Metodologia bayesowska kosmologii współczesnej

Marek Szydłowski   Aleksandra Kurek

Obserwatorium Astronomiczne Uniwersytetu Jagiellońskiego

XXXIV Zjazd PTA
Kraków, 15 września 2009

Modele kosmologiczne akcelerującego Wszechświata

1. Modele z ciemną energią
   - Wszechświat jednorodny i izotropowy (metryka FRW)
   - ewolucja opisana równaniami Friedmanna
   - dodatkowy składnik łamiący silny warunek energetyczny $\rho + 3p > 0$

★ ΛCDM ($w = -1$)

★ quintessence ($w(z) > -1$), phantom ($w(z) < -1$)

Modele kosmologiczne akcelerującego Wszechświata

1. Modele z ciemną energią
2. Modele ze zmodyfikowaną teorią grawitacji
   - Wszechświat jednorodny i izotropowy (metryka FRW)

★ Brane models: DGP; Sahni-Shtanov

★ Cardassian model: $3H^2 = \rho + B\rho^n$

Modele kosmologiczne akcelerującego Wszechświata

1. Modele z ciemną energią
2. Modele ze zmodyfikowaną teorią grawitacji
3. Modele niejednorodne
   - brak założenia o jednorodności Wszechświata

   ★ wpływ na propagację światła
   ♢ sferycznie symetryczne rozwiązania równań Einsteina z pyłem (Lemaitre-Tolman-Bondi (LTB) model)
   ♢ znajdujemy się w obszarze o obniżonej gęstości 'Hubble bubble' (lokalnie $\Omega_m \in [0.15, 0.35]$, globalnie $\Omega_m = 1$)
   ★ ogólna (uśredniona) dynamika Wszechświata może się różnić od dynamiki wynikającej z modelu standardowego, $G(< g >) =< G(g) >$ tylko gdy niejednorodności mogą być modelowane jako liniowa perturbacja modelu FRW - tylko we wczesnym Wszechświecie (backreaction models)

Metodologia bayesowska

- Prawdopodobieństwo jako stopień racjonalnego zaufania w świetle informacji, które posiadamy
(nie bazuje na asymptotycznych własnościach estymatorów; rezygnuje z pojęcia zmiennej losowej, odpowiada na pytania, których nie można zadać w ujęciu klasycznym; umożliwia uwzględnienie wszystkich informacji, które posiadamy)
- Estymacja parametrów modelu, porównanie modeli

Bayesowska metoda estymacji parametrów modelu

Wnioski dotyczące wartości parametrów modelu $\bar{\theta}$ na podstawie rozkładu prawdopodobieństwa wynikowego

$$P(\bar{\theta}|D, M_i) = \frac{P(D|\bar{\theta}, M_i)P(\bar{\theta}|M_i)}{P(D|M_i)}$$

$P(D|\bar{\theta}, M_i) \equiv L_i$ – likelihood, prawdopodobieństwo otrzymania danych D, przy założeniu, że model M_i jest prawdziwy
$P(\bar{\theta}|M_i)$ – rozkład prawdopodobieństwa a priori na parametry modelu
$P(D|M_i)$ – stała normalizacyjna, istotna w bayesowskiej metodzie porównywania modeli (tzw. ewidencja)

Bayesowska metoda estymacji parametrów

- Posterior pdf dla konkretnego parametru otrzymujemy poprzez marginalizację: $P(\phi|D, M_i) = \int P(\bar{\theta}|D, M_i) d\bar{\theta}'$, gdzie $\bar{\theta} = (\phi, \bar{\theta}')$
- Wartość parametru – wartość oczekiwana rozkładu, moda rozkładu (najbardziej prawdopodobna wartość)
- Przedział ufności – przedział zawierający 68%, czy 95% prawdopodobieństwa wynikowego (nie ma takiego samego znaczenia jak klasyczne 1σ, czy 2σ)

Bayesowska metoda estymacji parametrów

Bayesowska metodologia krytykowana ze względu na obecność priorów:

- priory są subiektywne; teoria nie specyfikuje skąd mają pochodzić priory; różne priory prowadzą do różnych posteriorów, a więc różnych wniosków

ale

- priory odzwierciedlają naszą wiedzę o modelu (powinny zawierać nasze wcześniejsze doświadczenia: pochodzące z wcześniejszych eksperymentów, czy też z teoretycznych przewidywań); powinny bazować na zrozumieniu problemu; im lepsze dane tym słabsza zależność wyników od priorów; różny stan wiedzy początkowej powinien prowadzić do różnych wniosków

Bayesowska metoda porównywania modeli

Najlepszy model – największa wartość prawdopodobieństwa wynikowego

$$P(M_i|D) = \frac{P(D|M_i)P(M_i)}{P(D)}$$

P(**D**) - stała normalizacyjna

$$\sum_{i=1}^{K} P(M_i|D) = 1 \longrightarrow P(D) = \sum_{i=1}^{K} P(D|M_i)P(M_i)$$

wnioski oparte na prawdopodobieństwie wynikowym zależą od rozważanego zbioru modeli

Bayesowska metoda porównywania modeli

Najlepszy model – największa wartość prawdopodobieństwa wynikowego

$$P(M_i|D) = \frac{P(D|M_i)P(M_i)}{P(D)}$$

$\mathbf{P(M_i)}$ prawdopodobieństwo początkowe przypisane i-temu modelowi

- wartość zależy od wcześniejszych informacji
- POSTULAT BAYESA: jeśli nic nie jest nam wiadome a priori o poszczególnych możliwych hipotezach prawdopodobieństwa tych hipotez powinniśmy przyjąć równe $P(M_i) = 1/K$, K - liczba rozważanych modeli

Bayesowska metoda porównywania modeli

Najlepszy model – największa wartość prawdopodobieństwa wynikowego

$$P(M_i|D) = \frac{P(D|M_i)P(M_i)}{P(D)}$$

$\mathbf{P(D|M_i)}$ – marginal likelihood (ewidencja)

$$P(D|M_i) = \int P(D|\bar{\theta}_i, M_i)P(\bar{\theta}_i|M_i)d\bar{\theta}_i \equiv E_i$$

$P(D|\bar{\theta}_i, M_i) \equiv L_i$ – likelihood i-tego modelu
$P(\bar{\theta}_i|M_i)$ – rozkład prawdopodobieństwa a priori dla parametrów i-tego modelu
$\bar{\theta}_i$ – wektor parametrów i-tego modelu

Bayesowska metoda porównywania modeli

$$O_{ij} = \frac{P(M_i|D)}{P(M_j|D)} = \frac{P(M_i)}{P(M_j)} \frac{E_i}{E_j}$$

$$P(M_i) = P(M_j) \mapsto O_{ij} = \frac{E_i}{E_j} \equiv B_{ij} \text{ (czynnik Bayesa)}$$

$\ln \mathbf{B_{ij}}$	**ewidencja na korzyść modelu *i* wzgl modelu *j***
$[0, 1]$	niekonkluzywna
$[1, 2.5]$	słaba
$[2.5, 5]$	umiarkowana
> 5	silna

Bayesowska metoda porównywania modeli

wartość ewidencji

- przybliżenie $-2\ln E_i$
  $BIC_i = -2\ln \mathcal{L}_i + d\ln N$
  $\mathcal{L}_i$ – max likelihood; d – liczba parametrów modelu; N – liczba danych
- obliczenia numeryczne (np. nested sampling algorytm)
  zaleta: parametry słabo wyznaczone nie są uwzględnione

Porównanie

Dane

- SNIa $N = 192$ (Riess et al. 2007; Wood-Vasey et al. 2007; Davis et al. 2007)

$$\mathrm{L} \propto \exp\left[-\frac{1}{2}\left(\sum_{i=1}^{N}\frac{(\mu_i^{\mathrm{theor}}-\mu_i^{\mathrm{obs}})^2}{\sigma_i^2}\right)\right]$$

$$\mu_i^{\mathrm{obs}} = m_i - M,\ \mu_i^{\mathrm{theor}} = 5\log_{10} D_{Li} - 5\log_{10} H_0 + 25,$$

$$D_{Li} = H_0 d_{Li} = H_0(1+z_i)c\int_0^{z_i}\frac{dz'}{H(z')}\ \text{(przy założeniu } \Omega_k = 0)$$

- CMB R (Spergel et al. 2006; Wang Mukherjee 2006)

$$\mathrm{L} \propto \exp\left[-\frac{(R^{\mathrm{theor}}-R^{\mathrm{obs}})^2}{2\sigma_R^2}\right]$$

$$R^{\mathrm{theor}} = \sqrt{\Omega_{\mathrm{m},0}}\int_0^{z_{dec}}\frac{H_0}{H(z)}dz,\ R^{\mathrm{obs}} = 1.70 \pm 0.03 \text{ dla } z_{\mathrm{dec}} = 1089$$

- BAO A (Eisenstein et al. 2005)

$$\mathrm{L} \propto \exp\left[-\frac{(A^{\mathrm{theor}}-A^{\mathrm{obs}})^2}{2\sigma_A^2}\right]$$

$$A^{\mathrm{theor}} = \sqrt{\Omega_{\mathrm{m},0}}\left(\frac{H(z)}{H_0}\right)^{-\frac{1}{3}}\left[\frac{1}{z_A}\int_0^{z_A}\frac{H_0}{H(z)}\right]^{\frac{2}{3}},\ A^{obs} = 0.469 \pm 0.017 \text{ dla } z_A = 0.35$$

Założone równe prawdopodobieństwa początkowe dla każdego modelu; porównanie na podstawie wielkości BIC

Porównanie

1. ΛCDM model

$$H^2 = H_0^2\{\Omega_{\mathrm{m},0}(1+z)^3 + (1-\Omega_{\mathrm{m},0})\}$$

2. model with generalized Chaplygin gas

$$H^2 = H_0^2\left\{\Omega_{\mathrm{m},0}(1+z)^3 + (1-\Omega_{\mathrm{m},0})[A_S + (1-A_S)(1+z)^{3(1+\alpha)}]^{\frac{1}{1+\alpha}}\right\}$$

3. model with phantom dark energy

$$H^2 = H_0^2\{\Omega_{\mathrm{m},0}(1+z)^3 + (1-\Omega_{\mathrm{m},0})(1+z)^{3(1+w_X)}\}$$

4. model with dynamical E.Q.S

$$H^2 = H_0^2\left\{\Omega_{\mathrm{m},0}(1+z)^3 + (1-\Omega_{\mathrm{m},0})(1+z)^{3(w_0+w_1+1)}\exp[-\frac{3w_1 z}{1+z}]\right\}$$

5. quintessence model

$$H^2 = H_0^2\left\{\Omega_{\mathrm{m},0}(1+z)^3 + (1-\Omega_{\mathrm{m},0})(1+z)^{3(1+w_0(1+z)^{-\alpha})}\right\}$$

Porównanie

6. DGP model

$$H^2 = H_0^2 \left\{ \left[\sqrt{\Omega_{\mathrm{m},0}(1+z)^3 + \Omega_{rc,0}} + \sqrt{\Omega_{rc,0}} \right]^2 \right\}$$

$$\Omega_{rc,0} = \frac{(1-\Omega_{\mathrm{m},0})^2}{4}$$

7. BΛCDM model

$$H^2 = H_0^2 \left\{ \Omega_{\mathrm{m},0}(1+z)^3 - \Omega_{n,0}(1+z)^n + 1 - \Omega_{\mathrm{m},0} + \Omega_{n,0} \right\}$$

8. interacting model with Λ

$$H^2 = H_0^2 \{ \Omega_{\mathrm{m},0}(1+z)^3 + \Omega_{\mathrm{int},0}(1+z)^n + 1 - \Omega_{\mathrm{m},0} - \Omega_{\mathrm{int},0} \}$$

9. Cardassian model

$$H^2 = H_0^2 \left\{ \Omega_{r,0}(1+z)^4 + \Omega_{\mathrm{m},0}(1+z)^4 \left[\frac{1}{1+z} + (1+z)^{-4+4n} \left(\frac{1-\Omega_{r,0}-\Omega_{\mathrm{m},0}}{\Omega_{\mathrm{m},0}} \right) \left(\frac{\frac{1}{1+z} + \frac{\Omega_{r,0}}{\Omega_{\mathrm{m},0}}}{1+\frac{\Omega_{r,0}}{\Omega_{\mathrm{m},0}}} \right)^n \right] \right\}$$

$$\Omega_{r,0} = 10^{-4}$$

10. Sahni-Shtanov brane I model

$$H^2 = H_0^2 \left\{ \Omega_{\mathrm{m},0}(1+z)^3 + \Omega_{\sigma,0} + 2\Omega_{l,0} - 2\sqrt{\Omega_{l,0}}\sqrt{\Omega_{\mathrm{m},0}(1+z)^3 + \Omega_{\sigma,0} + \Omega_{l,0} + \Omega_{\Lambda b,0}} \right\}$$

$$\Omega_{\sigma,0} = 1 - \Omega_{\mathrm{m},0} + 2\sqrt{\Omega_{l,0}}\sqrt{1 + \Omega_{\Lambda b,0}}$$

Porównanie

model	prior	posterior	model	prior	posterior
1	0.20	**0.84**	6	0.20	0.07
2	0.20	0.02	7	0.20	0.03
3	0.20	0.06	8	0.20	0.13
4	0.20	0.04	9	0.20	**0.74**
5	0.20	0.04	10	0.20	0.03

Porównanie

model	prior	posterior
1	0.10	**0.74**
2	0.10	0.02
3	0.10	0.05
4	0.10	0.04
5	0.10	0.03
6	0.10	0.01
7	0.10	0.005
8	0.10	0.01
9	0.10	0.09
10	0.10	0.005

Porównanie

1. Oscillating DE model $w_X(z) = -1 + (1+z)^3 \left\{ C_1 \cos(\ln(1+z)) + C_2 \sin(\ln(1+z)) \right\}$

$$H = H_0 \sqrt{\Omega_{\Lambda,0} \exp(-B) \exp\left((1+z)^3 \left[A \sin(\ln(1+z)) + B \cos(\ln(1+z)) \right] \right) + \Omega_{m,0}(1+z)^3 + \Omega_{r,0}(1+z)^4},$$

$\Omega_{r,0} \simeq 0.5 * 10^{-4}$, $\Omega_{\Lambda,0} = 1 - \Omega_{m,0} - \Omega_{r,0}$, $A = 0.3(C1 + 3C2)$, $B = 0.3(3C1 \hat{a} C2)$

2. Oscillating DE 1 model $C_2 = 0 \Rightarrow w_X(z) = -1 + C_1(1+z)^3 \cos(\ln(1+z))$

$$H(z) = H_0 \sqrt{\Omega_{\Lambda,0} \exp(-0.9C_1) \exp\left(0.3C_1(1+z)^3 \left[3\cos(\ln(1+z)) + \sin(\ln(1+z)) \right] \right) + \Omega_{m,0}(1+z)^3 + \Omega_{r,0}(1+z)^4},$$

$\Omega_{r,0} \simeq 0.5 * 10^{-4}$, $\Omega_{\Lambda,0} = 1 - \Omega_{m,0} - \Omega_{r,0}$.

3. Oscillating DE 2 model $C_1 = 0 \Rightarrow w_X(z) = -1 + C_2(1+z)^3 \sin(\ln(1+z))$

$$H(z) = H_0 \sqrt{\Omega_{\Lambda,0} \exp(0.3C_2) \exp\left(-0.3C_2(1+z)^3 \left[\cos(\ln(1+z)) - 3\sin(\ln(1+z)) \right] \right) + \Omega_{m,0}(1+z)^3 + \Omega_{r,0}(1+z)^4},$$

$\Omega_{r,0} \simeq 0.5 * 10^{-4}$, $\Omega_{\Lambda,0} = 1 - \Omega_{m,0} - \Omega_{r,0}$

Porównanie

- Dane
  - SNIa $N = 192$ (Riess et al. 2007; Wood-Vasey et al. 2007; Davis et al. 2007)
  - CMB R (Spergel et al. 2006; Wang Mukherjee 2006)
  - BAO A (Eisenstein et al. 2005)
- założone równe prawdopodobieństwa początkowe dla każdego modelu
- porównanie na podstawie wielkości *BIC*

Porównanie

Model	$2 \ln B$
ΛCDM	0
Osc DE	5.81
Osc DE 1	3.37
Osc DE 2	3.37
Linear parametrization	5.76

Porównanie

1. backreaction model

$$H = H_0\sqrt{\Omega_m(1+z)^3 + (1-\Omega_m)(1+z)^{-n}}$$
$$\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{a(t)^2}{1-k(t)r^2}\mathrm{d}r^2 - a(t)^2 r^2\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta\,\mathrm{d}\varphi^2\right)$$
$$k(z) = -\frac{(n+6)(1-\Omega_m)(1+z)^{-(n+2)}}{|(n+6)(1-\Omega_m)|}$$
$$\frac{dr}{dz} = \sqrt{\frac{1-kr^2}{\Omega_m(1+z)^3+(1-\Omega_m)(1+z)^{-n}}}$$

2. analogiczny model z $k = 0$
3. ΛCDM model

Porównanie

- Dane
  - Union SNIa sample $N = 307$ (Kowalski et al. 2008)
  - położenie pików CMB (Hinshaw et al. 2007)

$$L \propto \exp\left[-\frac{1}{2}\left(\left(\frac{l_1 - 220.8}{0.7}\right)^2 + \left(\frac{l_{3/2} - 412.4}{1.9}\right)^2 + \left(\frac{l_2 - 530.9}{3.8}\right)^2\right)\right]$$

$$l_m = l_a(m - \phi_m);\ m = 1, 3/2, 2, \ldots;\ l_a = \pi \frac{r_*}{r_s}$$

  - BAO D_V (Eisenstein et al. 2005)

$$L \propto \exp\left[-\frac{(D_V^{\mathrm{theor}} - D_V^{\mathrm{obs}})^2}{2\sigma_{BAO}^2}\right]$$

$$D_V^{\mathrm{theor}} = \left[D_A^2 \frac{cz}{H(z)}\right]^{1/3},\ D_V^{\mathrm{obs}} = 1370 \pm 64 Mpc \text{ dla } z = 0.35$$

- założone równe prawdopodobieństwa początkowe dla każdego modelu
- wartość ewidencji numerycznie za pomocą nested sampling

Porównanie

Data sets	$\ln B_{31}$	$\ln B_{21}$	$\ln B_{32}$
Union+CMB	4.94	3.28	1.66
Union+CMB+BAO	4.23	2.58	1.65

Wnioski

- Pośród wielu teoretycznych propozycji ciemnej energii stała kosmologiczna jest najlepszą koncepcją. Z drugiej strony jest kluczowe pytanie: czym jest stała kosmologiczna, tzn jaka jest jej interpretacja fizyczna? Taką jaką byśmy chcieli widzieć - energia próżni kwantowej generuje najwieksza w historii fizyki rozbieżność rzędu 10^{120}.
- W każdej teorii fizycznej można wyróżnic sektor kinematyczny i dynamiczny. Kosmografia, którą posługujemy się w selekcji modeli kontroluje kinematyczny sektor teorii, podczas gdy stała kosmologiczna może zmieniać się z czasem. Pomiary CMB i widma zaburzeń sa dedykowane do estymacji zmienności Lambdy.
- Wszystkie koncepcje ciemnej energii można podzielić na 3 kategorie (substancjalne, ciemna energia jako wynik modyfikacji równań Einsteina, efekty ciemnej energii jako wynik uśredniania we Wszechświecie niejednorodnym). Metody Bayesowskie faworyzują koncepcję, że ciemna energia jest substancją.

1. M. Szydlowski, A. Kurek, A. Krawiec, Phys.Lett. B642, 171 (2006)
2. A. Kurek, M. Szydlowski, Astrophys. J. 675, 1 (2008)
3. A. Kurek, O. Hrycyna, M. Szydlowski, Phys.Lett. B659, 14 (2008)
4. A. Kurek, M. Szydlowski, NuovoCim. 122B, 1359 (2007)
5. A. Kurek, O. Hrycyna, M. Szydlowski, arXiv:0805.4005
6. K. Bolejko, A. Kurek, M. Szydlowski, arXiv:0811.4487