Predicting Exoplanet Mass-Radius Relationship: a Nonparametric Approach

Bo Ning[1]

with Angie Wolfgang[2,3] and Sujit Ghosh[1]

[1]North Carolina State University

[2]Pennsylvania State University

[3]NSF Astronomy & Astrophysics Postdoctoral Fellow

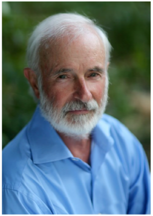
September 24, 2017

Project is supported by SAMSI and NSF.

Collaborators

Sujit Ghosh
(NC State, SAMSI)

Angie Wolfgang
(Penn State University)



Peter Bloomfield
(NC State University)

Subhashis Ghoshal
(NC State University)

Outline

Section 1

Background

- Astronomers have discovered thousands of exoplanets with either Mass or radius measurements
- Knowing a planet's mass and radius is important for understanding its compositions
- Only small portion planets have both mass and radius measurements
- Mass: radial velocity; Radius: transits
- To estimate the mass-radius relation (M-R relation) and use it to predict other planets' mass or radius

Background

- Astronomers have discovered thousands of exoplanets with either Mass or radius measurements
- Knowing a planet's mass and radius is important for understanding its compositions
- Only small portion planets have both mass and radius measurements
- Mass: radial velocity; Radius: transits
- To estimate the mass-radius relation (M-R relation) and use it to predict other planets' mass or radius

Background

- Astronomers have discovered thousands of exoplanets with either Mass or radius measurements
- Knowing a planet's mass and radius is important for understanding its compositions
- Only small portion planets have both mass and radius measurements
- Mass: radial velocity; Radius: transits
- To estimate the mass-radius relation (M-R relation) and use it to predict other planets' mass or radius

Background

- Astronomers have discovered thousands of exoplanets with either Mass or radius measurements
- Knowing a planet's mass and radius is important for understanding its compositions
- Only small portion planets have both mass and radius measurements
- Mass: radial velocity; Radius: transits
- To estimate the mass-radius relation (M-R relation) and use it to predict other planets' mass or radius

Background

- Astronomers have discovered thousands of exoplanets with either Mass or radius measurements
- Knowing a planet's mass and radius is important for understanding its compositions
- Only small portion planets have both mass and radius measurements
- Mass: radial velocity; Radius: transits
- To estimate the mass-radius relation (M-R relation) and use it to predict other planets' mass or radius

A hierarchical Bayesian power-law model

Model (HBM, WRF16)

$$M_i^{\text{obs}} \overset{ind}{\sim} \mathcal{N}(M_i, \sigma_{M,i}^{\text{obs}}),$$

$$R_i^{\text{obs}} \overset{ind}{\sim} \mathcal{N}(R_i, \sigma_{R,i}^{\text{obs}}),$$

$$M_i | R_i, C, \gamma, \sigma_M \sim \mathcal{N}(C R_i^{\gamma}, \sigma_M)$$

M_i is the planet mass divided by the Earth's mass, R_i is the planet radius divided by the Earth's radius.

A hierarchical Bayesian power-law model

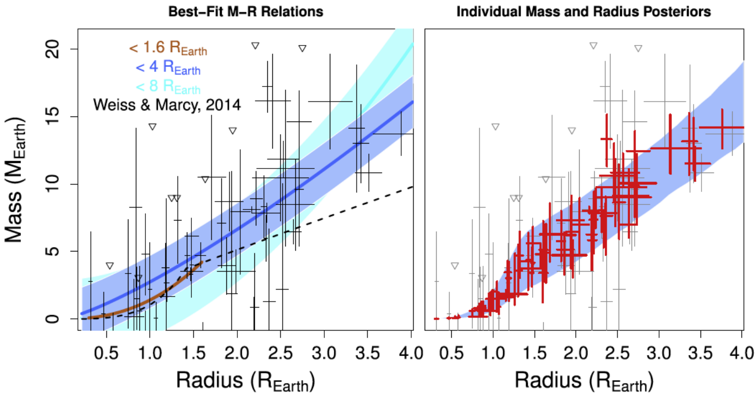


Figure: M-R relation using power-law model. (Copy from WRF16)

A hierarchical Bayesian power-law model

Model (HBM, WRF16)

$$M_i^{\text{obs}} \overset{ind}{\sim} \mathcal{N}(M_i, \sigma_{M,i}^{\text{obs}}),$$

$$R_i^{\text{obs}} \overset{ind}{\sim} \mathcal{N}(R_i, \sigma_{R,i}^{\text{obs}}),$$

$$M_i | R_i, C, \gamma, \sigma_M \sim \mathcal{N}(C R_i^{\gamma}, \sigma_M)$$

M_i is the planet mass divided by the Earth's mass, R_i is the planet radius divided by the Earth's radius.

- Normal distributed?
- Constant intrinsic scatter?
- Only one power-law?

Section 2

Nonparametric approaches

- Basis expansion, Gaussian process, Kernel methods, Dirichlet process, Pólya tree ....
- Basis expansion: spline functions, Bernstein polynomials, wavelets, trigonometric polynomials ....

Definition (Bernstein polynomial)

For a continuous function $F : [0,1] \to \mathbb{R}$*, the associated Bernstein polynomial is defined as*

$$B(x;k,F) = \sum_{k=0}^{d} F\left(\frac{k}{d}\right)\binom{d}{k} x^k (1-x)^{d-k}.$$

- As $d \to \infty$, $B(x;d,F)$ converge to F (uniformly) by Weierstrass approximation theorem

Nonparametric approaches

- Basis expansion, Gaussian process, Kernel methods, Dirichlet process, Pólya tree ....
- Basis expansion: spline functions, Bernstein polynomials, wavelets, trigonometric polynomials ....

Definition (Bernstein polynomial)

For a continuous function $F : [0, 1] \to \mathbb{R}$, the associated Bernstein polynomial is defined as

$$B(x; k, F) = \sum_{k=0}^{d} F\left(\frac{k}{d}\right)\binom{d}{k} x^k (1-x)^{d-k}.$$

- As $d \to \infty$, $B(x; d, F)$ converge to F (uniformly) by Weierstrass approximation theorem

Nonparametric approaches

- Basis expansion, Gaussian process, Kernel methods, Dirichlet process, Pólya tree ....
- Basis expansion: spline functions, Bernstein polynomials, wavelets, trigonometric polynomials ....

Definition (Bernstein polynomial)

For a continuous function $F : [0,1] \to \mathbb{R}$, the associated Bernstein polynomial is defined as

$$B(x; k, F) = \sum_{k=0}^{d} F\left(\frac{k}{d}\right)\binom{d}{k} x^k (1-x)^{d-k}.$$

- As $d \to \infty$, $B(x; d, F)$ converge to F (uniformly) by Weierstrass approximation theorem

Bernstein polynomial

- A Bernstein polynomial density can be obtained by taking derivative on $B(x; k, F)$, such that

$$b(x; k, f) = \sum_{k=1}^{d} \left(F\left(\frac{k}{d}\right) - F\left(\frac{k-1}{d}\right) \right) \beta_k(x; k, d-k+1),$$

where $\beta_k(x; k, d-k+1)$ is a beta density.

- One often estimates the density by rewriting it corresponding to a weight sequence $\boldsymbol{w} = (w_1, \ldots, w_d)$, such that

$$f_N(x|\boldsymbol{w}) \equiv b(x; k, f) = \sum_{k=1}^{d} w_k \beta_k(x; k, d-k+1), \ \sum_k w_k = 1, \ w_k \geq 0.$$

Bernstein polynomial

- A Bernstein polynomial density can be obtained by taking derivative on $B(x; k, F)$, such that

$$b(x; k, f) = \sum_{k=1}^{d} \left(F\left(\frac{k}{d}\right) - F\left(\frac{k-1}{d}\right) \right) \beta_k(x; k, d-k+1),$$

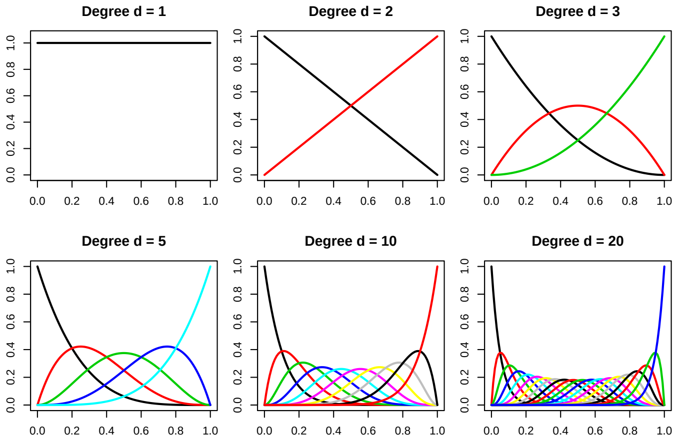
where $\beta_k(x; k, d-k+1)$ is a beta density.

- One often estimates the density by rewriting it corresponding to a weight sequence $\boldsymbol{w} = (w_1, \ldots, w_d)$, such that

$$f_N(x|\boldsymbol{w}) \equiv b(x; k, f) = \sum_{k=1}^{d} w_k \beta_k(x; k, d-k+1), \ \sum_k w_k = 1, \ w_k \geq 0.$$

Bernstein polynomial (cont'd)

$$f_N(x|\boldsymbol{w}) = \sum_{k=1}^{d} w_k \beta_k(x; k, d-k+1), \; \sum_k w_k = 1, \; w_k \geq 0.$$



Connections

- Connection to mixture models: $f_N(x|\boldsymbol{w}) = \sum_{k=1}^{d} w_k \beta_k(x; k, d-k+1)$
  - Clustering: number of power-laws
  - d is not the number of clusters
  - Gaussian mixture models: requires to estimate parameters in each Gaussian component
- Connection to a multivariate density estimation. For $x, y \in [0, 1]$, a bivariate Bernstien polynomial density is,

$$f(x, y; k, F) = \sum_{k=1}^{d_1} \sum_{l=1}^{d_2} w_{kl} \beta_k(x; k, d_1 - k + 1) \beta_l(y; l, d_2 - l + 1),$$

$$\sum_{k=1}^{d_1} \sum_{l=1}^{d_2} w_{kl} = 1, \ w_{kl} \geq 0$$

  - Modeling the joint density: when both masses and radii have measurement errors.
  - The conditional and marginal distributions are mixture of beta distributions

Connections

- Connection to mixture models: $f_N(x|\boldsymbol{w}) = \sum_{k=1}^{d} w_k \beta_k(x; k, d-k+1)$
  - Clustering: number of power-laws
  - d is not the number of clusters
  - Gaussian mixture models: requires to estimate parameters in each Gaussian component
- Connection to a multivariate density estimation. For $x, y \in [0, 1]$, a bivariate Bernstien polynomial density is,

$$f(x, y; k, F) = \sum_{k=1}^{d_1} \sum_{l=1}^{d_2} w_{kl} \beta_k(x; k, d_1 - k + 1) \beta_l(y; l, d_2 - l + 1),$$

$$\sum_{k=1}^{d_1} \sum_{l=1}^{d_2} w_{kl} = 1, \; w_{kl} \geq 0$$

  - Modeling the joint density: when both masses and radii have measurement errors.
  - The conditional and marginal distributions are mixture of beta distributions

Connections (cont'd)

- Connection to Bayesian nonparametric: we could put a Dirichlet prior on $\boldsymbol{w}$
  - Further connections to Bayesian density estimation
  - **Spectral density estimation: smoothing the periodogram*
- Other connections will not mention in details: i.e., B-spline

Connections (cont'd)

- Connection to Bayesian nonparametric: we could put a Dirichlet prior on $\boldsymbol{w}$
  - Further connections to Bayesian density estimation
  - **Spectral density estimation: smoothing the periodogram*
- Other connections will not mention in details: i.e., B-spline

Section 3

Bernstein polynomials model

Model (Bernstein polynomials model)

$$M_i^{\mathrm{obs}} \overset{ind}{\sim} \mathcal{N}(M_i, \sigma_{M_i}^{\mathrm{obs}}),$$

$$R_i^{\mathrm{obs}} \overset{ind}{\sim} \mathcal{N}(R_i, \sigma_{R_i}^{\mathrm{obs}}),$$

$$(M_i, R_i) \overset{iid}{\sim} f(m, r|\boldsymbol{w}, d),$$

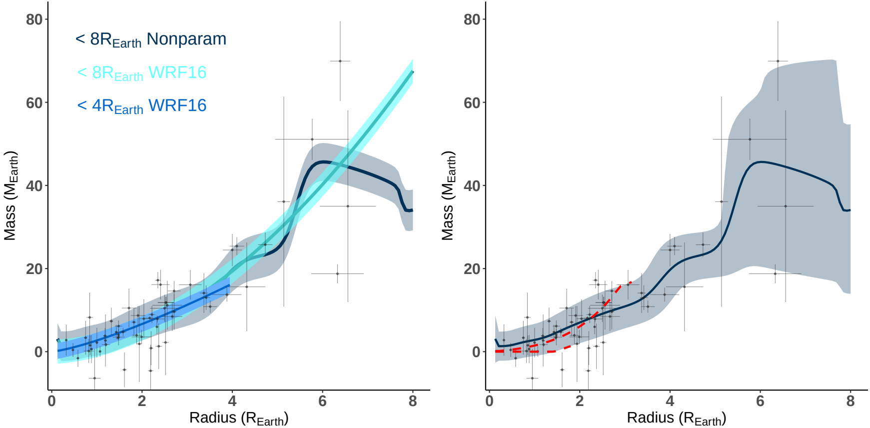
$$f(m, r|\boldsymbol{w}, d) = \sum_{k=1}^{d}\sum_{l=1}^{d} w_{kl} \frac{\beta_k(\frac{m-\underline{M}}{\overline{M}-\underline{M}})}{\overline{M}-\underline{M}} \frac{\beta_l(\frac{r-\underline{R}}{\overline{R}-\underline{R}})}{\overline{R}-\underline{R}},$$

where $\boldsymbol{w} = (w_{11}, \ldots, w_{dd})$, $\sum_{k=1}^{d}\sum_{l=1}^{d} w_{kl} = 1$, $w_{kl} \geq 0$.

- Estimate d using 10-fold cross validation
- Estimate $\boldsymbol{w}$ using convex programming package “Rsonlp” in R

Section 4

M-R relations: comparison of two models



WRF16:

RV only $< 4R_{\oplus}$: $M_i|R_i \sim N(2.7R_i^{1.3}, 1.9^2)$; **RV only** $< 8R_{\oplus}$: $M_i|R_i \sim N(1.6R_i^{1.8}, 2.9^2)$

M-R relations: comparison of two models

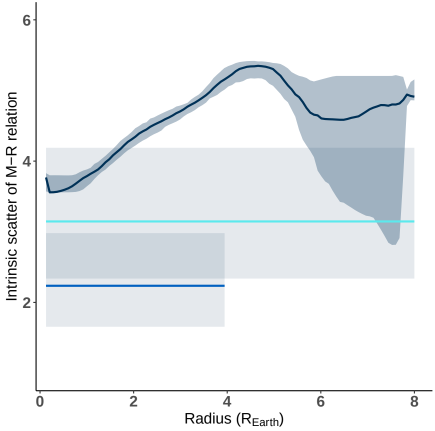


Figure: Instrinsic scatter plot for M-R relations.

Blue and light blue line: Power-law model. Dark blue line: Bernstein polynomials model

M-R relation for full Kepler dataset

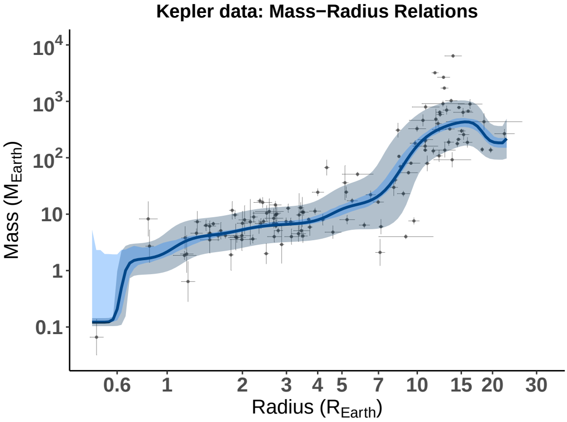


Figure: M-R relation for Kepler dataset.

Dark line: mean M-R relation. Grey area: 16% and 84% prediction intervals. Blue area: 16% and 84% bootstrap confidence intervals

M-R relation for full Kepler dataset: conditional densities

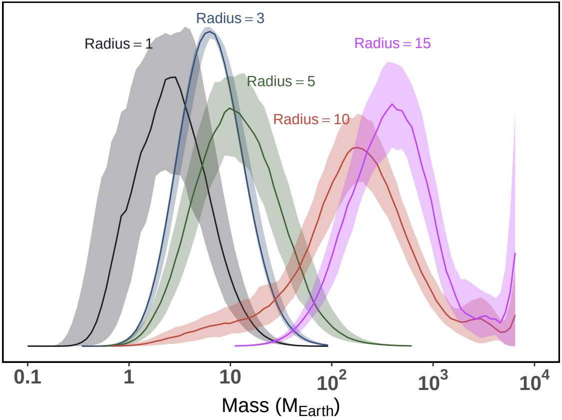


Figure: The conditional distributions for mass given radius

The uncertainty region are16% and 84% bootstrap confidence intervals.

Section 5

Conclusion and comments

- We considered a more flexible Bernstein polynomial model to estimate the M-R relation.
- Bernstein polynomials model is a mixture beta model
- Compares to the power-law model, the power-law model is underfitting the data, thus have smaller s.d.
- Easy to extent the model to incorporate a third variable
- Statistics properties of this model is under investigation

Looking for more details? Our draft is coming soon: Ning, Wolfgang & Ghosh (2017).

Reference

Bernstein polynomials

NWG17 Ning, B., Wolfgang, A. and Ghosh, S. K. 2017, in preparation

TK13 Turnbull, B. C. and Ghosh, S. K. 2014. Computational Statistics & Data Analysis, 13, 72

BCC02 Babua, G. J., Cantyb, A. J., Chaubeyb, Y. P. 2002, Journal of Statistical Planning and Inference, 105, 377

Spectral density estimation

CGR04 Chuoudhuri, N., Ghosal, S., Roy, A. 2004, Journal of the American Statistics Association, 99, 486

EMC17 Edwards, M. C., Meyer, R., Christensen, N. 2017, arXiv: 1707.04878v1

Bayesian density estimation using Bernstein polynomials

P99-a Petrone, S.1999, the Scandinavian Journal of Statistics, 26, 373

P99-b Petrone, S.1999, the Canadian Journal of Statistics, 27, 105

PW02 Petrone, S., Wasserman, L. 2002, Journal of the Royal Statistical Society, Series B, 64, 79

G01 Ghosal, S. 2001, the Annals of Statistics, 29, 5

GV17 Ghosal, S., van der Vaart, A. 2017, Cambridge university press.