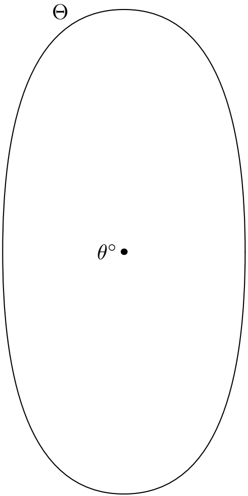
Frequentist and Bayesian methods for ill-posed inverse problems

Oracle and minimax optimality

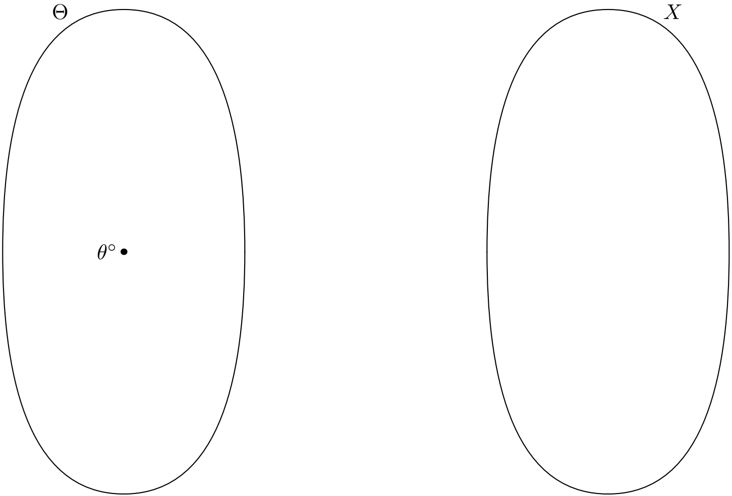
Contents

Contents

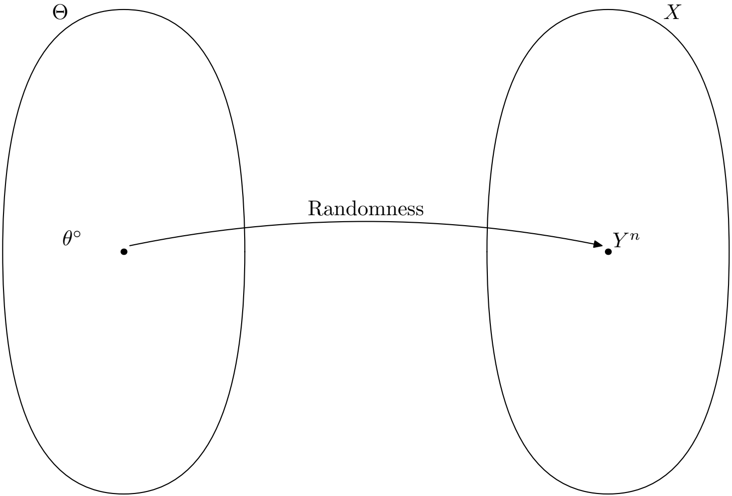
Frequentist estimation, mainstay



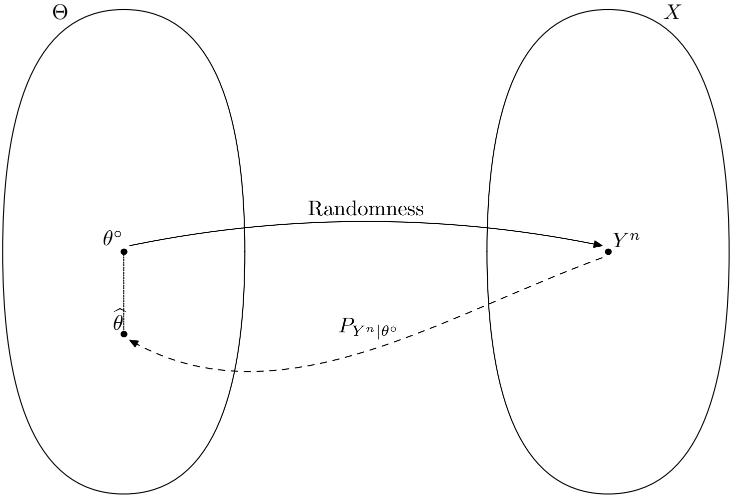
Frequentist estimation, mainstay



Frequentist estimation, mainstay



Frequentist estimation, mainstay



Frequentist estimation, data example

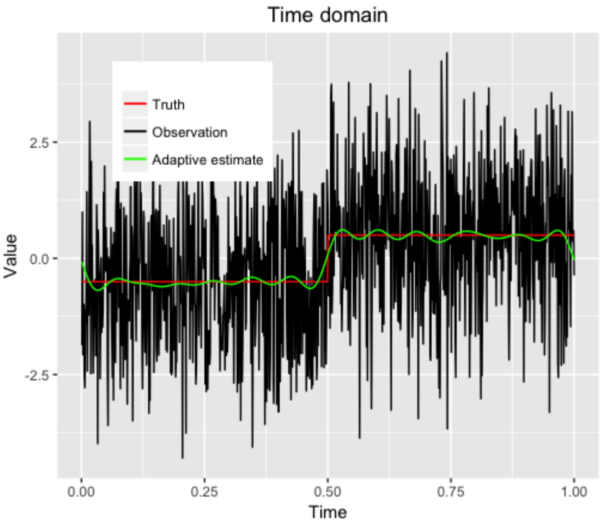
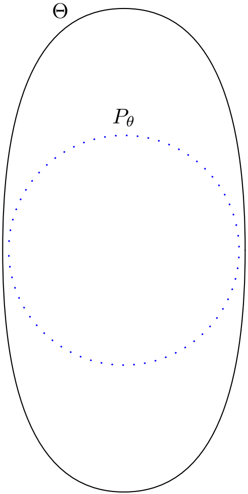


Figure: Estimating the drift in a Gaussian precess by aggregation of projection estimators

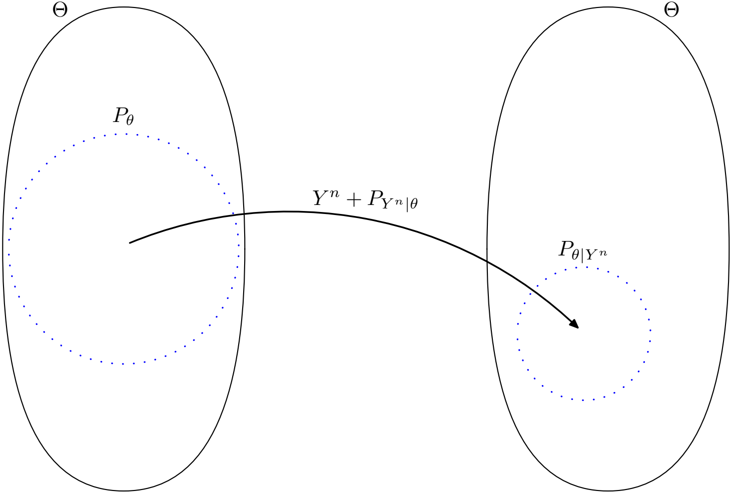
Frequentist estimation, mathematical formulation

$$
\begin{aligned}
dY_x &= \theta^\circ_x dx + \frac{1}{\sqrt{n}} dW(x); \\
Y_j &= \mathscr{F}(Y)(j) = \mathscr{F}(\theta^\circ)(j) + \frac{1}{\sqrt{n}} \xi_j; \\
\xi_j &\sim_{iid} \mathscr{N}(0,1); \\
\widehat{\theta}_j &= \mathbb{1}_{j \leq m} Y_j; \\
\widetilde{\theta}_x &= \mathscr{F}^{-1}(\widehat{\theta})(x).
\end{aligned}
$$

Bayesian estimation, mainstay



Bayesian estimation, mainstay



Bayesian estimation, data example

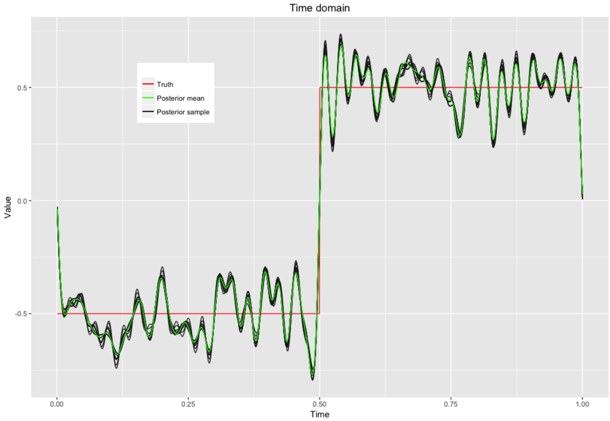
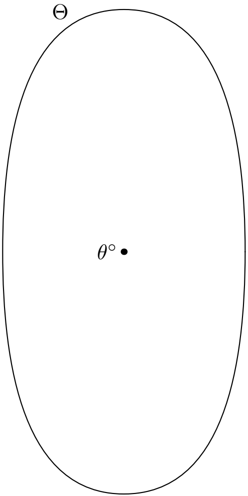


Figure: Sample from the posterior distribution of a hierarchical sieve prior

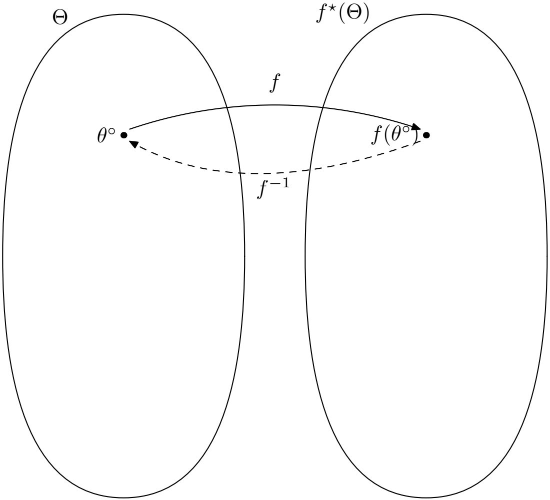
$$
\begin{aligned}
\boldsymbol{\theta}_j \quad &\sim \mathcal{N}(0,1)\mathbb{1}_{j\leq m} + \delta_0 \mathbb{1}_{j>m};\\
Y_j|\boldsymbol{\theta}_j &\sim \mathcal{N}(\boldsymbol{\theta}, \frac{1}{n});\\
\boldsymbol{\theta}_j|Y_j &\sim \mathcal{N}(\widehat{\theta}_j, \sigma_j)\mathbb{1}_{j\leq m} + \delta_0 \mathbb{1}_{j>m};\\
\widehat{\theta}_j \quad &= \frac{n \cdot Y_j}{1+n};\\
\sigma_j \quad &= \frac{1}{1+n}.
\end{aligned}
$$

Contents

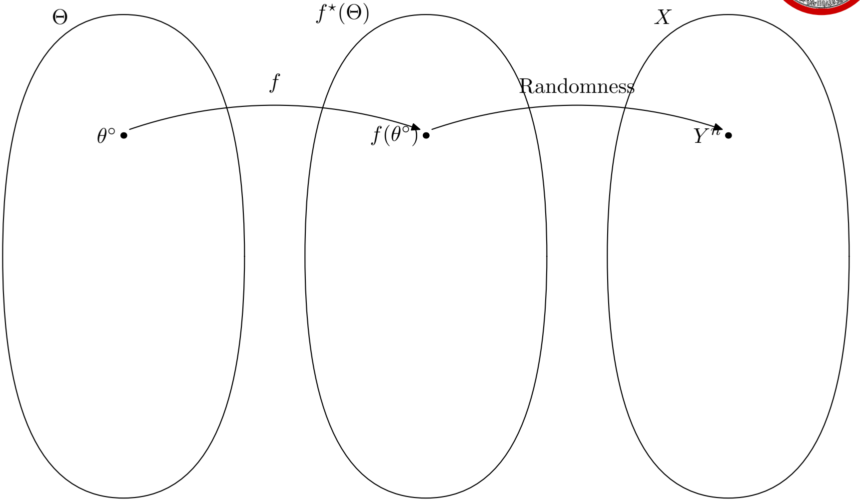
Ill posed-inverse problems, mainstay



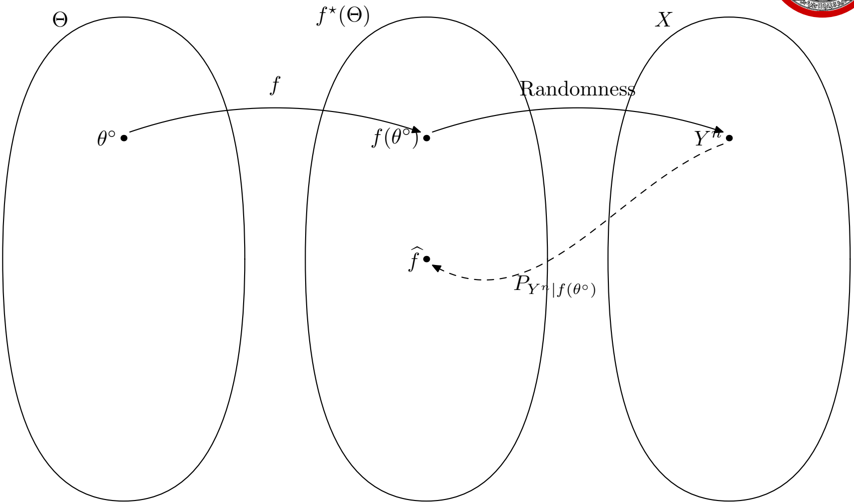
Ill posed-inverse problems, mainstay



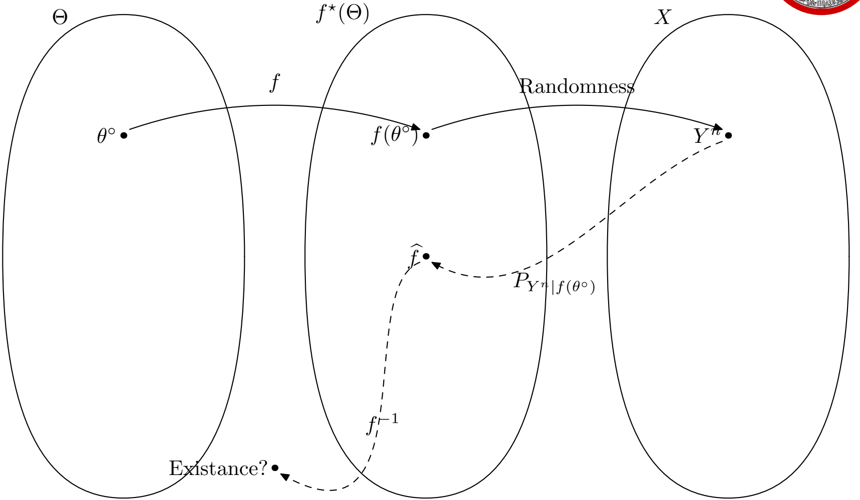
Ill posed-inverse problems, mainstay



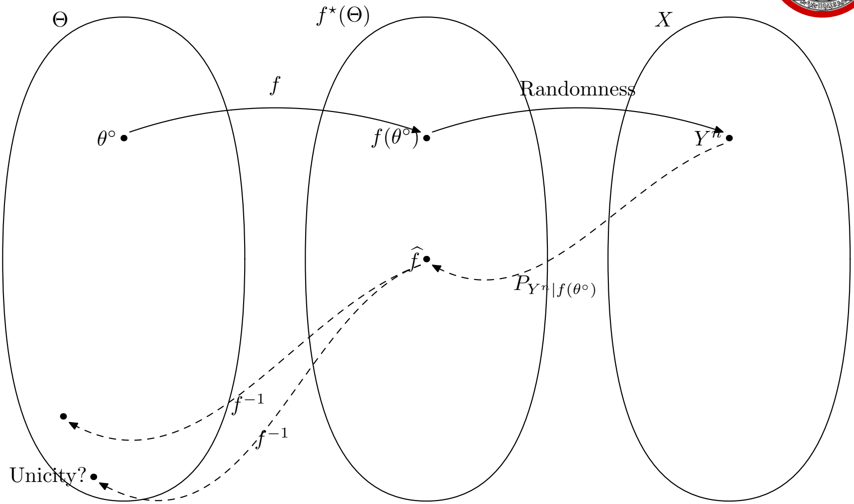
Ill posed-inverse problems, mainstay

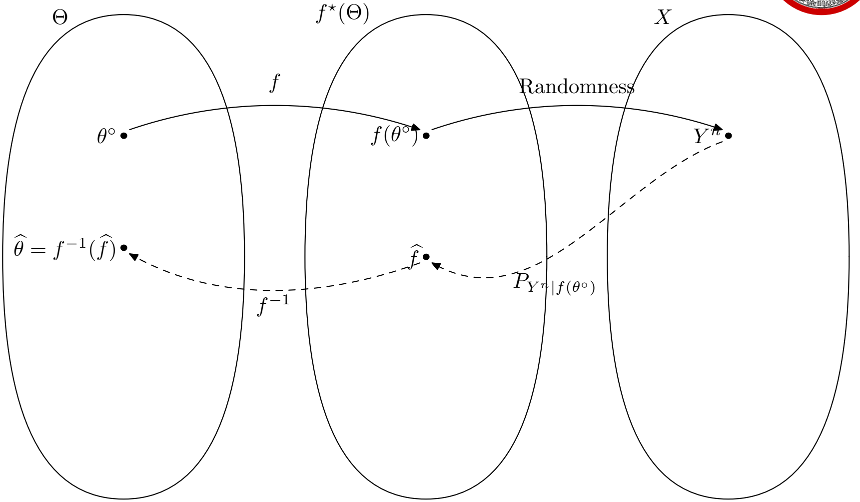


Ill posed-inverse problems, mainstay



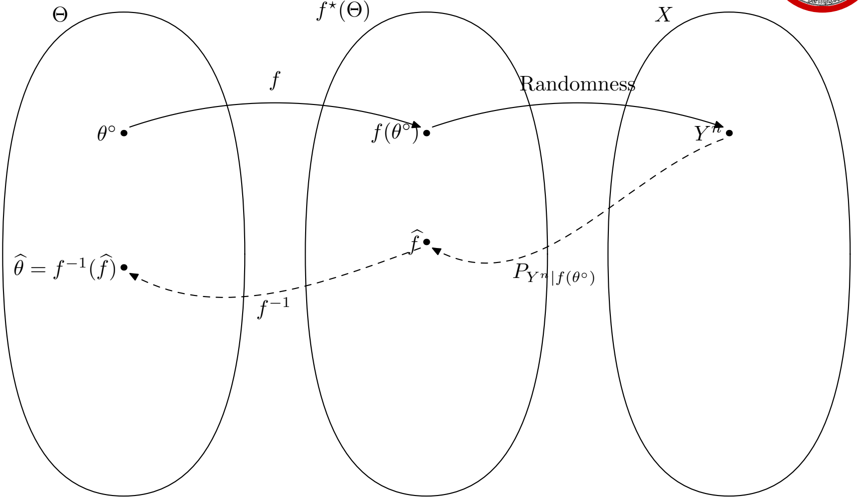
Ill posed-inverse problems, mainstay



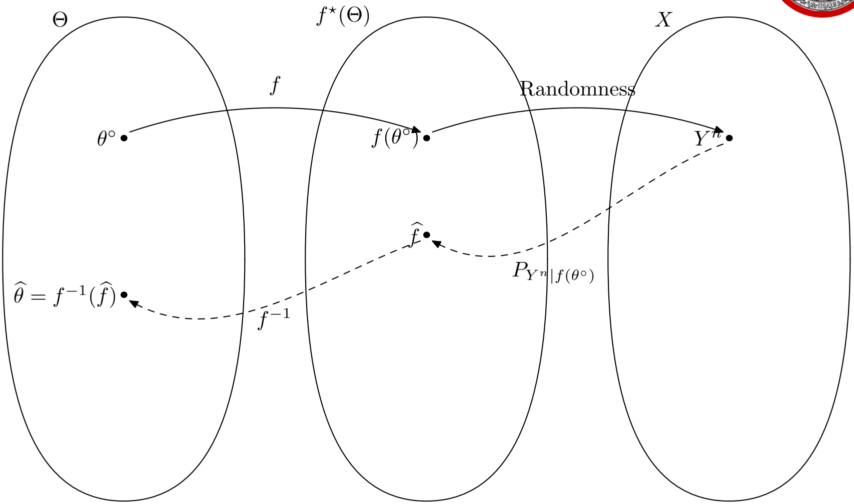


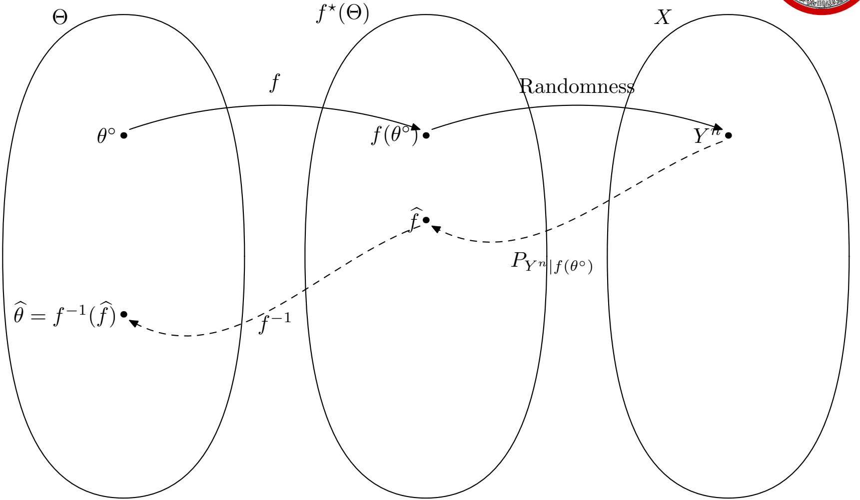
Θ
$f^{\star}(\Theta)$
X
f
Randomness
θ°
$f(\theta^{\circ})$
Y^{n}
$\widehat{\theta} = f^{-1}(\widehat{f})$
$\widehat{f}$
$P_{Y^{n}|f(\theta^{\circ})}$
f^{-1}

Ill posed-inverse problems, mainstay

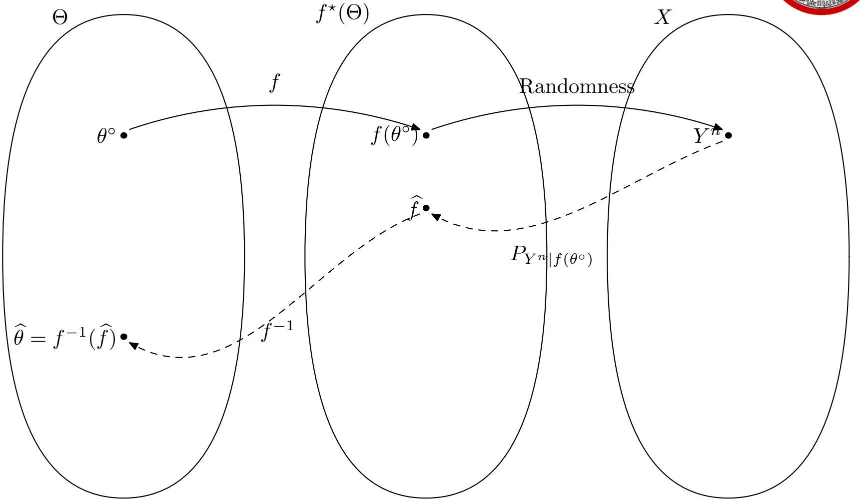


Ill posed-inverse problems, mainstay

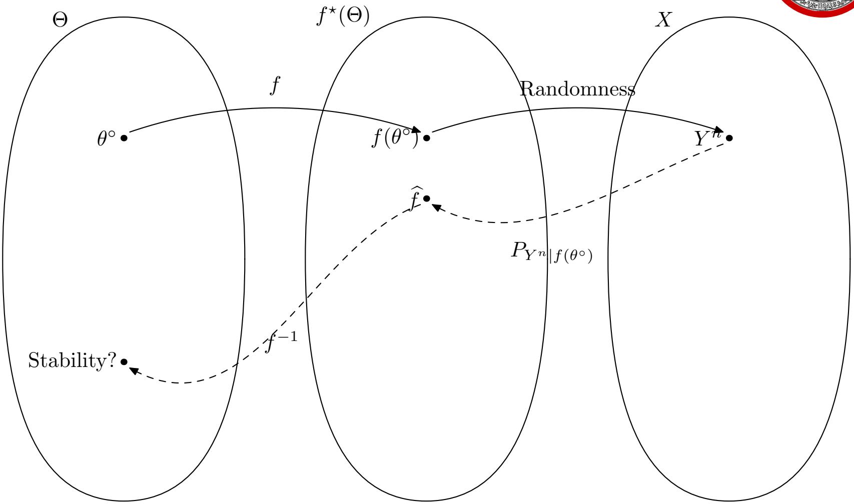


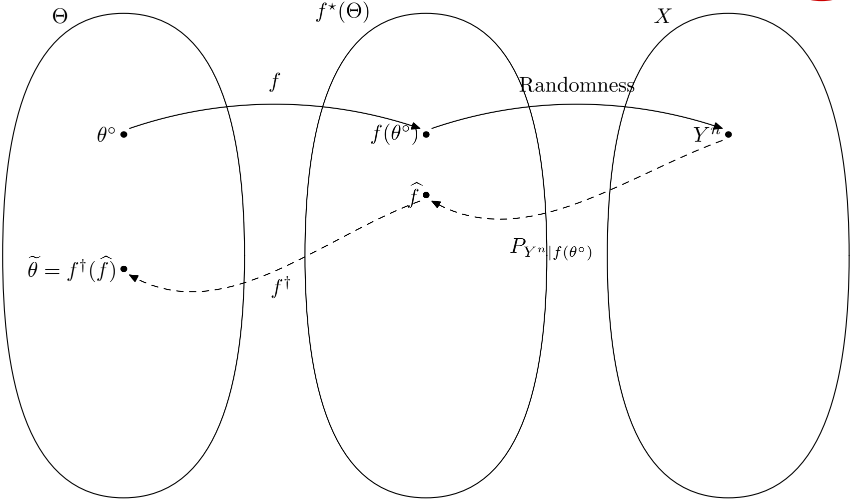

Θ
$f^{\star}(\Theta)$
X
f
Randomness
θ°
$f(\theta^{\circ})$
Y^{n}
$\widehat{f}$
$P_{Y^n|f(\theta^{\circ})}$
$\widehat{\theta} = f^{-1}(\widehat{f})$
f^{-1}

Ill posed-inverse problems, mainstay

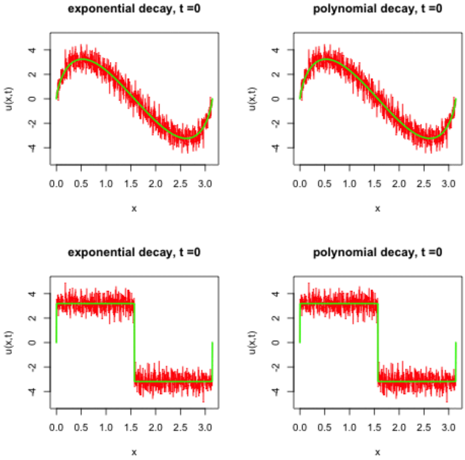


Ill posed-inverse problems, mainstay



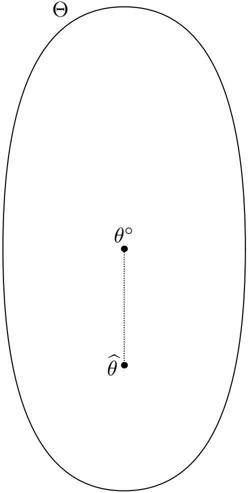

Θ
$f^\star(\Theta)$
X
f
Randomness
θ°
$f(\theta^\circ)$
Y^n
$\widehat{f}$
$P_{Y^n|f(\theta^\circ)}$
$\widetilde{\theta} = f^\dagger(\widehat{f})$
$f^\dagger$

Ill posed-inverse problems, data example



$$
\begin{aligned}
&\text{Process:} && \frac{\partial Y_{x,t}}{\partial t} &&= \frac{\partial^2 Y_{x,t}}{\partial x^2} + \frac{1}{n}\frac{dW(x,t)}{dt};\\
&\text{Solution:} && Y_j(t) &&= \theta_j^\circ(t) + \frac{1}{n}\xi_j(t)\\
& && \theta_j^\circ(t) &&= \mathscr{F}(\theta^\circ(\cdot,0))(j)\cdot\exp[-j^2 t]\\
&\text{Observation:} && Y_j &&= \theta_j^\circ(T) + \frac{1}{n}\xi_j;\\
&\text{Interest:} && \theta_j^\circ(0) &&= \theta_j^\circ(T)\exp[j^2 T].
\end{aligned}
$$

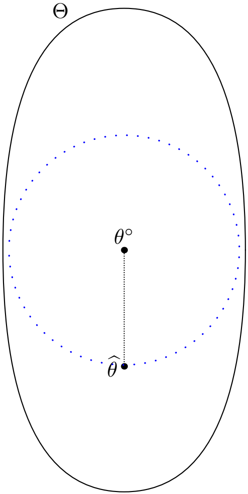
Contents



- Distance to the truth:

$$l(\widehat{\theta}, \theta^\circ)$$

Frequentist convergence rate, mainstay

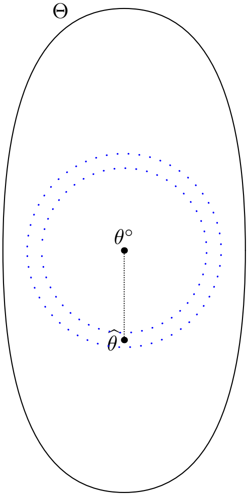


- Distance to the truth:

$$l(\widehat{\theta}, \theta^\circ)$$

- Average distance to the truth:

$$\mathbb{E}_{\theta^\circ}\left[l(\widehat{\theta}, \theta^\circ)\right]$$



- Distance to the truth:

$$l(\widehat{\theta}, \theta^\circ)$$

- Average distance to the truth:

$$\mathbb{E}_{\theta^\circ}\left[l(\widehat{\theta}, \theta^\circ)\right]$$

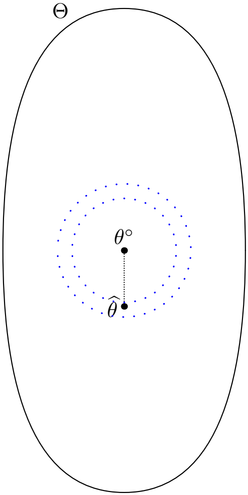
- Lower bound: $\Phi_n(\theta^\circ) \to 0$, $K > 1$

$$\inf_{\widetilde{\theta} \in \mathscr{E}} \mathbb{E}_{\theta^\circ}\left[l(\widetilde{\theta}, \theta^\circ)\right] \geq K^{-1}\Phi_n(\theta^\circ).$$

  Upper bound:

$$\mathbb{E}_{\theta^\circ}\left[l(\widehat{\theta}, \theta^\circ)\right] \leq K\Phi_n(\theta^\circ).$$

Frequentist convergence rate, mainstay



- Distance to the truth:

$$l(\widehat{\theta}, \theta^\circ)$$

- Average distance to the truth:

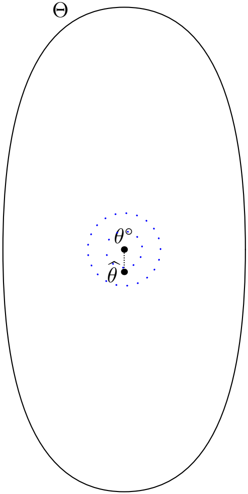
$$\mathbb{E}_{\theta^\circ}\left[l(\widehat{\theta}, \theta^\circ)\right]$$

- Lower bound: $\Phi_n(\theta^\circ) \to 0$, $K > 1$

$$\inf_{\widetilde{\theta} \in \mathscr{E}} \mathbb{E}_{\theta^\circ}\left[l(\widetilde{\theta}, \theta^\circ)\right] \geq K^{-1}\Phi_n(\theta^\circ).$$

Upper bound:

$$\mathbb{E}_{\theta^\circ}\left[l(\widehat{\theta}, \theta^\circ)\right] \leq K\Phi_n(\theta^\circ).$$



- Distance to the truth:

$$l(\widehat{\theta}, \theta^\circ)$$

- Average distance to the truth:

$$\mathbb{E}_{\theta^\circ}\left[l(\widehat{\theta}, \theta^\circ)\right]$$

- Lower bound: $\Phi_n(\theta^\circ) \to 0$, $K > 1$

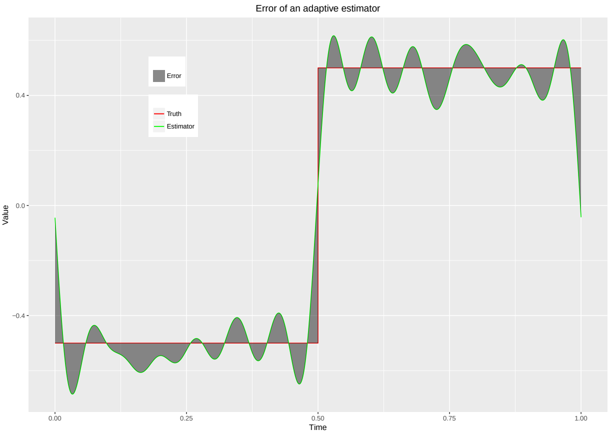
$$\inf_{\widetilde{\theta} \in \mathscr{E}} \mathbb{E}_{\theta^\circ}\left[l(\widetilde{\theta}, \theta^\circ)\right] \geq K^{-1} \Phi_n(\theta^\circ).$$

Upper bound:

$$\mathbb{E}_{\theta^\circ}\left[l(\widehat{\theta}, \theta^\circ)\right] \leq K \Phi_n(\theta^\circ).$$

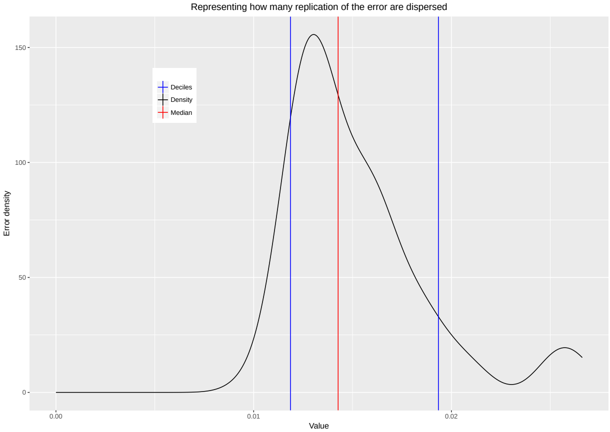
Frequentist convergence rate, example

- $\|\widehat{\theta}-\theta^\circ\|^2 = \int_{[0,1]} (\widehat{\theta}(x)-\theta^\circ(x))^2 dx;$



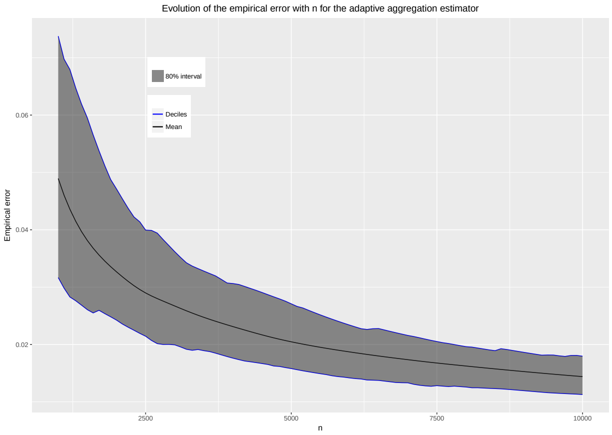
Frequentist convergence rate, example

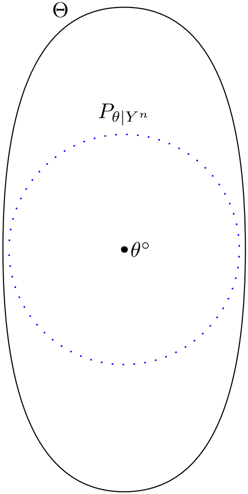
- $\mathbb{E}_{\theta^\circ}[\|\widehat{\theta}-\theta^\circ\|^2] = \int\limits_{X} \|\widehat{\theta}(Y^n)-\theta^\circ\| d\mathbb{P}_{Y^n};$



Frequentist convergence rate, example

- $\mathbb{E}_{\theta^\circ}[\|\widehat{\theta}-\theta^\circ\|^2] = \mathcal{O}_n(\Phi_n^\circ)$

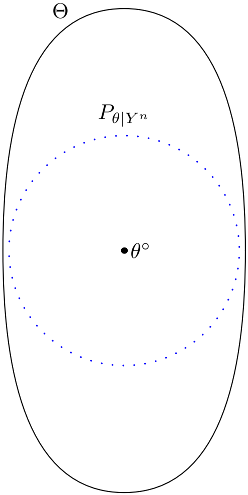




- Probability to give a guess within a distance of the truth:

$$\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta}, \theta^\circ) \leq \Phi_n\right);$$

Bayesian contraction rate, principle



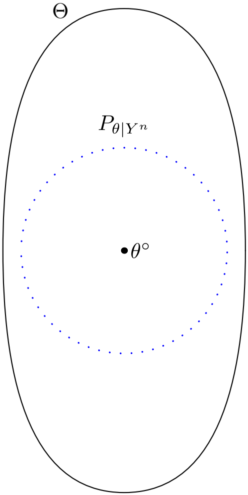
- Probability to give a guess within a distance of the truth:

$$\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta}, \theta^\circ) \leq \Phi_n\right);$$

- Average probability to give a guess within a distance of the truth:

$$\mathbb{E}_{\theta^\circ}\left[\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta}, \theta^\circ) \leq \Phi_n\right)\right];$$

Bayesian contraction rate, principle



- Probability to give a guess within a distance of the truth:

$$\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta}, \theta^\circ) \leq \Phi_n\right);$$

- Average probability to give a guess within a distance of the truth:

$$\mathbb{E}_{\theta^\circ}[\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta}, \theta^\circ) \leq \Phi_n\right)];$$

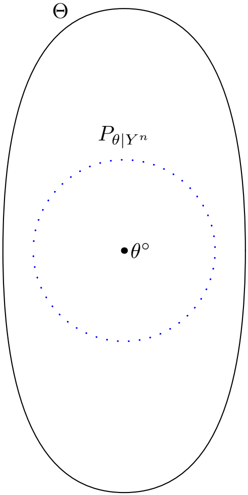
- Lower bound:

$$\lim_{n\to\infty} \sup_{\mathbb{Q}_{\boldsymbol{\theta}} \in \mathscr{G}} \mathbb{E}_{\theta^\circ}[\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta}, \theta^\circ) \leq K^{-1}\Phi_n\right)] < 1;$$

upper bound:

$$\lim_{n\to\infty} \mathbb{E}_{\theta^\circ}[\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta}, \theta^\circ) \leq K\Phi_n\right)] = 1.$$

Bayesian contraction rate, principle



- Probability to give a guess within a distance of the truth:

$$\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta},\theta^\circ) \leq \Phi_n\right);$$

- Average probability to give a guess within a distance of the truth:

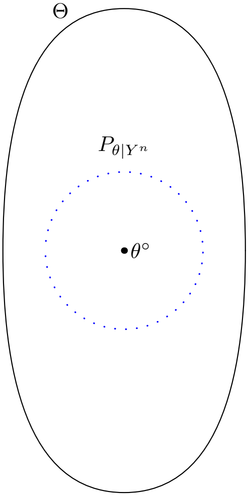
$$\mathbb{E}_{\theta^\circ}\left[\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta},\theta^\circ) \leq \Phi_n\right)\right];$$

- Lower bound:

$$\lim_{n\to\infty} \sup_{\mathbb{Q}_{\boldsymbol{\theta}}\in\mathcal{G}} \mathbb{E}_{\theta^\circ}\left[\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta},\theta^\circ) \leq K^{-1}\Phi_n\right)\right] < 1;$$

upper bound:

$$\lim_{n\to\infty} \mathbb{E}_{\theta^\circ}\left[\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta},\theta^\circ) \leq K\Phi_n\right)\right] = 1.$$



- Probability to give a guess within a distance of the truth:

$$\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta},\theta^\circ) \leq \Phi_n\right);$$

- Average probability to give a guess within a distance of the truth:

$$\mathbb{E}_{\theta^\circ}[\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta},\theta^\circ) \leq \Phi_n\right)];$$

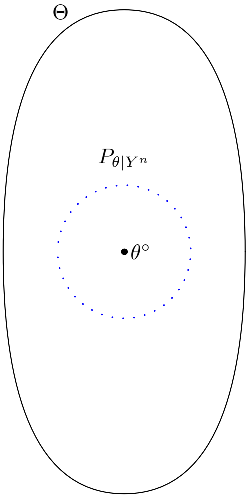
- Lower bound:

$$\lim_{n\to\infty} \sup_{\mathbb{Q}_{\boldsymbol{\theta}}\in\mathscr{G}} \mathbb{E}_{\theta^\circ}[\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta},\theta^\circ) \leq K^{-1}\Phi_n\right)] < 1;$$

upper bound:

$$\lim_{n\to\infty} \mathbb{E}_{\theta^\circ}[\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta},\theta^\circ) \leq K\Phi_n\right)] = 1.$$

Bayesian contraction rate, principle



- Probability to give a guess within a distance of the truth:

$$\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta},\theta^\circ) \leq \Phi_n\right);$$

- Average probability to give a guess within a distance of the truth:

$$\mathbb{E}_{\theta^\circ}[\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta},\theta^\circ) \leq \Phi_n\right)];$$

- Lower bound:

$$\lim_{n\to\infty} \sup_{\mathbb{Q}_{\boldsymbol{\theta}}\in\mathcal{G}} \mathbb{E}_{\theta^\circ}[\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta},\theta^\circ) \leq K^{-1}\Phi_n\right)] < 1;$$

upper bound:

$$\lim_{n\to\infty} \mathbb{E}_{\theta^\circ}[\mathbb{P}_{\boldsymbol{\theta}|Y^n}\left(\boldsymbol{\theta} : l(\boldsymbol{\theta},\theta^\circ) \leq K\Phi_n\right)] = 1.$$

Bibliography

Comte, F., Cuenod, C.-A., Pensky, M., and Rozenholc, Y. (2017). Laplace deconvolution on the basis of time domain data and its application to dynamic contrast-enhanced imaging. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(1):69–94.

Engl, H. W. and Kügler, P. (2005). Nonlinear inverse problems: theoretical aspects and some industrial applications. In *Multidisciplinary methods for analysis optimization and control of complex systems*, pages 3–47. Springer.

Mohammad-Djafari, A. (2002). Bayesian inference for inverse problems. In *AIP Conference Proceedings*, volume 617, pages 477–496. AIP.

Mohammad-Djafari, A., Giovannelli, J.-F., Demoment, G., and Idier, J. (2002). Regularization, maximum entropy and probabilistic methods in mass spectrometry data processing problems. *International Journal of Mass Spectrometry*, 215(1):175 – 193. Detectors and the Measurement of Mass Spectra.

Naidu, P. S. and Westphal, K. O. (1965). The inverse problem of particle motion and its application. *Applied Scientific Research, Section B*, 12(5):435–450.