

Inferência Bayesiana de parâmetros Cosmológicos

GABRIEL SILVA COSTA, N°USP: 11298601

Instituto de Física, Universidade de São Paulo - SP, Brasil

Resumo

Neste trabalho, são aplicados métodos bayesianos a dados observacionais simulados de redshifts de supernovas do tipo Ia e suas respectivas magnitudes, utilizando um modelo simplificado de brinquedo de SNIa, a fim de estimar os valores de parâmetros cosmológicos que descrevem o modelo Λ CDM

I. INTRODUÇÃO

O Λ CDM é o modelo atual para descrever o universo com base na teoria da relatividade geral. Um dos testes desse modelo é comparar as distâncias de luminosidade de supernovas do tipo Ia, que são consideradas velas-padrão, com os seus respectivos redshifts. A distância de luminosidade depende de parâmetros cosmológicos como a densidade da matéria, da energia escura e a constante de Hubble, os quais podem ser estimados através da inferência bayesiana a partir dos dados observacionais, usando uma verossimilhança e priores. O resultado é uma distribuição posterior que reflete o grau de plausibilidade dos parâmetros após observar os dados.

Foram utilizados os dados SNIa SimulatedData (arquivo de dados fornecidos no artigo (1), no problema (ix), pg. 47), que contém observações de redshift e magnitude aparente simuladas do modelo simplificado de brinquedo de SNIa, de $N = 300$ SNIa, cuja as características são projetadas para imitar conjuntos de dados atualmente disponíveis (um plot do conjunto de dados é mostrado na Figura 1). Os parâmetros cosmológicos que estamos interessados são: $\{\Omega_m, \Omega_\Lambda, h\}$, onde Ω_m é a densidade da matéria (em unidades da densidade de energia crítica) e Ω_Λ é a densidade de energia escura, assumida aqui como sendo na forma de uma constante cosmológica, isto é, $w = -1$ em todos os redshifts. A seguir, foi fixado $h = 0.72$, onde a constante de Hubble hoje é dada por $H_0 = 100h \text{ km/s/Mpc}$, pois o valor de H_0 é degenerado com a magnitude absoluta (desconhecida) da SNIas, M . Em uma cosmologia de FRW definida pelos parâmetros Ω_m, Ω_Λ e h , o módulo de distância μ (isto é, a diferença entre as magnitudes aparente e absoluta, $\mu = m - M$) para uma SNIa localizada no redshift z é dado por:

$$\mu(z, \Omega_m, \Omega_\Lambda, h) = \eta + 5 \log d_L(z, \Omega_m, \Omega_\Lambda) \quad (1)$$

onde

$$\eta = -5 \log \frac{100h}{c} + 25$$

e

$$d_L(z, \Omega_m, \Omega_\Lambda) = \frac{(1+z)}{\sqrt{|1 - \Omega_m - \Omega_\Lambda|}} \text{sinn} \left\{ \sqrt{|1 - \Omega_m - \Omega_\Lambda|} \int_0^z dz' \left[(1+z')^3 \Omega_m + \Omega_\Lambda + (1+z')^2 (1 - \Omega_m - \Omega_\Lambda) \right]^{-\frac{1}{2}} \right\}$$

em que

$$\text{sinn}(x) = \begin{cases} x & \text{para } (\Omega_m + \Omega_\Lambda = 0) \\ \sin(x) & \text{para } (\Omega_m + \Omega_\Lambda > 1) \\ \sinh(x) & \text{para } (\Omega_m + \Omega_\Lambda < 1) \end{cases}$$

Assume-se que de cada SNIa na amostra, temos uma medida de seu módulo de distância com ruído gaussiano, ou seja, que a função de verossimilhança para cada SNIa i ($i = 1, \dots, N$) é da forma:

$$L_i(z, \Omega_m, \Omega_\Lambda, M) = -\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2} \frac{(\hat{\mu}_i - \mu_i(z, \Omega_m, \Omega_\Lambda, h))^2}{\sigma^2}\right)$$

(O erro observacional é o mesmo para todas as SN, $\sigma_i = \sigma = 0,4 \text{ mag}$ para $i = 1, \dots, N$.)

O módulo de distância observado é dado por:

$$\hat{\mu}_i = m_i - M$$

onde $\hat{\mu}_i$ é a magnitude aparente observada e M é a magnitude intrínseca do SNIa.

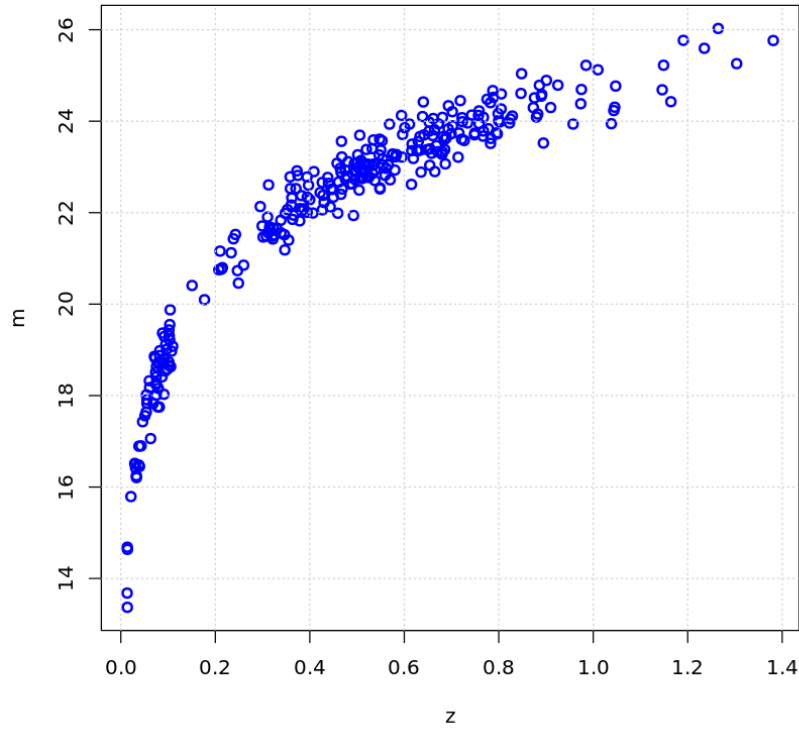


Figura 1: plot dos dados observacionais simulados (y =magnitude aparente, x =redshift).

II. METODOLOGIA

A verossimelhança dos dados é:

$$L = \prod_i L_i = \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp \left(\sum_i -\frac{1}{2} \frac{(\hat{\mu}_i - \mu_i)^2}{\sigma^2} \right)$$

Trabalhando com o logarítmo dessa expressão, temos:

$$\ln L = n \ln \left(\frac{1}{\sqrt{2\pi}\sigma} \right) - \frac{1}{2\sigma^2} \sum_i (\hat{\mu}_i - \mu_i)^2$$

Usando o teorema de Bayes (veja (2)):

$$P(\Omega_m, \Omega_\Lambda | \text{Dados}) \propto P(\text{Dados} | \Omega_m, \Omega_\Lambda) P(\Omega_m, \Omega_\Lambda)$$

em que $P(\text{Dados} | \Omega_m, \Omega_\Lambda)$ é a verossimelhança e $P(\Omega_m, \Omega_\Lambda) = P(\Omega_m) \cdot P(\Omega_\Lambda)$ o prior.

Na forma logarítima, temos:

$$\ln P(\Omega_m, \Omega_\Lambda | \text{Dados}) = \ln P(\text{Dados} | \Omega_m, \Omega_\Lambda) + \ln P(\Omega_m) + \ln P(\Omega_\Lambda)$$

Dessa forma, a probabilidade posterior pode ser determinada por meio de MCMC.

III. RESULTADOS

Fixando $M = -19.3 \text{ mag}$, foi determinada a probabilidade posterior de $(\Omega_m, \Omega_\Lambda)$ por meio de simulações de MCMC usando priors uniformes, através da qual foi obtido um plot dos contornos das regiões de 68% e 95% de credibilidade, como pode ser visto na Figura 2.

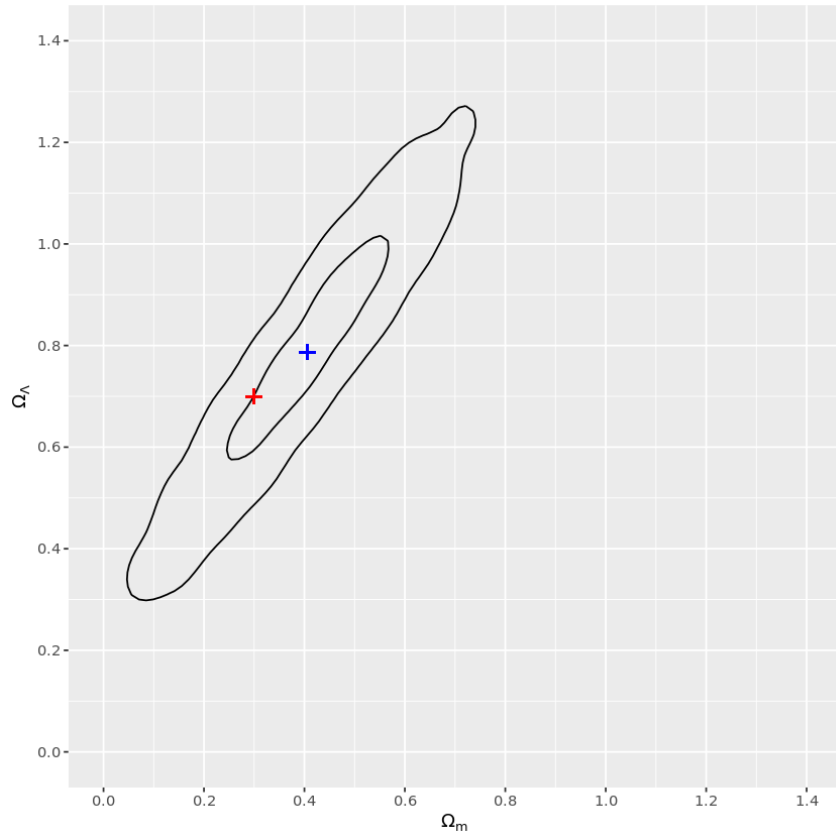


Figura 2: plot dos contornos das regiões de 68% e 95% de credibilidade (A cruz em vermelha é o ponto verdadeiro e, em azul é a média da distribuição obtida, $[\Omega_m = 0.405, \Omega_\Lambda = 0.787]$)

Em seguida, foi adicionada a quantidade σ (o erro observacional) ao conjunto de parâmetros desconhecidos, o qual foi estimado a partir dos dados utilizando-se o prior $P(\sigma) \propto 1/\sigma$, o qual é apropriado dado que σ é um “parâmetro escalar”, veja (1), pg. 48 para uma justificativa). Portanto, realizando-se novamente simulações de MCMC, pode-se estimar a distribuição para σ , como pode ser vista no histograma na Figura 3.

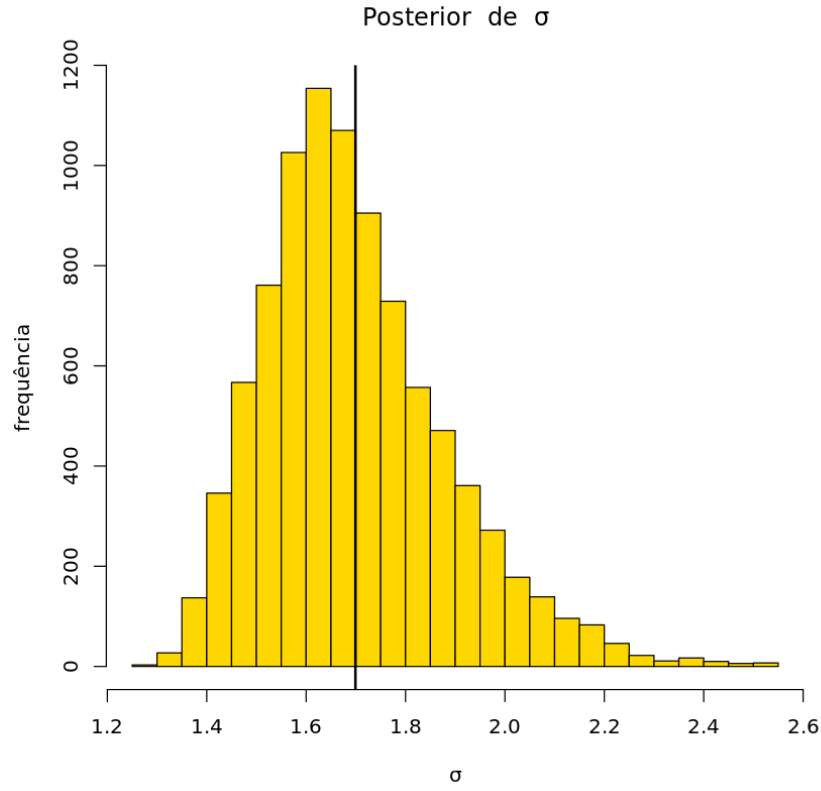


Figura 3: histograma do posterior de σ , média = 1.699

E por fim, utilizando dados inferidos da CBM (veja (3), Fig. 20), em que notou-se uma restrição efetiva na combinação de Ω_m e Ω_Λ , dada por:

$$1.41\Omega_\Lambda + \Omega_m = 1.30 \pm 0.04$$

A qual foi adicionada à verossimelhança, o que implicou na distribuição do posterior dos parâmetros que pode ser observada na Figura 4.

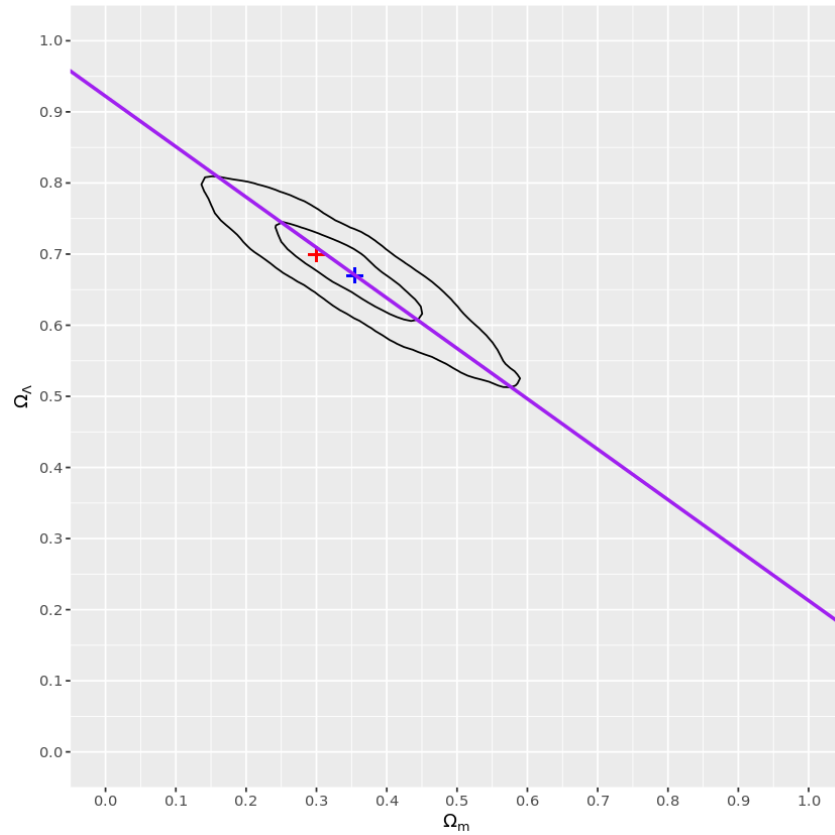


Figura 4: plot dos contornos das regiões de 68% e 95% de credibilidade (A cruz em vermelha é o ponto verdadeiro e, em azul é a média da distribuição obtido [$\Omega_m = 0.405, \Omega_\Lambda = 0.787$]) adicionado ao plot 1D da restrição (reta em roxo)

IV. CONCLUSÃO

Nota-se como métodos bayesianos são ferramentas poderosas de análise e inferência de parâmetros cosmológicos, haja visto que, a partir dos dados fornecidos, foi possível a determinação estatísticas dos parâmetros cosmológicos Ω_m , Ω_Λ e σ com certa acurácia e precisão, uma vez que os valores verdadeiros estavam contidos nas regiões de 68% e 95% de credibilidade, como pode ser averiguado nas simulações dos posteriores, representados nas figuras 2, 3 e 4.

A.

APENDICE: CÓDIGO USADO

O código usado pode ser acessado por meio do link: <https://datalore.jetbrains.com/notebook/qeduIPAivhqfZhio3VN1U1/wN5cK6kGCuguQWdrYcAA3I>

REFERÊNCIAS

- [1] <https://arxiv.org/abs/1701.01467>
- [2] http://www.astro.iag.usp.br/~laerte/aga0505_23/R7.html
- [3] <https://arxiv.org/abs/astro-ph/0603449>