High-Dimensional:

The Barrier and Bayes  and Bias

D A S Fraser
Statistics
U Toronto

International Workshop on Perspectives on
High-dimensional Data Analysis

Fields Institute
Toronto June 9–11 2011

http://www.utstat.toronto.edu/dfraser/documents/fields11.pdf

Some references ⇒ /xxx.pdf

Very special thanks to:

- Ejaz 

- the Organizing Committee: Ejaz Ahmed
  Peter Song
  Mu Zhu

- and the supporting Organizations...

Statistical Society of Canada
Société statistique du Canada

NELSON EDUCATION

FIELDS

University of Windsor

MICHIGAN

University of Waterloo

- and  Bayes

Very special thanks to:

- Ejaz 

- the Organizing Committee: Ejaz Ahmed
  Peter Song
  Mu Zhu

- and the supporting Organizations...

Statistical Society of Canada
Société statistique du Canada

NELSON EDUCATION

FIELDS

University of Windsor

MICHIGAN

University of Waterloo

- and  Bayes

0. High-dimension Barrier

1 A first encounter: high dimension-Barrier

---

2 The Barrier: in data processing; in model-based

3 Gradient of a prior ... Bayes

4 Standardize and calibrate θ

5 Origins: Welch & Peers 1963

6 Gradient: Fine tuning...

7 Summary: "Check the bias from your prior" High Risks!

8 Thank you!

Slice it --- Bayes
Bridge it --- asymptotics
Calibrate: .... Risks

0. High-dimension Barrier

1 A first encounter: high dimension-Barrier

2 The Barrier: in data processing; in model-based analysis; ..

3 Gradient of a prior ... Bayes

4 Standardize and calibrate θ

5 Origins: Welch & Peers 1963

6 Gradient: Fine tuning...

7 Summary: "Check the bias from your prior" High Risks!

8 Thank you!

Slice it --- Bayes
Bridge it --- asymptotics
Calibrate: .... Risks

0. High-dimension Barrier

1 A first encounter: high dimension-Barrier

2 The Barrier: in data processing; in model-based

3 Gradient of a prior ... Bayes

---

4 Standardize and calibrate θ

5 Origins: Welch & Peers 1963

6 Gradient: Fine tuning...

7 Summary: "Check the bias from your prior" High Risks!

8 Thank you!

Slice it --- Bayes

Bridge it... asymptotics

Calibrate: .... Risks

0. High-dimension Barrier

1 A first encounter: high dimension-Barrier

2 The Barrier: in data processing; in model-based

3 Gradient of a prior ... Bayes

4 Standardize and calibrate θ

---

5 Origins: Welch & Peers 1963

6 Gradient: Fine tuning...

7 Summary: "Check the bias from your prior" High Risks!

8 Thank you!

Slice it --- Bayes

Bridge it--- asymptotics

Calibrate: .... Risks

0. High-dimension Barrier

1 A first encounter: high dimension-<u>Barrier</u>

2 The Barrier: in data processing; in <u>model-based</u>

3 <u>Gradient</u> of a prior ... Bayes

4 Standardize and <u>calibrate</u> θ

5 <u>Origins</u>: Welch & Peers 1963

---

6 Gradient: <u>Fine</u> tuning...

7 Summary: "Check the bias from your prior" High Risks!

8 Thank you!

<u>Slice</u> it --- Bayes

<u>Bridge</u> it--- asymptotics

<u>Calibrate</u>: .... Risks

0. High-dimension Barrier

1 A first encounter: high dimension-Barrier

2 The Barrier: in data processing; in model-based

3 Gradient of a prior ... Bayes

4 Standardize and calibrate θ

5 Origins: Welch & Peers 1963

6 Gradient: Fine tuning...

---

7 Summary: "Check the bias from your prior" High Risks!

8 Thank you!

Slice it --- Bayes

Bridge it --- asymptotics

Calibrate: .... Risks

0. High-dimension Barrier

1 A first encounter: high dimension-Barrier
2 The Barrier: in data processing; in model-based
3 Gradient of a prior ... Bayes
4 Standardize and calibrate θ
5 Origins: Welch & Peers 1963
6 Gradient: Fine tuning...
7 Summary: "Check the bias from your prior" High Risks!
8 Thank you!

Slice it --- Bayes
Bridge it --- asymptotics
Calibrate: .... Risks

0. High-dimension Barrier

1 A first encounter: high dimension-Barrier
2 The Barrier: in data processing; in model-based
3 Gradient of a prior ... Bayes
4 Standardize and calibrate θ
5 Origins: Welch & Peers 1963
6 Gradient: Fine tuning...
7 Summary: "Check the bias from your prior" High Risks!
8 Thank you!

Slice it --- Bayes
Bridge it--- asymptotics
Calibrate: .... Risks

(1) First Encounter

- I won't count back ...

- Importance sampling had emerged ... "the answer"
- modest problem, grad student, try it out!
- Spectacular! Get value: y_i
  " weight: w_i importance ... LR
- Looked nice for a while ... nice w_i's .7, 1.5, ...
  then 3×10^5
  blew-up!
- Sabbatical - a grad student, then staff
  - continued ... broke the Barrier ... McMC
  - Keith Hastings ... of Metropolis-Hastings
- but Barrier really still there!
  Just a little farther away!
- That's why we are here ...

① First Encounter

- I won't count back ...
- Importance sampling had emerged ... "the answer"

- modest problem, grad student, try it out!
- Spectacular! Get value: y_i
  " weight: w_i importance ... LR
- Looked nice for a while ... nice w_i's .7, 1.5, ...
  then 3×10^5
  blew-up!
- Sabbatical - a grad student, then staff
  - continued ... broke the Barrier ... McMC
  - Keith Hastings ... of Metropolis-Hastings
- but Barrier really still there!
  Just a little farther away!
- That's why we are here ...

① First Encounter

- I won't count back ...
- Importance sampling had emerged ... "the answer"
- modest problem, grad student, try it out !

- Spectacular! Get value: y_i
  " weight: w_i importance ... LR
- Looked nice for a while ... nice w_i's .7, 1.5, ...
  then 3×10^5
  blew-up !
- Sabbatical - a grad student, then staff
  - continued ... broke the Barrier ... McMC
  - Keith Hastings ... of Metropolis-Hastings
- but Barrier really still there !
  Just a little farther away !
- That's why we are here ...

(1) First Encounter

- I won't count back ...
- Importance sampling had emerged ... "the answer"
- modest problem, grad student, try it out !
- Spectacular ! Get value: y_i
  " weight: w_i importance ... LR

- Looked nice for a while ... nice w_i's .7, 1.5, ...
  then 3×10^5
  blew-up !
- Sabbatical - a grad student, then staff
  - continued ... broke the Barrier ... McMC
  - Keith Hastings ... of Metropolis-Hastings
- but Barrier really still there !
  Just a little farther away !
- That's why we are here ...

(1) First Encounter

- I won't count back ...
- Importance sampling had emerged ... "the answer"
- modest problem, grad student, try it out!
- Spectacular! Get value: y_i
  " weight: w_i importance ... LR
- Looked nice for a while ... nice w_i's .7, 1.5, ...
  then 3×10^5
  blew-up!

- Sabbatical - a grad student, then staff
  - continued ... broke the Barrier ... McMC
  - Keith Hastings ... of Metropolis-Hastings
- but Barrier really still there!
  Just a little farther away!
- That's why we are here ...

(1) First Encounter

- I won't count back ...
- Importance sampling had emerged ... "the answer"
- modest problem, grad student, try it out!
- Spectacular! Get value: y_i
  " weight: w_i importance ... LR
- Looked nice for a while ... nice w_i's .7, 1.5, ...
  then 3×10^5
  blew-up!
- Sabbatical - a grad student, then staff
  - continued ... broke the Barrier ... MCMC

- Keith Hastings ... of Metropolis-Hastings

- but Barrier really still there!
  Just a little farther away!

- That's why we are here ...

(1) First Encounter

- I won't count back ...
- Importance sampling had emerged ... "the answer"
- modest problem, grad student, try it out !
- Spectacular ! Get value: y_i
  " weight: w_i importance ... LR
- Looked nice for a while ... nice w_i's .7, 1.5, ...
  then 3×10^5
  blew-up !
- Sabbatical - a grad student, then staff
  - continued ... broke the Barrier ... McMC
  - Keith Hastings ... of Metropolis-Hastings

- but Barrier really still there !
  Just a little farther away !
- That's why we are here ...

(1) First Encounter

- I won't count back ...
- Importance sampling had emerged ... "the answer"
- modest problem, grad student, try it out !
- Spectacular! Get value: y_i
  " weight: w_i importance ... LR
- Looked nice for a while ... nice w_i's .7, 1.5, ...
  then 3×10^5
  blew-up !
- Sabbatical - a grad student, then staff
  - continued ... broke the Barrier ... McMC
  - Keith Hastings ... of Metropolis-Hastings

- but Barrier really still there !
  Just a little farther away !

- That's why we are here ...

(1) First Encounter

- I won't count back ---
- Importance sampling had emerged ... "the answer"
- modest problem, grad student, try it out !
- Spectacular ! Get value : y_i
  " weight : w_i importance ... LR
- Looked nice for a while ... nice w_i's .7, 1.5, ...
  then 3×10^5
  blew-up !
- Sabbatical - a grad student, then staff
  - continued ... broke the Barrier ... McMC
  - Keith Hastings ... of Metropolis - Hastings

- but Barrier really still there !
  Just a little farther away !
- That's why we are here ...

(2) The <u>Barrier</u>? in "data-processing"; in "<u>model</u>-based analysis"; ...

Bring dimension down!

On data space: Conditioning Integration Laplace & <u>Asymptotics</u>

On model space: <u>Bayes</u> (<u>Slice</u>); Integration <u>Asymptotics</u> & McMC

<u>Slice</u>: y^0- slice · model $L(\theta) = f(y^0; \theta)$

<u>Bridge</u> (asymptotics) <u>Taylor</u> series "about ∞"

(2) The Barrier? in "data-processing"; in "model-based analysis"; ...

Bring dimension down!

On data space: [1] Conditioning [2] Integration: Laplace & Asymptotics

On model space: Bayes (Slice); Integration Asymptotics & MCMC

Slice: y^0- slice model $L(\theta) = f(y^0; \theta)$

Bridge (asymptotics) Taylor series "about ∞"

(2) The <u>Barrier</u>? in "data-processing"; in "<u>model</u>-based analysis"; ...

Bring dimension down!

On data space:	Conditioning	Integration	Laplace & <u>Asymptotics</u>
On model space:	<u>Bayes</u> (<u>Slice</u>) ;	Integration	<u>Asymptotics</u> & McMC

<u>Slice</u>: y^0- slice · model $L(\theta) = f(y^0; \theta)$

<u>Bridge</u> (asymptotics) <u>Taylor</u> series "about ∞"

② The Barrier? in "data-processing"; in "model-based analysis"; ...

Bring dimension down!

On data space: Conditioning Integration Laplace & Asymptotics

On model space: Bayes (Slice); Integration Asymptotics & McMC

Slice: y°- slice . model $L(\theta) = f(y^\circ; \theta)$

Bridge (asymptotics) Taylor series "about ∞"

(2) The Barrier? in "data-processing"; in "model-based analysis"; ...

Bring dimension down!

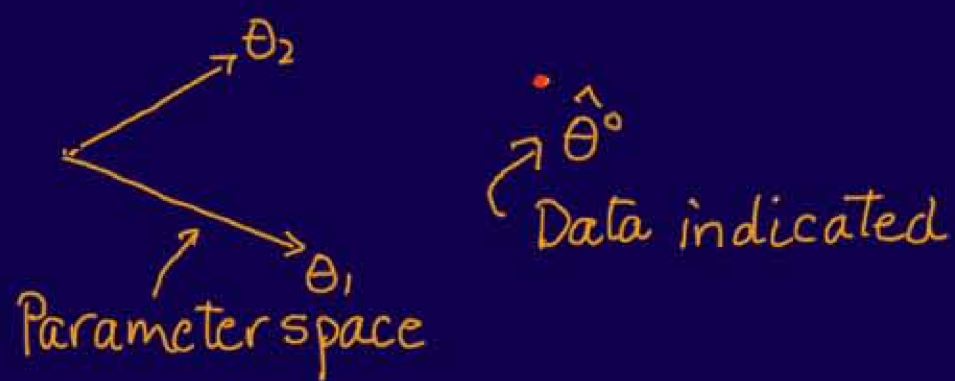
On data space: Conditioning Integration Laplace & Asymptotics

On model space: Bayes (Slice) ; Integration Asymptotics & MCMC

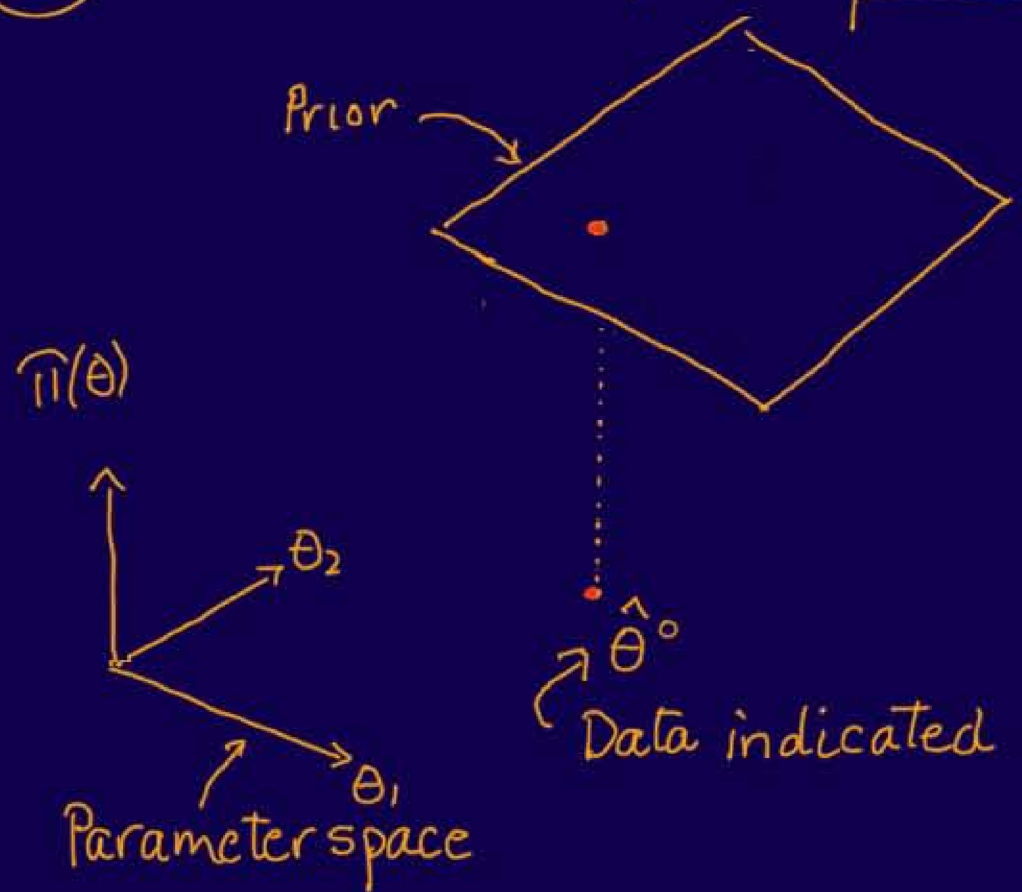
Slice: y^0- slice . model $L(\theta) = f(y^0; \theta)$

Bridge (asymptotics) Taylor series "about ∞"

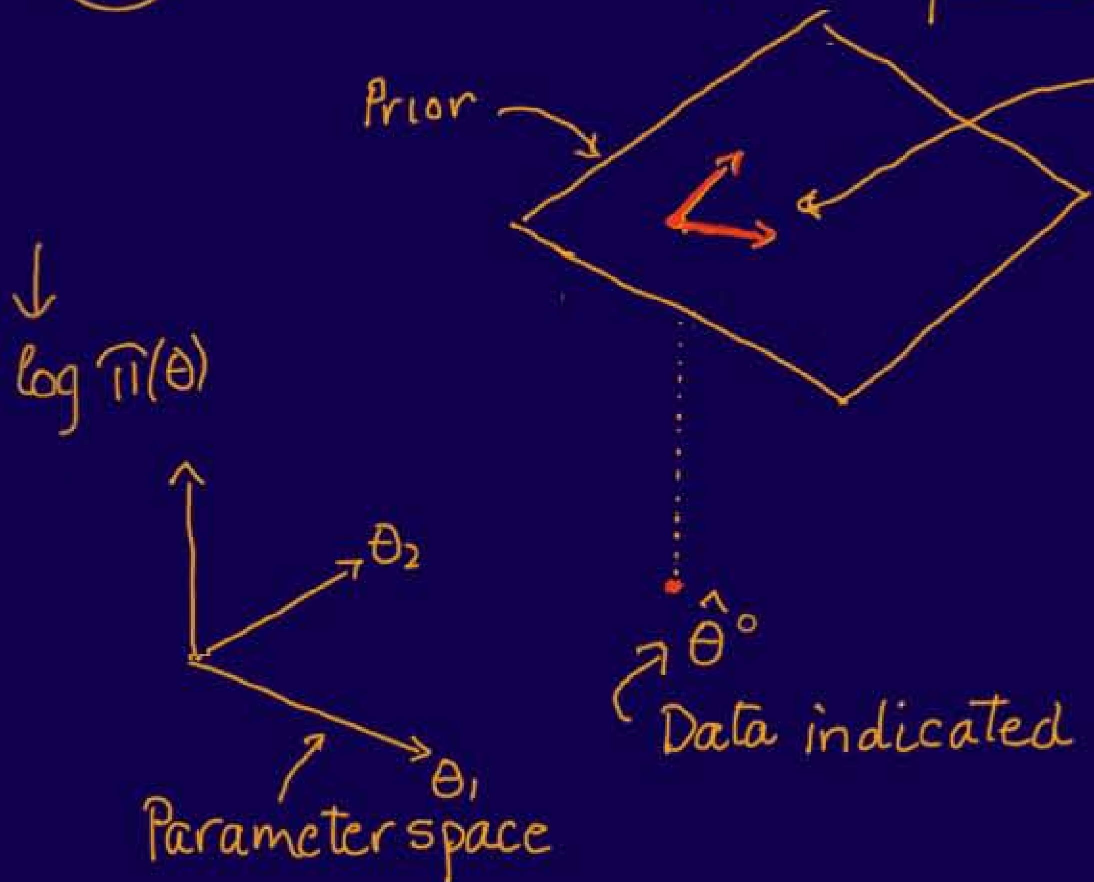
③ Gradient of the prior regularity



③ Gradient of the prior regularity



③ Gradient of the prior regularity

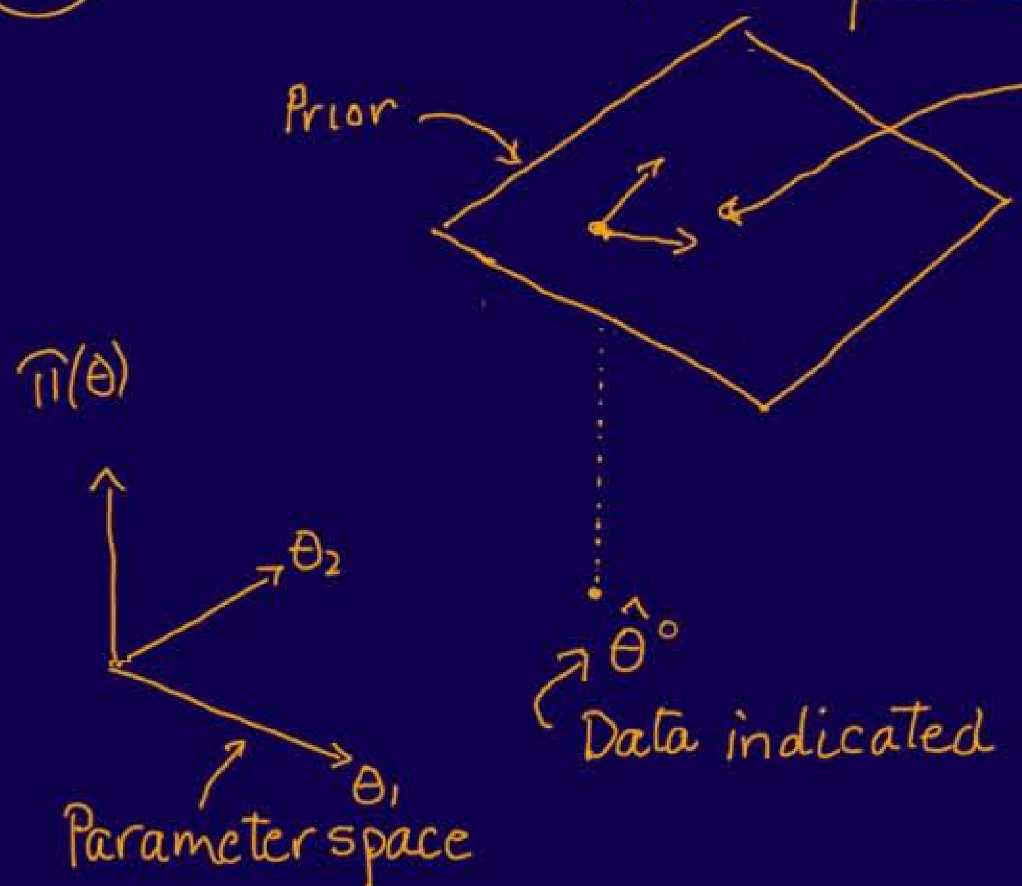


Gradient = $\underset{\sim}{a} = \frac{\partial}{\partial \theta} \log \pi(\theta) \Big|_{\hat{\theta}^{\circ}}$

log ?

$\pi(\theta)$, $c\pi(\theta)$

③ Gradient of the prior regularity

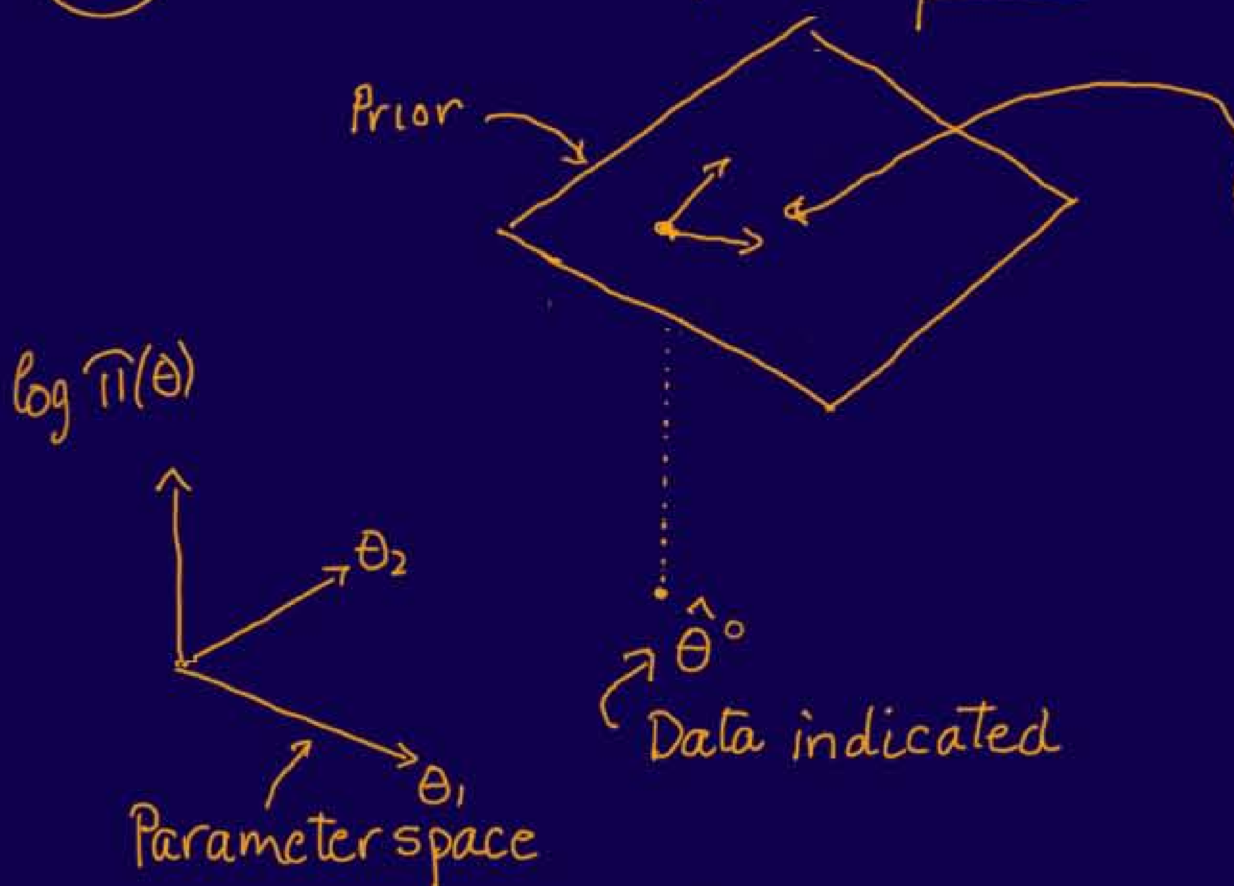


Gradient = $\underset{\sim}{a} = \frac{\partial}{\partial \theta} \log \pi(\theta) \Big|_{\hat{\theta}}$ log ? $\pi(\theta)$, $c\pi(\theta)$

What does it tell you?

a lot! $O(n^{-1})$

③ Gradient of the prior regularity



Gradient = $\underset{\sim}{a} = \frac{\partial}{\partial \theta} \log \pi(\theta) \Big|_{\hat{\theta}}$ log ? $\pi(\theta)$, $c\pi(\theta)$

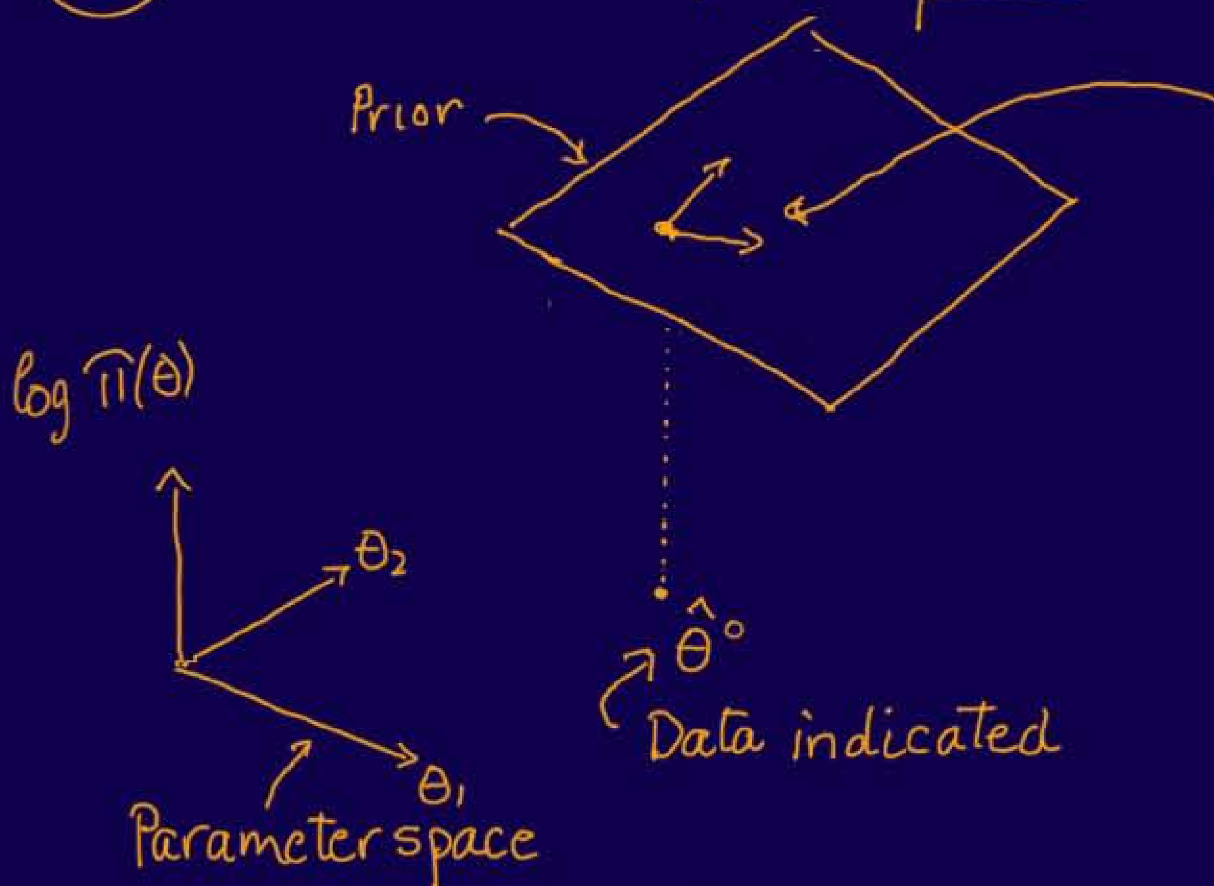
What does it tell you?

a lot! $O(n^{-1})$

Pure case

$N(\hat{\theta}^{\circ}; 1)$

③ Gradient of the prior regularity

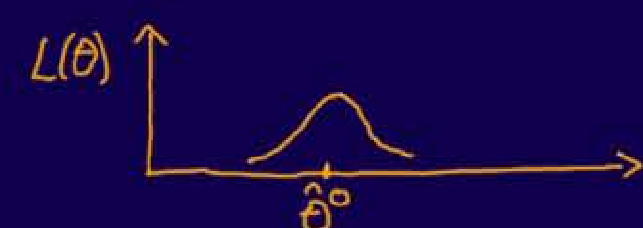


Gradient = $\underset{\sim}{a} = \frac{\partial}{\partial \theta} \log \pi(\theta) \Big|_{\hat{\theta}}$ log ? $\pi(\theta)$, $c\pi(\theta)$

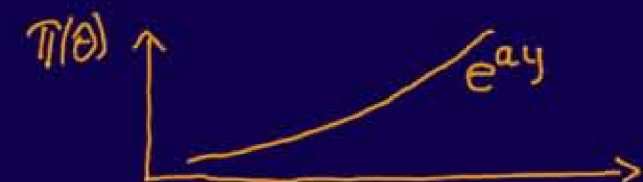
What does it tell you?

a lot! $O(n^{-1})$

Pure case

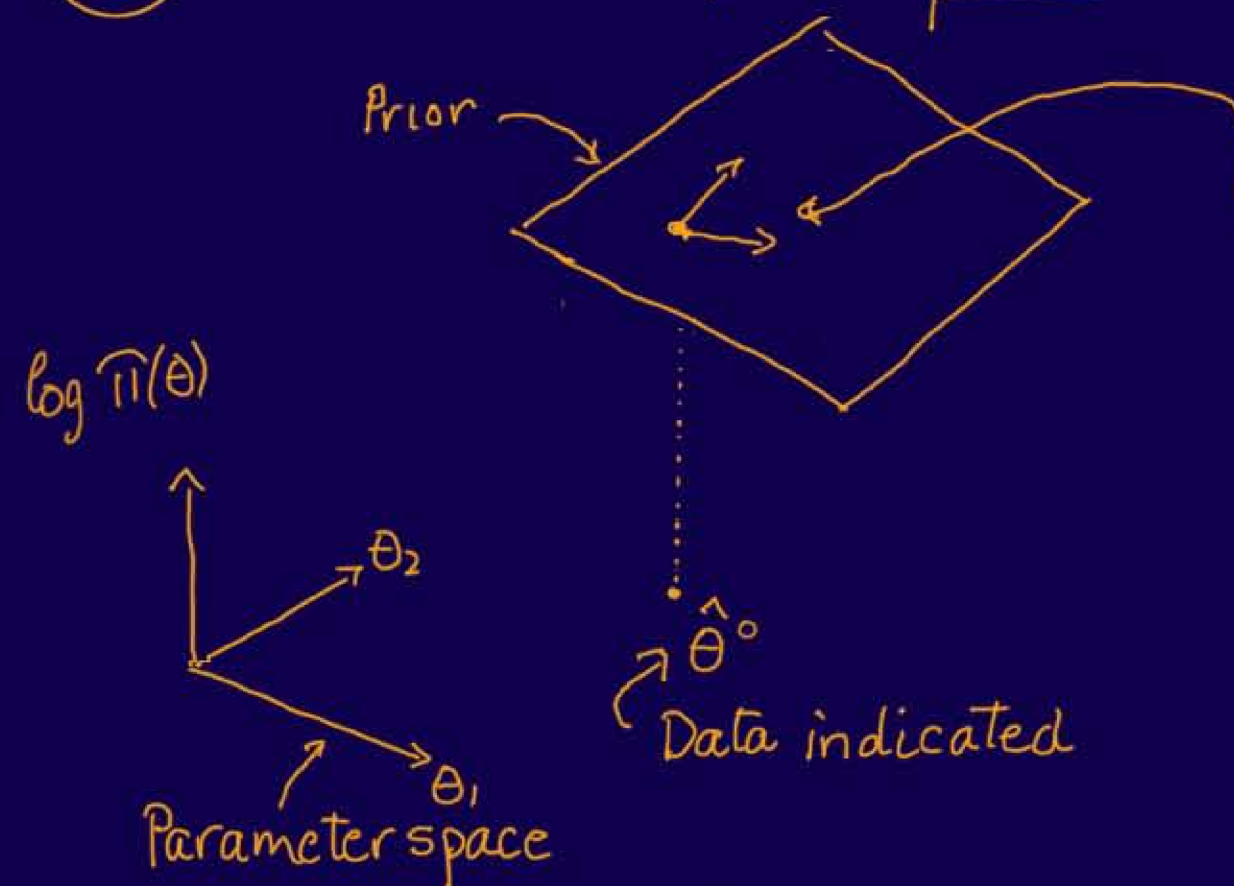


$N(\hat{\theta}^{\circ}; 1)$



e^{ay} = Exp't'l tilt

③ Gradient of the prior regularity

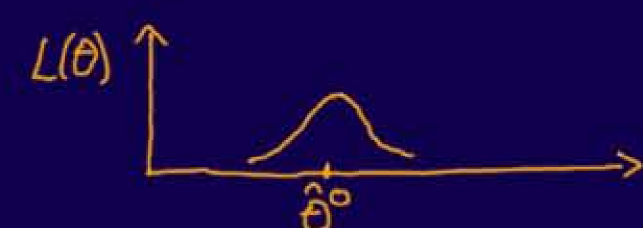


Gradient = $\underset{\sim}{a} = \frac{\partial}{\partial\theta} \log \pi(\theta)\Big|_{\hat{\theta}}$ log ? $\pi(\theta)$, $c\pi(\theta)$

What does it tell you?

a lot! $O(n^{-1})$

Pure case

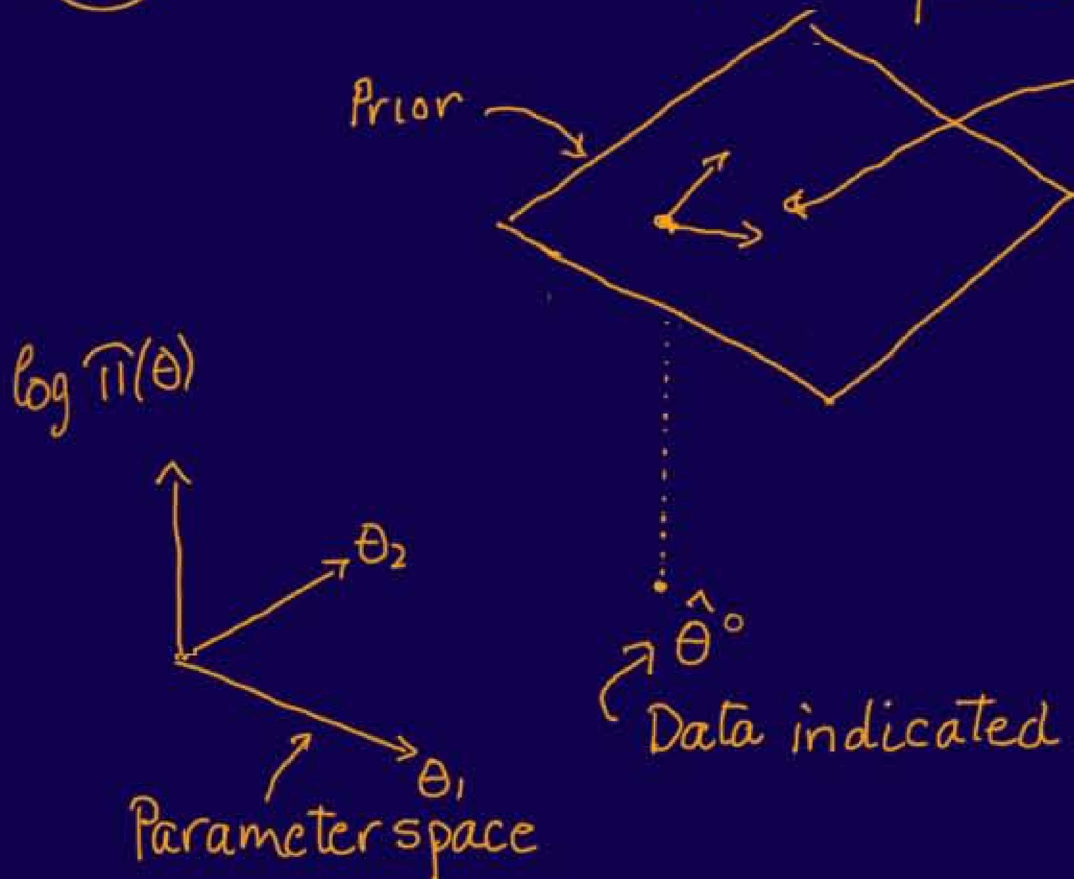


$N(\hat{\theta}^\circ; 1)$

e^{ay} = Exp't'l tilt

$N(\hat{\theta}^\circ + a; 1)$ Displaced Likelihood

③ Gradient of the prior   regularity

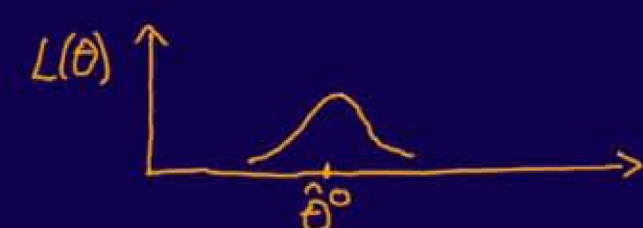


Gradient = $\underset{\sim}{a} = \frac{\partial}{\partial \theta} \log \pi(\theta) \big|_{\hat{\theta}}$ log ? $\pi(\theta)$, $c\pi(\theta)$

What does it tell you?

a lot! $O(n^{-1})$

<u>Pure case</u>



$N(\hat{\theta}^{\circ}; 1)$

e^{ay} = Exp't'l tilt

$N(\hat{\theta}^{\circ} + a; 1)$ Displaced Likelihood

<u>Widely</u> available ⇒ calculate <u>bias</u> $\underset{\sim}{a}$, but need to calibrate θ

(4) Standardize & Calibrate θ

$f(y;\theta)$ regular, data y^0

(a) Standardize: Centre at obs. mle $\hat{\theta}^0$ (Approx by Exp'l model)

Scale by obs. info. $\hat{J}^0_{\theta\theta}$

In new θ: Obs log likelihood $\ell^0(\theta) = -\theta^2/2 + O(n^{-1/2})$

log-prior $\log \pi(\theta) = a\theta/n^{1/2} + O(n^{-1})$

(b) Linearize How θ affects model at y^0 ... Crucial! (near slice)

Differentiate ... model at y^0; use "Dist'n fns" or "Quantiles" ~~Densities~~

Examine at y^0 $\quad \underset{\text{at } y^0}{dy} = \underset{n\times p}{V(\theta)}\, \underset{p\times 1}{d\theta}$ $\quad V(\theta) = \left(\frac{\partial y}{\partial \theta}\right)^0$ Fixed p-values Like "$X_{n\times p}$"

$\underset{\text{at } y^0}{d\hat{\theta}} = \underset{p\times p}{\bar{V}(\theta)}\, d\theta$ $\quad \bar{V}(\theta)$ from $V(\theta)$ easy general

Gives: Location parameterization $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$ Integration, $W = \bar{V}_\theta(\hat{\theta}^0)$

Location model at $\hat{\theta}^0 \pm d\hat{\theta}$ $\quad O(n^{-1})$ Can see what's there!
(use θ for β now)

get log-likelihood $\ell(\theta;\hat{\theta}) = -\frac{1}{2}(\theta-\hat{\theta})^2 - \alpha_3(\theta-\hat{\theta})^3/6n^{1/2} + O(n^{-1})$

JRSSB 2010 F Reid Marras Yi $\quad$ Bernoulli 2010 FFS

(4) Standardize & Calibrate θ

$f(y;\theta)$ regular, data y^0

(a) Standardize: Centre at obs. mle $\hat{\theta}^0$ ⇒ (Approx by Exp'l model)

Scale by obs. info. $\hat{j}^0_{\theta\theta}$

In new θ: Obs log likelihood $\ell^0(\theta) = -\theta^2/2 + O(n^{-1/2})$

log-prior $\log \pi(\theta) = a\theta/n^{1/2} + O(n^{-1})$

(b) Linearize How θ affects model at y^0... Crucial! (near slice)

Differentiate ... model at y^0; use "Dist'n fns" or "Quantiles" ~~Densities~~

Examine at y^0 $\underset{\text{at } y^0}{dy} = \underset{n\times p}{V(\theta)} \underset{p\times 1}{d\theta}$ $V(\theta) = \left(\frac{\partial y}{\partial \theta}\right)^0$ Fixed p-values Like "$\underset{n\times p}{X}$"

$\underset{\text{at } y^0}{d\hat{\theta}} = \underset{p\times p}{\bar{V}(\theta)}\, d\theta$ $\bar{V}(\theta)$ from $V(\theta)$ easy general

Gives: Location parameterization $\beta = I\theta + \theta' W\theta/2n^{1/2} +$ Integration, $W = \bar{V}_\theta(\hat{\theta})$

Location model at $\hat{\theta}^0 \pm d\hat{\theta}$ $O(n^{-1})$ Can see what's there!
(use θ for β now)

get log-likelihood $\ell(\theta;\hat{\theta}) = -\frac{1}{2}(\theta-\hat{\theta})^2 - \alpha_3(\theta-\hat{\theta})^3/6n^{1/2} + O(n^{-1})$

JRSSB 2010 F Reid Marras Yi Bernoulli 2010 FFS

(4) Standardize & Calibrate θ

$f(y;\theta)$ regular, data y^0

(a) Standardize: Centre at obs. mle $\hat{\theta}^0$ (Approx by Exp'l model)

Scale by obs. info. $\hat{j}^0_{\theta\theta}$

In new θ: Obs log likelihood $\ell^0(\theta) = -\theta^2/2 + O(n^{-1/2})$

log-prior $\log \pi(\theta) = a\theta/n^{1/2} + O(n^{-1})$

(b) Linearize How θ affects model at y^0 ... Crucial! (near slice)

Differentiate ... model at y^0; use "Dist'n fns" or "Quantiles" ~~Densities~~

Examine at y^0 $\quad \underset{\text{at } y^0}{dy} = \underset{n \times p}{V(\theta)} \, \underset{p \times 1}{d\theta}$ $\quad V(\theta) = \left(\frac{\partial y}{\partial \theta}\right)^0$ Fixed p-values Like "$\underset{n \times p}{X}$"

$\underset{\text{at } y^0}{d\hat{\theta}} = \underset{p \times p}{\bar{V}(\theta)} \, d\theta$ $\quad \bar{V}(\theta)$ from $V(\theta)$ easy general

Gives: Location parameterization $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$ Integration, $W = \bar{V}_\theta(\hat{\theta}^0)$

Location model at $\hat{\theta}^0 \pm d\hat{\theta}$ $\quad O(n^{-1})$ Can see what's there!

(use θ for β now)

get log-likelihood $\ell(\theta; \hat{\theta}) = -\frac{1}{2}(\theta - \hat{\theta})^2 - \alpha_3 (\theta - \hat{\theta})^3 / 6n^{1/2} + O(n^{-1})$

JRSSB 2010 F Reid Marras Yi $\quad$ Bernoulli 2010 FFS

④ Standardize & Calibrate θ

$f(y;\theta)$ regular, data y^0

(a) Standardize: Centre at obs. mle $\hat{\theta}^0$ (Approx by Expt'l model)

Scale by obs. info. $\hat{J}^0_{\theta\theta}$

In new θ: Obs log likelihood $\ell^0(\theta) = -\theta^2/2 + O(n^{-1/2})$

log-prior $\log \pi(\theta) = a\theta/n^{1/2} + O(n^{-1})$

(b) Linearize How θ affects model at y^0... Crucial!

Differentiate ... model at y^0; use "Dist'n fns" or "Quantiles" ~~Densities~~

Examine at y^0 $\quad \underset{\text{at } y^0}{dy} = \underset{n\times p}{V(\theta)}\, \underset{p\times 1}{d\theta}$ $\quad V(\theta) = \left(\frac{\partial y}{\partial \theta}\right)^0$ Fixed p-values Like "$X_{n\times p}$"

$\underset{\text{at } y^0}{d\hat{\theta}} = \underset{p\times p}{\bar{V}(\theta)}\, d\theta$ $\quad \bar{V}(\theta)$ from $V(\theta)$ easy general

Gives: Location parameterization $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$ Integration, $W = \bar{V}_\theta(\hat{\theta}^0)$

Location model at $\hat{\theta}^0 \pm d\hat{\theta}$ $\quad O(n^{-1})$ Can see what's there!

(use θ for β now)

get log-likelihood $\ell(\theta;\hat{\theta}) = -\frac{1}{2}(\theta-\hat{\theta})^2 - \alpha_3(\theta-\hat{\theta})^3/6n^{1/2} + O(n^{-1})$

JRSSB 2010 F Reid Marras Yi $\quad$ Bernoulli 2010 FFS

④ Standardize & Calibrate θ

$f(y;\theta)$ regular, data y^0

(a) Standardize: Centre at obs. mle $\hat{\theta}^0$

Scale by obs. info. $\hat{J}^0_{\theta\theta}$

(Approx by Expt'l model)

In new θ: Obs log likelihood $\ell^0(\theta) = -\theta^2/2 + O(n^{-1/2})$

log-prior $\log \pi(\theta) = a\theta/n^{1/2} + O(n^{-1})$

(b) Linearize How θ affects model at y^0... Crucial! (near slice)

"Differentiate ... model" at y^0; use "~~Dist~~ fns" or "Quantiles" ~~Densities~~

Examine at y^0 $\underset{\text{at } y^0}{dy} = \underset{n\times p}{V(\theta)}\,\underset{p\times 1}{d\theta}$ $V(\theta) = \left(\frac{\partial y}{\partial \theta}\right)^0$ Fixed p-values Like "$\underset{n\times p}{X}$"

$\underset{\text{at } y^0}{d\hat{\theta}} = \underset{p\times p}{\bar{V}(\theta)}\, d\theta$ $\bar{V}(\theta)$ from $V(\theta)$ easy general

Gives: Location parameterization $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$ Integration, $W = \bar{V}_\theta(\hat{\theta}^0)$

Location model at $\hat{\theta}^0 \pm d\hat{\theta}$ $O(n^{-1})$ Can see what's there!

(use θ for β now)

get log-likelihood $\ell(\theta;\hat{\theta}) = -\frac{1}{2}(\theta-\hat{\theta})^2 - \alpha_3(\theta-\hat{\theta})^3/6n^{1/2} + O(n^{-1})$

JRSSB 2010 F Reid Marras Yi Bernoulli 2010 FFS

(4) Standardize & Calibrate θ

$f(y;\theta)$ regular, data y^0

(a) Standardize: Centre at obs. mle $\hat{\theta}^0$ (Approx. by Exp'l model)

Scale by obs. info. $\hat{j}^0_{\theta\theta}$

In new θ: Obs log likelihood $\ell^0(\theta) = -\theta^2/2 + O(n^{-1/2})$

log-prior $\log \pi(\theta) = a\theta/n^{1/2} + O(n^{-1})$

(b) Linearize How θ affects model at y^0... Crucial! (near slice)

Differentiate ... model at y^0; use "Dist~~n~~ fns" or "Quantiles" Den~~si~~ties

Examine at y^0

$$\underset{\text{at } y^0}{dy} = \underset{n\times p}{V(\theta)}\, \underset{p\times 1}{d\theta}$$

$V(\theta) = \left(\dfrac{\partial y}{\partial \theta}\right)^0$ Fixed p-values Like "$\underset{n\times p}{X}$"

$$\underset{\text{at } y^0}{d\hat{\theta}} = \underset{p\times p}{\bar{V}(\theta)}\, \underset{p\times 1}{d\theta}$$

$\bar{V}(\theta)$ from $V(\theta)$ (easy general)

Gives: Location parameterization $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$ Integration, $W = \bar{V}_\theta(\hat{\theta}^0)$

Location model at $\hat{\theta}^0 \pm d\hat{\theta}$ $O(n^{-1})$ Can see what's there!

(use θ for β now)

get log-likelihood $\ell(\theta;\hat{\theta}) = -\frac{1}{2}(\theta-\hat{\theta})^2 - \alpha_3(\theta-\hat{\theta})^3/6n^{1/2} + O(n^{-1})$

JRSSB 2010 F Reid Marras Yi   Bernoulli 2010 FFS

(4) Standardize & Calibrate θ

$f(y;\theta)$ regular, data y^0

(a) Standardize: Centre at obs. mle $\hat{\theta}^0$ (Approx. by Expt'l model)

Scale by obs. info. $\hat{J}^0_{\theta\theta}$

In new θ: Obs log likelihood $\ell^0(\theta) = -\theta^2/2 + O(n^{-1/2})$

log-prior $\log \pi(\theta) = a\theta/n^{1/2} + O(n^{-1})$

(b) Linearize How θ affects model at y^0... Crucial! (near slice)

Differentiate ... model at y^0; use "Dist ~~fns~~" or "Quantiles" Den~~sities~~

Examine at y^0 $\underset{\text{at } y^0}{dy} = \underset{n\times p}{V(\theta)}\, \underset{p\times 1}{d\theta}$ $V(\theta) = \left(\frac{\partial y}{\partial \theta}\right)^0$ Fixed p-values Like "$\underset{n\times p}{X}$"

$\underset{\text{at } y^0}{d\hat{\theta}} = \underset{p\times p}{\bar{V}(\theta)}\, d\theta$ $\bar{V}(\theta)$ from $V(\theta)$ easy general

Gives: Location parameterization $\beta = I\theta + \theta' W \theta / 2n^{1/2} + \ldots$ Integration, $W = \bar{V}_\theta(\hat{\theta}^0)$

Location model at $\hat{\theta}^0 \pm d\hat{\theta}$ $O(n^{-1})$ Can see what's there!

(use θ for β now)

get log-likelihood $\ell(\theta;\hat{\theta}) = -\frac{1}{2}(\theta-\hat{\theta})^2 - \alpha_3(\theta-\hat{\theta})^3/6n^{1/2} + O(n^{-1})$

JRSSB 2010 F Reid Marras Yi Bernoulli 2010 FFS

(4) Standardize & Calibrate θ

$f(y;\theta)$ regular, data y^0

(a) Standardize: Centre at obs. mle $\hat{\theta}^0$ (Approx. by Exp'l model)

Scale by obs. info. $\hat{J}^0_{\theta\theta}$

In new θ: Obs log likelihood $\ell^0(\theta) = -\theta^2/2 + O(n^{-1/2})$

log-prior $\log \pi(\theta) = a\theta/n^{1/2} + O(n^{-1})$

(b) Linearize How θ affects model at y^0... Crucial! (near slice)

Differentiate ... model at y^0; use "Dist~~n~~ fns" or "Quantiles" Den~~s~~ities

Examine at y^0 $\underset{\text{at } y^0}{dy} = \underset{n\times p}{V(\theta)} \underset{p\times 1}{d\theta}$ $V(\theta) = \left(\frac{\partial y}{\partial \theta}\right)^0$ Fixed p-values Like "$\underset{n\times p}{X}$"

$\underset{\text{at } y^0}{d\hat{\theta}} = \underset{p\times p}{\bar{V}(\theta)}\, d\theta$ $\bar{V}(\theta)$ from $V(\theta)$ easy general

Gives: Location parameterization $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$ Integration, $W = \bar{V}_\theta(\hat{\theta}^0)$

Location model at $\hat{\theta}^0 \pm d\hat{\theta}$ ← β $O(n^{-1})$ Can see what's there!

(Use θ for β now)

get log-likelihood $\ell(\theta;\hat{\theta}) = -\frac{1}{2}(\theta-\hat{\theta})^2 - \alpha_3(\theta-\hat{\theta})^3/6n^{1/2} + O(n^{-1})$

(4) Standardize & Calibrate θ

$f(y;\theta)$ regular, data y^o

(a) Standardize: Centre at obs. mle $\hat{\theta}^o$ (Approx. by Expt'l model)
Scale by obs. info. $\hat{J}^o_{\theta\theta}$

In new θ: Obs log likelihood $\ell^o(\theta) = -\theta^2/2 + O(n^{-1/2})$
log-prior $\log \pi(\theta) = a\theta/n^{1/2} + O(n^{-1})$

(b) Linearize How θ affects model at y^o... Crucial! (near slice)

Differentiate ... model at y^o; use "~~Dist fns~~" or "Quantiles" ~~Densities~~

Examine at y^o $\underset{\text{at } y^o}{dy} = \underset{n\times p}{V(\theta)}\, \underset{p\times 1}{d\theta}$ $V(\theta) = \left(\frac{\partial y}{\partial \theta}\right)^o$ Fixed p-values Like "$\underset{n\times p}{X}$"

$\underset{\text{at } y^o}{d\hat{\theta}} = \underset{p\times p}{\bar{V}(\theta)}\, d\theta$ $\bar{V}(\theta)$ from $V(\theta)$ easy general

Gives: Location parameterization $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$ Integration, $W = \bar{V}_\theta(\hat{\theta}^o)$

Location model at $\hat{\theta}^o \pm d\hat{\theta}$ $O(n^{-1})$ Can see what's there!
(Use θ for β now)

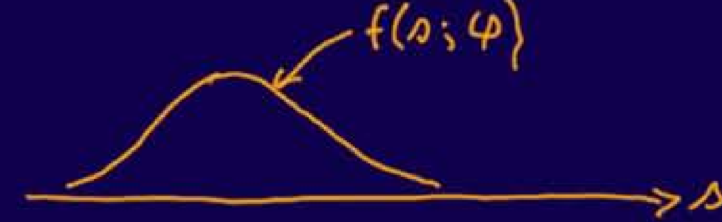
get log-likelihood $\ell(\theta;\hat{\theta}) = -\frac{1}{2}(\theta-\hat{\theta})^2 - \alpha_3(\theta-\hat{\theta})^3/6n^{1/2} + O(n^{-1})$

JRSSB (2010) F Reid Marras Yi; Bernoulli (2010) F F Staicu; F F F (2010) J Stat. R | xxx = 239 240 249

(5) Origins: Welch & Peers 1963 (Jeffreys, 1946; B, 1763) JRSSB

Scalar exponential model; scalar

$$f(s;\varphi) = \exp\{\varphi s + k(\varphi)\}\, h(s)$$



Used $\pi(\varphi) = j^{1/2}(\varphi)$ root info information $= j(\varphi) = -\ell_{\varphi\varphi}(\varphi; y)$

$\Rightarrow$ Bayes = Confidence i.e. Bayes is reproducible!

Deeper: $\beta(\varphi) = \int^{\varphi} j^{1/2}(\varphi)\, d\varphi$ "Constant info reparameterization"

...gives Location model prescient..

Extend: Any regular model; vector y, θ

$$V = \left(\frac{\partial \hat{y}}{\partial \theta}\right)^{0} \Rightarrow \varphi(\theta) = \left(\frac{\partial \ell}{\partial V}\right)^{0}$$

Treat as exponential $f(s;\varphi)$ ... just a tool 2nd/3rd

$\Rightarrow$ p-values $p(\varphi)$ available to $O(n^{-3/2})$ 3rd ..

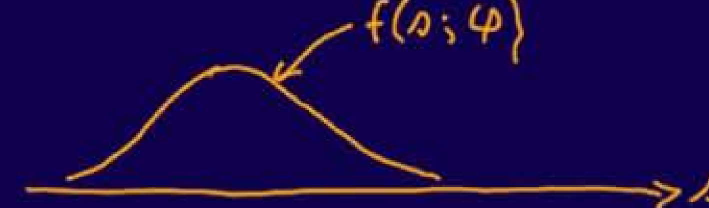
Also $\Rightarrow$ Linear parameter $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$ 2nd

F Reid Wu. Bka 1999 Reid, F Bka 2010 xxx = 196, 251

(5) Origins: Welch & Peers 1963 (Jeffreys, 1946) JRSSB

Scalar exponential model

$$f(s;\varphi) = \exp\{\varphi s + k(\varphi)\}\, h(s)$$



Used $\pi(\varphi) = j^{1/2}(\varphi)$ root info information $= j(\varphi) = -\ell_{\varphi\varphi}(\varphi; y)$

$\Rightarrow$ Bayes = Confidence i.e. Bayes is reproducible!

Deeper: $\beta(\varphi) = \int^{\varphi} j^{1/2}(\varphi)\, d\varphi$ "Constant info reparameterization"

...gives Location model prescient..

Extend: Any regular model; vector y, θ

$$V = \left(\frac{\partial \hat{y}}{\partial \theta}\right)^{\circ} \Rightarrow \varphi(\theta) = \left(\frac{\partial \ell}{\partial V}\right)^{\circ}$$

Treat as exponential $f(s;\varphi)$ ... just a tool 2nd/3rd

$\Rightarrow$ p-values $p(\varphi)$ available to $O(n^{-3/2})$ 3rd ..

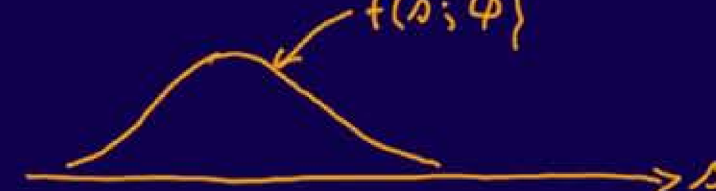
Also $\Rightarrow$ Linear parameter $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$ 2nd

F Reid Wu Bka 1999 Reid, F Bka 2010 xxx = 196, 251

(5) Origins: Welch & Peers 1963 (Jeffreys, 1946) JRSSB

Scalar exponential model

$f(s;\varphi) = \exp\{\varphi s + k(\varphi)\} h(s)$



Used $\pi(\varphi) = j^{1/2}(\varphi)$ root info information = $j(\varphi) = -\ell_{\varphi\varphi}(\varphi; y)$

$\Rightarrow$ Bayes = Confidence i.e. Bayes is reproducible!

Deeper: $\beta(\varphi) = \int^{\varphi} j^{1/2}(\varphi)\, d\varphi$ "Constant info reparameterization"

...gives Location model prescient ..

Extend: Any regular model; vector y, θ

$V = (\frac{\partial y}{\partial \theta})^{\circ} \Rightarrow \varphi(\theta) = (\frac{\partial \ell}{\partial V})^{\circ}$

Treat as exponential $f(s;\varphi)$ ... just a tool 2nd/3rd

$\Rightarrow$ p-values $p(\varphi)$ available to $O(n^{-3/2})$ 3rd ..

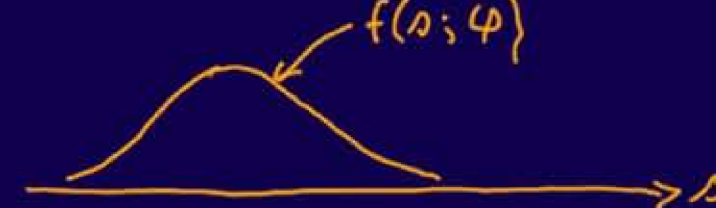
Also $\Rightarrow$ Linear parameter $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$ 2nd

F Reid Wu Bka 1999 Reid, F Bka 2010 xxx = 196, 251

(5) Origins: Welch & Peers 1963 (Jeffreys, 1946) JRSSB

Scalar exponential model

$f(s;\varphi) = \exp\{\varphi s + k(\varphi)\} h(s)$



Used $\pi(\varphi) = j^{1/2}(\varphi)$ root info information $= j(\varphi) = -\ell_{\varphi\varphi}(\varphi; y)$

$\Rightarrow$ Bayes = Confidence i.e. Bayes is reproducible!

Deeper: $\beta(\varphi) = \int^{\varphi} j^{1/2}(\varphi)\, d\varphi$ "Constant info reparameterization"

...gives Location model prescient..

Extend: Any regular model; vector y, θ

$V = \left(\frac{\partial y}{\partial \theta}\right)^{\circ} \Rightarrow \varphi(\theta) = \left(\frac{\partial \ell}{\partial V}\right)^{\circ}$

Treat as exponential $f(s;\varphi)$ ... just a "tool" 2nd / 3rd

$\Rightarrow$ p-values $p(\varphi)$ available to $O(n^{-3/2})$ 3rd

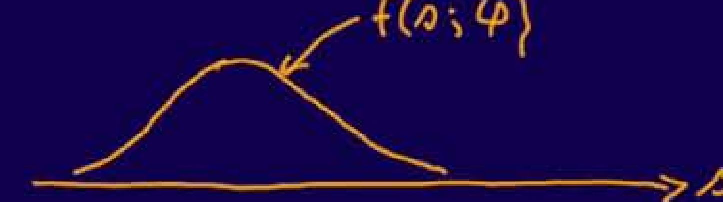
Also $\Rightarrow$ Linear parameter $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$ 2nd

F Reid Wu. Bka 1999 Reid, F Bka 2010 xxx = 196, 251

(5) <u>Origins</u>: Welch & Peers 1963 (Jeffreys, 1946) JRSSB

<u>Scalar</u> exponential model

$f(s;\varphi) = \exp\{\varphi s + k(\varphi)\}\, h(s)$



Used $\pi(\varphi) = j^{1/2}(\varphi)$ root info information $= j(\varphi) = -\ell_{\varphi\varphi}(\varphi; y)$

$\Rightarrow$ Bayes = Confidence i.e. Bayes is <u>reproducible</u>!

<u>Deeper</u>: $\beta(\varphi) = \int^{\varphi} j^{1/2}(\varphi)\, d\varphi$ "Constant info reparameterization"

...gives <u>Location</u> model prescient..

<u>Extend</u>: <u>Any</u> <u>regular</u> model; vector y, θ

$V = (\widehat{\frac{\partial y}{\partial \theta}})^{\circ} \Rightarrow \varphi(\theta) = (\frac{\partial \ell}{\partial V})^{\circ}$

Treat as <u>exponential</u> $f(s;\varphi)$ ... just a tool 2nd / 3rd

$\Rightarrow$ p-values $p(\varphi)$ available to $O(n^{-3/2})$ 3rd ..

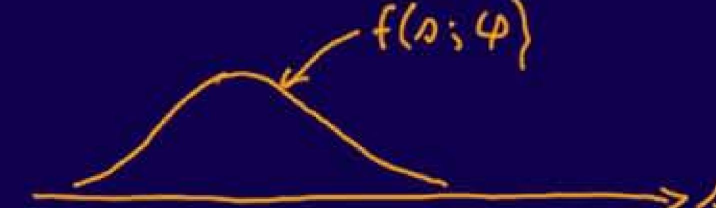
Also $\Rightarrow$ Linear parameter $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$ 2nd

F Reid Wu Bka 1999 Reid, F Bka 2010 xxx = 196, 251

(5) <u>Origins</u>: Welch & Peers 1963 (Jeffreys, 1946) JRSSB

<u>Scalar</u> exponential model

$f(s;\varphi) = \exp\{\varphi s + k(\varphi)\}\, h(s)$



Used $\pi(\varphi) = j^{1/2}(\varphi)$ root info information $= j(\varphi) = -\ell_{\varphi\varphi}(\varphi; y)$

$\Rightarrow$ Bayes = Confidence i.e. Bayes is <u>reproducible</u>!

<u>Deeper</u>: $\beta(\varphi) = \int^{\varphi} j^{1/2}(\varphi)\, d\varphi$ "Constant info reparameterization"

...gives <u>Location</u> model prescient..

<u>Extend</u>: <u>Any</u> <u>regular</u> model; vector y, θ

$V = \left(\frac{\partial \hat{y}}{\partial \theta}\right)^{\circ} \Rightarrow \varphi(\theta) = \left(\frac{\partial \ell}{\partial V}\right)^{\circ}$

Treat as <u>exponential</u> $f(s;\varphi)$ ... just a tool 2nd/3rd

$\Rightarrow$ p-values $p(\varphi)$ available to $O(n^{-3/2})$ $\bar{V}_\theta$ 3rd ..

Also $\Rightarrow$ Linear parameter $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$ 2nd

F Reid Wu Bka 1999 Reid, F Bka 2010 xxx = 196, 251

⑥ Gradient of the prior; fine tuning

Centered, Std'zd Coords: to see what's happening...

Prior: $\log \pi(\theta) = \Sigma a_i \theta_i / n^{1/2} = \underset{\sim}{a}' \theta / \sqrt{n} \quad + \quad O(n^{-1})$ ... 2nd order

Likelihood: $\ell(\theta) = -\frac{1}{2} \Sigma \theta_i^2 - \frac{1}{6 n^{1/2}} \Sigma \alpha_{ijk} \theta_i \theta_j \theta_k$ " "

(In linear para β) $\ell(\theta; \hat{\theta}) = -\frac{1}{2} \Sigma (\theta_i - \hat{\theta}_i)^2 - \frac{1}{6 n^{1/2}} \Sigma \alpha_{ijk} (\theta_i - \hat{\theta}_i)(\cdot)(\cdot)$ ... 1st deriv at y^0

Posterior (In linear para β) $\log \pi(\theta | y^0) = -\frac{1}{2} \Sigma (\theta_i - a_i / n^{1/2})^2 - \frac{1}{6 n^{1/2}} \Sigma \alpha_{ijk} (\theta_i - a_i / n^{1/2})(\theta_j - a_j / n^{1/2})(\theta_k - a_k / n^{1/2})$

Bayes is confidence/is Reproducible if θ is "location/linear"

But (sneaker) Prior (re β) $\log \pi_\beta(\theta) = \log \pi(\theta) + \log \left| \frac{\partial \theta}{\partial \beta} \right|$

$$\frac{\partial \beta}{\partial \theta} = \frac{\partial}{\partial \theta} \left\{ I (\theta - \hat{\theta}^0) + \frac{1}{2} (\theta - \hat{\theta}^0) \underset{\sim}{W} (\theta - \hat{\theta}^0) \right\} = I + W(\theta - \hat{\theta}^0) \quad \leftarrow n^{-1/2}$$

$$\log \left| \frac{\partial \theta}{\partial \beta} \right| = -\mathrm{Tr}\, \underset{\sim}{W} (\theta - \hat{\theta}^0) = -\underset{\sim}{c}' (\theta - \hat{\theta}^0) \qquad O(n^{-1})$$

Reproducible if $\underset{\sim}{a} - \underset{\sim}{c} = 0$ — Prior moves $L(\underset{\sim}{\theta})$ to $L(\underset{\sim}{\theta} - \underset{\sim}{a} + \underset{\sim}{c})$ — Std'zed β coords

- $\underset{\sim}{c}$: Effect of curvature
- $\underset{\sim}{a}$: Tilt in prior

} must compensate

(6) Gradient of the prior; fine tuning

Centered, Std'zd Coord's: to see what's happening...

Prior: $\log \pi(\theta) = \sum a_i \theta_i / n^{1/2} = \underset{\sim}{a}' \theta / \sqrt{n} + O(n^{-1})$ ... 2nd order

---

Likelihood: $\ell(\theta) = -\frac{1}{2}\sum \theta_i^2 - \frac{1}{6n^{1/2}} \sum \alpha_{ijk} \theta_i \theta_j \theta_k$ " "

(in linear para β) $\ell(\theta; \hat{\theta}) = -\frac{1}{2}\sum (\theta_i - \hat{\theta}_i)^2 - \frac{1}{6n^{1/2}} \sum \alpha_{ijk} (\theta_i - \hat{\theta}_i)(\cdot)(\cdot)$ ... 1st deriv at y^0

Posterior (in linear para β) $\log \pi(\theta | y^0) = -\frac{1}{2}\sum (\theta_i - a_i/n^{1/2})^2 - \frac{1}{6n^{1/2}} \sum \alpha_{ijk} (\theta_i - a_i/n^{1/2})(\theta_j - a_j/n^{1/2})(\theta_k - a_k/n^{1/2})$

Bayes is confidence/is Reproducible if θ is "location/linear"

But Prior (re β) $\log \pi_\beta(\theta) = \log \pi(\theta) + \log \left| \frac{\partial \theta}{\partial \beta} \right|$

sneaker

$\frac{\partial \beta}{\partial \theta} = \frac{\partial}{\partial \theta} \left\{ I(\theta - \hat{\theta}^0) + \frac{1}{2} (\theta - \hat{\theta}^0) \underset{\sim}{W} (\theta - \hat{\theta}^0) \right\} = I + W(\theta - \hat{\theta}^0)$ ← $n^{-1/2}$

$\log \left| \frac{\partial \theta}{\partial \beta} \right| = -\mathrm{tr}\, \underset{\sim}{W} (\theta - \hat{\theta}^0) = -\underset{\sim}{c}' (\theta - \hat{\theta}^0)$ $O(n^{-1})$

Reproducible if $\underset{\sim}{a} - \underset{\sim}{c} = 0$ Prior moves $L(\underset{\sim}{\theta})$ to $L(\underset{\sim}{\theta} - \underset{\sim}{a} + \underset{\sim}{c})$ Std'zed β coords

- Effect of curvature } must compensate
- Tilt in prior

⑥ Gradient of the prior; fine tuning

Centered, Std'zd Coord's: to see what's happening...

Prior: $\log \pi(\theta) = \Sigma a_i \theta_i / n^{1/2} = \underset{\sim}{a}'\theta/\sqrt{n} + O(n^{-1})$ ... 2nd order

Likelihood: $\ell(\theta) = -\frac{1}{2}\Sigma\theta_i^2 - \frac{1}{6n^{1/2}}\Sigma\alpha_{ijk}\theta_i\theta_j\theta_k$ " "

(in linear para β) $\ell(\theta;\hat{\theta}) = -\frac{1}{2}\Sigma(\theta_i - \hat{\theta}_i)^2 - \frac{1}{6n^{1/2}}\Sigma\alpha_{ijk}(\theta_i - \hat{\theta}_i)(\cdot)(\cdot)$ ... 1st deriv at y^0

Posterior (in linear para β) $\log \pi(\theta|y^0) = -\frac{1}{2}\Sigma(\theta_i - a_i/n^{1/2})^2 - \frac{1}{6n^{1/2}}\Sigma\alpha_{ijk}(\theta_i - a_i/n^{1/2})(\theta_j - a_j/n^{1/2})(\theta_k - a_k/n^{1/2})$

Bayes is confidence/is Reproducible if θ is "location/linear"

But (sneaker) Prior (re β) $\log \pi_\beta(\theta) = \log \pi(\theta) + \log\left|\frac{\partial\theta}{\partial\beta}\right|$

$$\frac{\partial\beta}{\partial\theta} = \frac{\partial}{\partial\theta}\left\{ I(\theta - \hat{\theta}^0) + \frac{1}{2}(\theta - \hat{\theta}^0)\underset{\sim}{W}(\theta - \hat{\theta}^0)\right\} = I + W(\theta - \hat{\theta}^0) \quad \leftarrow n^{-1/2}$$

$$\log\left|\frac{\partial\theta}{\partial\beta}\right| = -\mathrm{Tr}\,\underset{\sim}{W}(\theta - \hat{\theta}^0) = -\underset{\sim}{c}'(\theta - \hat{\theta}^0) \qquad O(n^{-1})$$

Reproducible if $\underset{\sim}{a} - \underset{\sim}{c} = 0$ Prior moves $L(\underset{\sim}{\theta})$ to $L(\underset{\sim}{\theta} - \underset{\sim}{a} + \underset{\sim}{c})$ Std'zed β coords

- $\underset{\sim}{c}$: Effect of curvature
- $\underset{\sim}{a}$: Tilt in prior

} must compensate

⑥ Gradient of the prior; fine tuning

Centered, Std'z'd Coord's: to see what's happening...

Prior: $\log \pi(\theta) = \sum a_i \theta_i / n^{1/2} = \underset{\sim}{a}' \theta / \sqrt{n} \quad + \quad O(n^{-1})$ ... 2nd order

Likelihood: $\ell(\theta) = -\frac{1}{2} \sum \theta_i^2 - \frac{1}{6n^{1/2}} \sum \alpha_{ijk} \theta_i \theta_j \theta_k$ " "

(in linear para β) $\ell(\theta; \hat{\theta}) = -\frac{1}{2} \sum (\theta_i - \hat{\theta}_i)^2 - \frac{1}{6n^{1/2}} \sum \alpha_{ijk} (\theta_i - \hat{\theta}_i)(\cdot)(\cdot)$ ... 1st deriv at y^0

Posterior (in linear para β) $\log \pi(\theta | y^0) = -\frac{1}{2} \sum (\theta_i - a_i/n^{1/2})^2 - \frac{1}{6n^{1/2}} \sum \alpha_{ijk} (\theta_i - a_i/n^{1/2})(\theta_j - a_j/n^{1/2})(\theta_k - a_k/n^{1/2})$

Bayes is confidence/is Reproducible if θ is "location/linear"

But (sneaker) Prior (re β) $\log \pi_\beta(\theta) = \log \pi(\theta) + \log \left| \frac{\partial \theta}{\partial \beta} \right|$

$$\frac{\partial \beta}{\partial \theta} = \frac{\partial}{\partial \theta} \left\{ I(\theta - \hat{\theta}^0) + \frac{1}{2} (\theta - \hat{\theta}^0) \underset{\sim}{W} (\theta - \hat{\theta}^0) \right\} = I + W(\theta - \hat{\theta}^0) \quad \leftarrow n^{-1/2}$$

$$\log \left| \frac{\partial \theta}{\partial \beta} \right| = -\mathrm{Tr}\, \underset{\sim}{W} (\theta - \hat{\theta}^0) = -\underset{\sim}{c}' (\theta - \hat{\theta}^0) \qquad O(n^{-1})$$

Reproducible if $\underset{\sim}{a} - \underset{\sim}{c} = 0$ Prior moves $L(\underset{\sim}{\theta})$ to $L(\underset{\sim}{\theta} - \underset{\sim}{a} + \underset{\sim}{c})$ Std'zed β coords

- Effect of curvature
- Tilt in prior

} must compensate

⑥ Gradient of the prior; fine tuning

Centered, Std'zd Coord's: to see what's happening...

Prior: $\log \pi(\theta) = \Sigma a_i\theta_i / n^{1/2} = \underset{\sim}{a}'\theta/\sqrt{n} \quad + \quad O(n^{-1})$ ... 2nd order

Likelihood: $\ell(\theta) = -\frac{1}{2}\Sigma\theta_i^2 - \frac{1}{6n^{1/2}}\Sigma\alpha_{ijk}\theta_i\theta_j\theta_k$ " "

(in linear para β) $\ell(\theta;\hat\theta) = -\frac{1}{2}\Sigma(\theta_i - \hat\theta_i)^2 - \frac{1}{6n^{1/2}}\Sigma\alpha_{ijk}(\theta_i - \hat\theta_i)(\cdot)(\cdot)$ ... 1st deriv at y^0

Posterior (in linear para β) $\log \pi(\theta|y^0) = -\frac{1}{2}\Sigma(\theta_i - a_i/n^{1/2})^2 - \frac{1}{6n^{1/2}}\Sigma\alpha_{ijk}(\theta_i - a_i/n^{1/2})(\theta_j - a_j/n^{1/2})(\theta_k - a_k/n^{1/2})$

Bayes is confidence/is Reproducible if θ is "location/linear"

But sneaker Prior (re β) $\log\pi_\beta(\theta) = \log\pi(\theta) + \log\left|\frac{\partial\theta}{\partial\beta}\right|$

$\frac{\partial\beta}{\partial\theta} = \frac{\partial}{\partial\theta}\left\{ I(\theta - \hat\theta^0) + \frac{1}{2}(\theta-\hat\theta^0)\underset{\sim}{W}(\theta-\hat\theta^0)\right\} = I + W(\theta - \hat\theta^0)$ ← $n^{-1/2}$

$\log\left|\frac{\partial\theta}{\partial\beta}\right| = -\mathrm{Tr}\,\underset{\sim}{W}(\theta-\hat\theta^0) = -\underset{\sim}{c}'(\theta-\hat\theta^0)$ $\quad O(n^{-1})$

Reproducible if $\underset{\sim}{a} - \underset{\sim}{c} = 0$ Prior moves $L(\underset{\sim}{\theta})$ to $L(\underset{\sim}{\theta} - \underset{\sim}{a} + \underset{\sim}{c})$ Std'zed β coords

- $\underset{\sim}{c}$: Effect of curvature
- $\underset{\sim}{a}$: Tilt in prior

} must compensate

(6) Gradient of the prior; fine tuning

Centered, Std'zd Coord's: to see what's happening...

Prior: $\log \pi(\theta) = \sum a_i \theta_i / n^{1/2} = \underset{\sim}{a}'\theta/\sqrt{n} \quad + \quad O(n^{-1})$ ... 2nd order

Likelihood: $\ell(\theta) = -\frac{1}{2}\sum \theta_i^2 - \frac{1}{6n^{1/2}}\sum \alpha_{ijk}\theta_i\theta_j\theta_k$ " "

(in linear para β) $\ell(\theta;\hat\theta) = -\frac{1}{2}\sum(\theta_i - \hat\theta_i)^2 - \frac{1}{6n^{1/2}}\sum \alpha_{ijk}(\theta_i - \hat\theta_i)(\cdot)(\cdot)$ ... 1st deriv at y^0

Posterior (in linear para β) $\log \pi(\theta|y^0) = -\frac{1}{2}\sum(\theta_i - a_i/n^{1/2})^2 - \frac{1}{6n^{1/2}}\sum\alpha_{ijk}(\theta_i - a_i/n^{1/2})(\theta_j - a_j/n^{1/2})(\theta_k - a_k/n^{1/2})$

Bayes is confidence/is Reproducible if θ is "location/linear"

But (sneaker) Prior (re β) $\log \pi_\beta(\theta) = \log \pi(\theta) + \log\left|\frac{\partial\theta}{\partial\beta}\right|$

$$\frac{\partial\beta}{\partial\theta} = \frac{\partial}{\partial\theta}\left\{ I(\theta-\hat\theta^0) + \frac{1}{2}(\theta-\hat\theta^0)\underset{\sim}{W}(\theta-\hat\theta^0)\right\} = I + W(\theta - \hat\theta^0) \quad \leftarrow n^{-1/2}$$

$$\log\left|\frac{\partial\theta}{\partial\beta}\right| = -\mathrm{tr}\, \underset{\sim}{W}(\theta-\hat\theta^0) = -\underset{\sim}{c}'(\theta-\hat\theta^0) \qquad O(n^{-1})$$

Reproducible if $\underset{\sim}{a} - \underset{\sim}{c} = 0$ Prior moves $L(\underset{\sim}{\theta})$ to $L(\underset{\sim}{\theta} - \underset{\sim}{a} + \underset{\sim}{c})$ Std'zed β coords

- $\underset{\sim}{c}$: Effect of curvature
- $\underset{\sim}{a}$: Tilt in prior

} must compensate

⑥ Gradient of the prior; fine tuning

Centered, Std'zd Coords: to see what's happening...

Prior: $\log \pi(\theta) = \sum a_i \theta_i / n^{1/2} = a'\theta/\sqrt{n} + O(n^{-1})$ ... 2nd order

Likelihood: $\ell(\theta) = -\frac{1}{2}\sum \theta_i^2 - \frac{1}{6n^{1/2}} \sum \alpha_{ijk}\theta_i\theta_j\theta_k$ " "

(in linear para β) $\ell(\theta;\hat\theta) = -\frac{1}{2}\sum(\theta_i - \hat\theta_i)^2 - \frac{1}{6n^{1/2}}\sum \alpha_{ijk}(\theta_i - \hat\theta_i)(\cdot)(\cdot)$ ... 1st deriv at y^0

Posterior (in linear para β) $\log \pi(\theta|y^0) = -\frac{1}{2}\sum(\theta_i - a_i/n^{1/2})^2 - \frac{1}{6n^{1/2}}\sum\alpha_{ijk}(\theta_i - a_i/n^{1/2})(\theta_j - a_j/n^{1/2})(\theta_k - a_k/n^{1/2})$

Bayes is confidence/is Reproducible if θ is "location/linear"

But (sneaker) Prior (re β) $\log \pi_\beta(\theta) = \log \pi(\theta) + \log\left|\frac{\partial\theta}{\partial\beta}\right|$

$$\frac{\partial\beta}{\partial\theta} = \frac{\partial}{\partial\theta}\left\{ I(\theta-\hat\theta^0) + \frac{1}{2}(\theta-\hat\theta^0)W(\theta-\hat\theta^0)\right\} = I + W(\theta - \hat\theta^0) \quad \leftarrow n^{-1/2}$$

$$\log\left|\frac{\partial\theta}{\partial\beta}\right| = -\mathrm{tr}\, W(\theta-\hat\theta^0) = -c'(\theta-\hat\theta^0) \qquad O(n^{-1})$$

Reproducible if $a - c = 0$ Prior moves $L(\theta)$ to $L(\theta - a + c)$ Std'zed β coords

- a: Tilt in prior
- c: Effect of curvature

} must compensate

(6) Gradient of the prior; fine tuning

Centered, Std'zd Coord's: to see what's happening...

Prior: $\log \pi(\theta) = \Sigma a_i \theta_i / n^{1/2} = \underset{\sim}{a}' \theta / \sqrt{n} \quad + \quad O(n^{-1})$ ... 2nd order

Likelihood: $\ell(\theta) = -\frac{1}{2}\Sigma \theta_i^2 - \frac{1}{6n^{1/2}} \Sigma \alpha_{ijk} \theta_i \theta_j \theta_k$ " "

(in linear para β) $\ell(\theta; \hat\theta) = -\frac{1}{2}\Sigma(\theta_i - \hat\theta_i)^2 - \frac{1}{6n^{1/2}} \Sigma \alpha_{ijk} (\theta_i - \hat\theta_i)(\cdot)(\cdot)$ ... 1st deriv at y^0

Posterior (in linear para β) $\log \pi(\theta | y^0) = -\frac{1}{2}\Sigma(\theta_i - a_i/n^{1/2})^2 - \frac{1}{6n^{1/2}} \Sigma \alpha_{ijk} (\theta_i - a_i/n^{1/2})(\theta_j - a_j/n^{1/2})(\theta_k - a_k/n^{1/2})$

Bayes is confidence/is Reproducible if θ is "location/linear"

But sneaker Prior (re β) $\log \pi_\beta(\theta) = \log \pi(\theta) + \log \left| \frac{\partial \theta}{\partial \beta} \right|$

$$\frac{\partial \beta}{\partial \theta} = \frac{\partial}{\partial \theta} \left\{ I(\theta - \hat\theta^0) + \frac{1}{2}(\theta - \hat\theta^0) \underset{\sim}{W} (\theta - \hat\theta^0) \right\} = I + W(\theta - \hat\theta^0)$$

($W(\theta - \hat\theta^0) \leftarrow n^{-1/2}$)

$$\log \left| \frac{\partial \theta}{\partial \beta} \right| = -\mathrm{Tr}\, \underset{\sim}{W}(\theta - \hat\theta^0) = -\underset{\sim}{c}'(\theta - \hat\theta^0) \qquad O(n^{-1})$$

Reproducible if $\underset{\sim}{a} - \underset{\sim}{c} = 0$ % Prior moves $L(\underset{\sim}{\theta})$ to $L(\underset{\sim}{\theta} - \underset{\sim}{a} + \underset{\sim}{c})$ Stdzed β coords

- $\underset{\sim}{c}$: Effect of curvature
- $\underset{\sim}{a}$: Tilt in prior

} must compensate

(7) Summary

a) Parameter can have curvature; can construct

$$\text{Linear: } \beta = I\theta + \theta' W \theta / 2n^{1/2} +$$

b) In standardized/linear coord's:

Prior directly displaces confidence: $L(\theta)$ to $L\{\theta - (a - c)\}$

bias

c) Posterior is Reproducible / Confidence

if $a - c = 0$ ... if prior "anticipates" curvature

d) Interest in $\psi(\theta) = \Sigma d_i \theta_i$?

Marginal posterior is a direct displacement of confidence:

$$\text{Bias} = d_1(a_1 - c_1) + \cdots + d_p(a_p - c_p)$$

"Use a prior?" $\Rightarrow$ Know the curvature!

⑦ Summary

a) Parameter can have curvature ; can construct

Linear: $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$

b) In standardized/linear coord's:

Prior directly displaces confidence: $L(\underset{\sim}{\theta})$ to $L\{\underset{\sim}{\theta} - (\underset{\sim}{a} - \underset{\sim}{c})\}$

Bias

c) Posterior is Reproducible / Confidence

if $\underset{\sim}{a} - \underset{\sim}{c} = 0$ ... if prior "anticipates" curvature

d) Interest in $\psi(\theta) = \sum d_i \theta_i$?

Marginal posterior is a direct displacement of confidence:

$$\text{Bias} = d_1(a_1 - c_1) + \cdots + d_p(a_p - c_p)$$

"Use a prior?" ⇒ Know the curvature!

7 Summary

a) Parameter can have curvature ; can construct

Linear: $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$

b) In standardized/linear coord's:

Prior directly displaces confidence: $L(\underset{\sim}{\theta})$ to $L\{\underset{\sim}{\theta} - (\underset{\sim}{a} - \underset{\sim}{c})\}$ bias

c) Posterior is Reproducible / Confidence

if $\underset{\sim}{a} - \underset{\sim}{c} = 0$ ... if prior "anticipates" curvature

d) Interest in $\psi(\theta) = \Sigma d_i \theta_i$?

Marginal posterior is a direct displacement of confidence:

$$\text{Bias} = d_1(a_1 - c_1) + \cdots + d_p(a_p - c_p)$$

"Use a prior?" ⇒ Know the curvature!

⑦ Summary

a) Parameter can have curvature; can construct

Linear: $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$

b) In standardized/linear coord's:

Prior directly displaces confidence: $L(\underset{\sim}{\theta})$ to $L\{\underset{\sim}{\theta} - (\underset{\sim}{a} - \underset{\sim}{c})\}$ bias

c) Posterior is Reproducible / Confidence

if $\underset{\sim}{a} - \underset{\sim}{c} = 0$ ... if prior "anticipates" curvature

d) Interest in $\psi(\theta) = \Sigma d_i \theta_i$?

Marginal posterior is a displacement of confidence:

$$\text{Bias} = d_1(a_1 - c_1) + \cdots + d_p(a_p - c_p)$$

"Use a prior?" ⇒ Know the curvature!

(7) Summary

a) Parameter can have <u>curvature</u>; can construct

<u>Linear</u>: $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$

b) In standardized/<u>linear</u> coord's:

Prior directly <u>displaces</u> confidence: $L(\underset{\sim}{\theta})$ to $L\{\underset{\sim}{\theta} - (\underset{\sim}{a} - \underset{\sim}{c})\}$ bias

c) Posterior is Reproducible / Confidence

if $\underset{\sim}{a} - \underset{\sim}{c} = 0$ ... if prior "anticipates" curvature

d) <u>Interest</u> in $\psi(\theta) = \sum d_i \theta_i$?

Marginal posterior is a direct <u>displacement</u> of confidence:

$$\text{Bias} = d_1(a_1 - c_1) + \cdots + d_p(a_p - c_p)$$

"Use a prior?" $\Rightarrow$ <u>Know</u> the curvature!

7 Summary

a) Parameter can have curvature; can construct

Linear: $\beta = I\theta + \theta' W \theta / 2n^{1/2} +$

b) In standardized/linear coord's:

Prior directly displaces confidence: $L(\theta)$ to $L\{\theta - (a - c)\}$ bias

c) Posterior is Reproducible / Confidence

if $a - c = 0$ ... if prior "anticipates" curvature

d) Interest in $\psi(\theta) = \Sigma d_i \theta_i$?

Marginal posterior is a direct displacement of confidence:

$$\text{Bias} = d_1(a_1 - c_1) + \cdots + d_p(a_p - c_p)$$

1) Calc. gradient a
2) Calc. curvature c
3) Prior displaces info. by $a - c$

"Use a prior?" ⇒ Know the curvature!

Thank you

- Ejaz

- the Organizing Committee: Ejaz Ahmed
  Peter Song
  Mu Zhu

- and the supporting Organizations...

Statistical Society of Canada
Société statistique du Canada

NELSON EDUCATION

FIELDS

University of Windsor

MICHIGAN

University of Waterloo

- and

Bayes

Thank you

- Ejaz 

- the Organizing Committee: Ejaz Ahmed
  Peter Song
  Mu Zhu

- and the supporting Organizations...

Statistical Society of Canada
Société statistique du Canada

NELSON EDUCATION

FIELDS

University of Windsor

MICHIGAN

University of Waterloo

- and  Bayes

An addendum? →

But Recall:

Have a prior $\pi(\theta)$

Posterior is right, if $\pi(\theta)$ is true!

---

but wrong otherwise!

Confidence is right, if $\pi(\theta)$ is true

and right otherwise!

$y = X\beta + \sigma z$ β σ^2

$V(\theta) = (X \;\; z^0(\theta)/2\sigma)$

$\tilde{V}(\theta) = \begin{pmatrix} I & \beta - \hat{\beta}^0/2\sigma^2 \\ 0 & \hat{\sigma}^2/\sigma^2 \end{pmatrix}$

$\frac{d\beta\, d\sigma^2}{2\sigma^2}$

$xxx = 247$

But Recall:

Have a prior $\pi(\theta)$

Posterior is right, if $\pi(\theta)$ is true!

but wrong otherwise!

Confidence is right, if $\pi(\theta)$ is true

and right otherwise!

$y = X\beta + \sigma z$ $\quad \beta \; \sigma^2$

$V(\theta) = (X \;\; z^0(\theta)/2\sigma)$ $\quad \frac{d\beta \, d\sigma^2}{2\sigma^2}$

$\tilde{V}(\theta) = \begin{pmatrix} I & \beta - \hat{\beta}^0/2\sigma^2 \\ 0 & \hat{\sigma}^2/\sigma^2 \end{pmatrix}$

$xxx = 247$

But Recall:

Have a prior $\pi(\theta)$

Posterior is right, if $\pi(\theta)$ is true!

but wrong otherwise!

Confidence is right, if $\pi(\theta)$ is true

and right otherwise!

$y = X\beta + \sigma z$ β σ^2

$V(\theta) = (X \quad z^0(\theta)/2\sigma)$

$\tilde{V}(\theta) = \begin{pmatrix} I & \beta - \hat{\beta}^0/2\sigma^2 \\ 0 & \hat{\sigma}^2/\sigma^2 \end{pmatrix}$

$\frac{d\beta \, d\sigma^2}{2\sigma^2}$

$xxx = 247$

But Recall:

Have a prior $\pi(\theta)$

Posterior is right, if $\pi(\theta)$ is true!
but wrong otherwise!

Confidence is right, if $\pi(\theta)$ is true
and right otherwise!

But Recall:

Have a prior $\pi(\theta)$

Posterior is right, if $\pi(\theta)$ is true!
but wrong otherwise!

Confidence is right, if $\pi(\theta)$ is true

and right otherwise!

$y = X\beta + \sigma z$ β σ^2

$V(\theta) = (X \;\; z^0(\theta)/2\sigma)$

$\tilde{V}(\theta) = \begin{pmatrix} I & \beta - \hat{\beta}^0/2\sigma^2 \\ 0 & \hat{\sigma}^2/\sigma^2 \end{pmatrix}$ $\frac{d\beta\, d\sigma^2}{2\sigma^2}$

$xxx = 247$