \partial_\beta\Gamma^\alpha_{\gamma\delta} - \partial_\gamma\Gamma^\alpha_{\beta\delta} + \Gamma^\sigma_{\gamma\delta}\Gamma^\alpha_{\beta\sigma} - \Gamma^\sigma_{\beta\delta}\Gamma^\alpha_{\gamma\sigma}$$

So the Ricci curvature tensor is

$$R_{\beta\delta} = R^\alpha_{\beta\alpha\delta} = \partial_\beta\Gamma^\alpha_{\alpha\delta} - \partial_\alpha\Gamma^\alpha_{\beta\delta} + \Gamma^\sigma_{\alpha\delta}\Gamma^\alpha_{\beta\sigma} - \Gamma^\sigma_{\beta\delta}\Gamma^\alpha_{\alpha\sigma}$$

And the Ricci scalar curvature is

$$R = R^\beta_\beta = g^{\beta\delta}R_{\beta\delta} = g^{\beta\delta}\partial_\beta\Gamma^\alpha_{\alpha\delta} - g^{\beta\delta}\partial_\alpha\Gamma^\alpha_{\beta\delta} + g^{\beta\delta}\Gamma^\sigma_{\alpha\delta}\Gamma^\alpha_{\beta\sigma} - g^{\beta\delta}\Gamma^\sigma_{\beta\delta}\Gamma^\alpha_{\alpha\sigma}$$

Since the fields only depend on time, we only need to consider the temporal component of the Ricci curvature tensor, so

$$R_{tt} = \partial_t\Gamma^\alpha_{\alpha t} - \partial_\alpha\Gamma^\alpha_{tt} + \Gamma^\sigma_{\alpha t}\Gamma^\alpha_{t\sigma} - \Gamma^\sigma_{tt}\Gamma^\alpha_{\alpha\sigma}$$

From the definition of the christoffel symbols, and noting $g = g(t)$ and g is diagonal, we have $\partial_t g_{aa} = 2g_{ba}\Gamma^b_{ta}$, thus the second term is necessarily 0 since $\partial_t g_{tt} = 0$. The first term is non-zero for $\alpha = 1, 2, 3$, in which it equals

$$\frac{1}{2}\partial_t(1/a^2\partial_t a^2) = \partial_t(\dot{a}/a) = -3\ddot{a}/a + 3\dot{a}^2/a^2$$

The third term is non-zero for $\alpha = \sigma$, In which case we get

$$\frac{3}{4a^4}(\partial_t(a^2))^2 = -3\dot{a}^2/a^2$$

The last term is vanishes since the temporal component of the metric is constant. Thus we find

$$R_{tt} = 3\frac{\ddot{a}}{a}$$

Now, if we consider spatial components of the Ricci curvature, we find that the first term now vanishes, the second term is equal to $-\ddot{a}/a - \dot{a}^2/a^2$, the thrid term is 0, and the last term is $-\dot{a}^2/a^2$, so the spacial components are

$$R_{xx} = -\ddot{a}a - 2\dot{a}^2$$

Thus the Ricci curvature scalar is

$$6\ddot{a}/a + 6\dot{a}^2/a^2$$

Hence the einstein equation gives

$$-3H^2 =$$

I need to write out the indices better if I want to get the signs correct.

$$\partial_a g_{bc} + \partial_b g_{ca} - \partial_c g_{ab} = 2g_{dc}\Gamma^d_{ab}$$

Considering the way this acts on the coordinate vector field ∂_α,

$$R\partial_\alpha = g^{\beta\delta}\partial_\beta\Gamma^\alpha_{\alpha\delta}\partial_\alpha - g^{\beta\delta}\partial_\alpha\Gamma^\alpha_{\beta\delta}\partial_\alpha + g^{\beta\delta}\Gamma^\sigma_{\alpha\delta}\Gamma^\alpha_{\beta\sigma}\partial_\alpha - g^{\beta\delta}\Gamma^\sigma_{\beta\delta}\Gamma^\alpha_{\alpha\sigma}\partial_\alpha$$

$$R\partial_\alpha = g^{\beta\delta}\partial_\beta\nabla_\alpha\partial_\delta - g^{\beta\delta}\partial_\alpha\nabla_\beta\partial_\delta + g^{\beta\delta}\Gamma^\sigma_{\alpha\delta}\nabla_\beta\partial_\sigma - g^{\beta\delta}\Gamma^\sigma_{\beta\delta}\nabla_\alpha\partial_\sigma$$