Hierarchical Bayes Modeling for Large-Scale Inference

Daniel Yekutieli and Asaf Weinstein

CIRM Virtual Conference June 2020

< ロ > < 同 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >

# Plan

- 1. Background
- 2. Optimal inference via the Empirical Oracle Distribution
  - Bayesian perspective
  - Frequentist perspective
- 3. Implementation by hierarchical Bayesian modeling
- 4. High-dimensional logistic regression Example of Sur and Candes (2019)

イロト イヨト イヨト イヨト

# Background

- Model selection (averaging) procedures that correspond to eliciting a shrinkage prior distribution to model parameters (Ridge Regression; LASSO; Spike and Slab; ABSLOPE )
- Empirical Bayes (Robbins, 1956; James and Stein, 1961; Brown, 1966; Efron et al., 2001; Sun and Cai, 2007; Brown and Greenshtein, 2009; Efron, 2011)
- Compound decision approach (Robbins, 1951; Zhang, 2003; Weinstein et al., 2018): marginal Bayes rules with respect to the empirical distribution of the parameter vector are optimal frequentist procedures (i.e. minimize the frequentist compound risk for any fixed  $\vec{\theta} = (\theta_1, \dots, \theta_n)$ )
- Our hierarchical Bayesian model uses a finite Polya tree on a dyadic partition to define a random distribution on  $[a_0, a_I]$  (Ferguson, 1974)

イロト 不得 とくき とくき とうき

# Setup and notations

- Parameter vector  $\vec{\theta} = (\theta_1 \cdots \theta_m)$ , we *initially* assume  $\vec{\theta} \sim \pi(\vec{\theta})$
- Observation  $\vec{Y} = (Y_1 \cdots Y_n)$ , with likelihood  $\vec{Y} \sim f(\vec{y}|\vec{\theta})$
- Let  $\mathcal{O}(\vec{\theta}) = (\theta_{(1)} \leq \cdots \leq \theta_{(m)})$  denote the order statistic of  $\vec{\theta}$
- Let  $Perm_m$  denote the set of permutations on  $\{1, 2, \dots, m\}$
- Thus  $\pi(\vec{\theta})$  specifies a distribution on parameter order statistics

$$ilde{\pi}(\mathcal{O}(ec{ heta})) = \sum_{\sigma \in \mathit{Perm}_m} \pi(\sigma(\mathcal{O}(ec{ heta}))).$$

• For all  $\sigma \in Perm_m$  and  $\tilde{\theta} = \mathcal{O}(\vec{\theta})$  we further define

$$\tilde{\pi}(\sigma(\tilde{\theta})|\tilde{\theta}) = \frac{\pi(\sigma(\tilde{\theta}))}{\tilde{\pi}(\tilde{\theta})}$$
(1)

イロン イロン イヨン イヨン 三日

# **Bayes Rules**

<u>Goal</u>: find  $\delta(\cdot)$  that minimizes the average risk  $E_{\vec{\theta} \sim \pi} E_{\vec{Y} \sim f(\vec{y}|\vec{\theta})} L(\delta(\cdot), \vec{\theta})$ 

• Observing  $\vec{Y}$  the average risk is minimized by Bayes rule

$$\delta^{Bayes}(\vec{y}) = \operatorname{argmin} E_{\vec{\theta} \sim \pi(\vec{\theta} | \vec{Y} = \vec{y})} L(\delta(\vec{y}), \vec{\theta})$$

• Observing  $\vec{Y}$  and  $\mathcal{O}(\vec{\theta}) = \tilde{\theta}$  we may define the Bayes rule

 $\delta^{Bayes}(\tilde{\theta}, \vec{y}) = \operatorname{argmin} E_{\vec{\theta} \sim \pi(\vec{\theta} \mid \mathcal{O}(\vec{\theta}) = \tilde{\theta}, \vec{Y} = \vec{y})} L(\delta(\tilde{\theta}, \vec{y}), \vec{\theta})$ 

that *per definition* yields a smaller average risk than  $\delta^{Bayes}(\vec{y})$ .

The empirical Oracle Distribution

#### The posterior distribution for symmetric priors

Assuming a symmetric prior

$$\forall \sigma \in \operatorname{Perm}_m, \ \pi(\vec{\theta}) = \pi(\sigma(\vec{\theta})) \quad \Rightarrow \quad \tilde{\pi}(\sigma(\tilde{\theta})|\tilde{\theta}) = \frac{1}{m!}$$

we may express

$$\begin{aligned} \pi(\vec{\theta} \mid \mathcal{O}(\vec{\theta}) &= \tilde{\theta}, \vec{Y} = \vec{y}) = \frac{f(\vec{\theta}, \tilde{\theta}, \vec{y})}{\sum_{\sigma \in Perm} f(\sigma(\tilde{\theta}), \tilde{\theta}, \vec{y})} \\ &= \frac{\tilde{\pi}(\vec{\theta} \mid \tilde{\theta}) \cdot \tilde{\pi}(\tilde{\theta}) \cdot f(\vec{y} \mid \vec{\theta})}{\sum_{\sigma \in Perm} \tilde{\pi}(\sigma(\tilde{\theta}) \mid \tilde{\theta}) \cdot \tilde{\pi}(\tilde{\theta}) \cdot f(\vec{y} \mid \sigma(\vec{\theta}))} \\ &= \frac{\tilde{\pi}(\vec{\theta} \mid \tilde{\theta}) \cdot f(\vec{y} \mid \vec{\theta})}{\sum_{\sigma \in Perm} \tilde{\pi}(\sigma(\tilde{\theta}) \mid \tilde{\theta})) \cdot f(\vec{y} \mid \sigma(\vec{\theta}))} = \frac{f(\vec{y} \mid \vec{\theta})}{\sum_{\sigma \in Perm_m} f(\vec{y} \mid \sigma(\vec{\theta}))} \end{aligned}$$

イロト 不得 トイヨト イヨト

The empirical Oracle Distribution

# The Empirical Oracle Distribution

• The Empirical Oracle posterior Distribution

$$\pi^{EOD}(\vec{\theta} \mid \mathcal{O}(\vec{\theta}), \vec{y}) = \frac{f(\vec{y} \mid \vec{\theta})}{\sum_{\sigma \in Perm_m} f(\vec{y} \mid \sigma(\tilde{\theta}))}$$

• e.g. EOD Bayes rule for  $L(\delta(\vec{Y}), \vec{\theta}) = \|\delta(\vec{Y}) - \vec{\theta}\|^2$  is

$$\delta^{EOD}(\tilde{\theta}, \vec{y}) = \frac{\sum_{\sigma \in Perm_m} \sigma(\tilde{\theta}) \cdot f(\vec{y} | \sigma(\tilde{\theta}))}{\sum_{\sigma \in Perm_m} f(\vec{y} | \sigma(\tilde{\theta}))},$$

• EOD Bayes rules yield minimal average risk of all Bayes rules for symmetric priors (in shrinkage priors  $\theta_i$  are iid, MLE is MAP for flat prior)

イロト イロト イヨト イヨト

# The frequentist perspective

 $\vec{\theta}$  fixed unknown and the goal is to find  $\delta(\vec{y})$  that minimizes

$$Risk(\vec{\theta}; \delta(\cdot)) = E_{\vec{Y} \sim f(\vec{y}|\vec{\theta})} L(\delta(\vec{Y}), \vec{\theta})$$

• Whereas we have  $\delta^{EOD}(\tilde{\theta}, \vec{y})$  that by construction minimizes

 $\sum_{\sigma \in Perm_m} Risk(\sigma(\tilde{\theta}); \delta(\cdot))/m!$ 

- Symmetric likelihood and loss:  $Risk(\sigma(\tilde{\theta}); \delta^{EOD}) \equiv Risk(\tilde{\theta}; \delta^{EOD})$  $\Rightarrow \delta^{EOD}$  yields minimal frequentist risk of all symmetric  $\delta(\vec{y})$
- In particular, for sequence model and compound loss  $\rightarrow$  Robbins (1951)

・ロト ・ 同ト ・ ヨト ・ ヨト ・ ヨ

# The L level hierarchical Beta model

Ferguson (1974): Generative model for distributions with step functions PDF's at endpoints  $\vec{a} = (a_0, \dots, a_{2^L})$  corresponding to dyadic partition of  $[a_0, a_{2^L}]$  for nonparametric density estimation

- Independent Beta random variables  $\phi_{l,j} \sim Beta(\alpha_{l,j}, \beta_{l,j})$  that specify the conditional subinterval probabilities for the dyadic partitions
- $\pi_{1,1} \cdots \pi_{L,2^L}$  are the probabilities of the subintervals in the dyadic partitions that are products of the Beta random variables
- Step function PDF for  $\theta \in [a_0, a_{2^L}]$

$$f(\theta; \vec{a}, \vec{\pi}) = \pi_{L,1} \cdot \frac{I_{[a_0, a_1]}(\theta)}{a_1 - a_0} + \dots + \pi_{L, 2^L} \cdot \frac{I_{[a_{2^L}, a_{2^L}]}(\theta)}{a_{2^L} - a_{2^L-1}}$$

Implementation by hBeta model

#### The 3 level hierarchical Beta model



# Hierarchical Bayes modeling for Large-Scale Inference

- a. Parameter  $\vec{\theta} = (\theta_1 \cdots \theta_m)$ , data  $\vec{Y} = (Y_1 \cdots Y_n)$ , likelihood  $\vec{Y} \sim f(\vec{y}|\vec{\theta})$
- b. We imbed likelihood in (made up) hierarchical model for the data:
  - 1. Generate  $f(\theta; \vec{a}, \vec{\pi})$  from hBeta model with  $\phi_{l,j} \sim Beta(1, 1)$
  - 2. For  $i = 1 \cdots m$  generate iid  $\theta_i \sim f(\theta; \vec{a}, \vec{\pi})$
  - 3. Generate  $\vec{Y} \sim p(\vec{y}; \vec{\theta})$

c. We use a Gibbs sampler to derive the posterior distribution of the hBeta model given  $\vec{Y} = \vec{y}$ , in which the Gibbs samples of  $f(\theta; \vec{a}, \vec{\pi})$  are deconvolution estimates for the empirical marginal distribution of  $\vec{\theta}$  and Gibbs samples of  $\vec{\theta}$  approximate posterior samples of  $\pi^{EOD}(\vec{\theta} | \mathcal{O}(\vec{\theta}), \vec{y})$ 

d. Our inferences are Bayes rules for on Gibbs sampling distribution of  $\vec{\theta}$ .

# Candes and Sur (2019): High dim. logistic regression

#### Simulation study

- $X_{4000\times800}$  with iid N(0, 1/n) entries
- $Y_j = Bernoulli(p_j)$ , for  $j = 1 \cdots 4000$  where
- $p_j = \exp(\mu_j)/(1 + \exp(\mu_j))$  with  $\vec{\mu}_{4000 \times 1} = X\vec{\beta}_{800 \times 1}$ :

a. 
$$\vec{\beta} = (-10, \cdots, -10, 10, \cdots, 10, 0, \cdots, 0)$$

- b.  $\beta_i \sim N(3, 4^2)$ ,  $i = 1 \cdots 800$
- c.  $\beta_i = 0$  or  $\beta_i \sim N(3, 4^2)$  with probability 0.5

イロト 不得 とくき とくき とうき

# Implementing hBeta approach

- For the hBeta approach we <u>assume</u> generative model for  $\vec{Y}$  is
  - 1. Generate  $f(\beta; \vec{a}, \vec{\pi})$  from hierarchical Beta model with L = 6,  $\phi_{l,j} \sim Beta(1, 1)$ , and  $\vec{a}$  is a regular 65 point grid on [-20, 20].
  - 2. For  $i = 1 \cdots 800$  generate  $\beta_i \sim f(\theta; \vec{a}, \vec{\pi})$ , and compute  $\vec{\mu} = X\vec{\beta}$
  - 3. For  $j = 1 \cdots 4000$ , compute  $p_j = \exp(\mu_j)/(1 + \exp(\mu_j))$  and then generate  $Y_j \sim Bernoulli(p_j)$
- We compare five estimates: MLE; "corrected" MLE of Candes and Sur (2019); LASSO and Ridge penalized likelihood estimates (R GLMNET); hBeta posterior means.

イロト 不得 とくき とくき とうき



Gibbs sampler disribution of the coefficient distribution CDF

Interval center

イロト イポト イヨト イヨト



#### Coefficients estimation plot

Index

< ロ > < 同 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >



Coefficient shrinkage plot

Beta MLE's

< ロ > < 同 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >



95% credible intervals for coefficients (FCP = 8/800)

# Summary of results

		adj.MLE	LASSO	Ridge	hBeta
Example a	$\vec{\beta}$	0.33	0.19	0.19	0.10
	$\vec{\mu}$	0.31	0.21	0.20	0.11
	$\vec{p}$	0.75	0.49	0.61	0.34
Example b	$\vec{\beta}$	0.34	0.38	0.26	0.17
	$\vec{\mu}$	0.32	0.38	0.27	0.17
	$\vec{p}$	0.75	0.80	0.64	0.32
Example c	$\vec{\beta}$	0.36	0.27	0.25	0.18
	$\vec{\mu}$	0.34	0.26	0.26	0.19
	$\vec{p}$	0.76	0.67	0.63	0.48

Table: MSE for single realization displayed as fractions of the MSE for the MLE.

イロト イポト イヨト イヨト

# Discussion

- Given a hierarchical model for the data, the risks and average risks for the different estimators can be readily computed
- We propose using the EOD average loss as a benchmark for specifying difficulty of inferential problem and comparing estimation methods
- Exchangeable priors make sense when there is no previous information on problem

< ロ > < 同 > < 回 > < 回 > .

Fin

◆□ → ◆□ → ◆ 三 → ◆ □ → ◆ ○ ◆ ○ ◆ ○ ◆ ○ ◆