STAT 403

May 12, 2025

l15 – TBPosted
project - TBPosted
Lecture Notes – on MCMC
. YC Chen
. MMP – t.b. posted

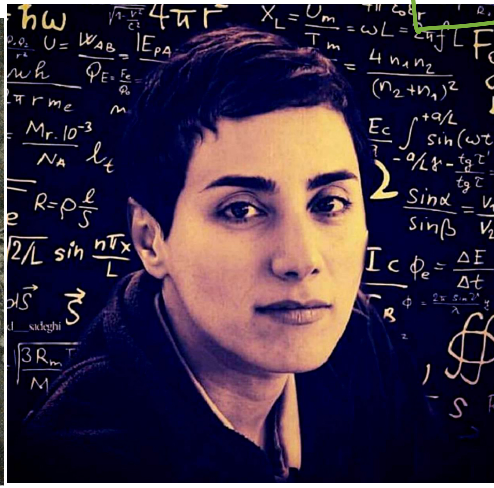
- MSE - $\text{Bias}^2 + \text{Var}$ decomposition
- MCMC

Lecture 16

The Secrets of the Surface

3pm
PDL C-301

International Nurses' Day



unofficial

International Day of Women in Statistics!

International Day of Women in Mathematics

Mean Square Error (MSE)

θ = true parameter

$\hat{\theta}_n$ = estimated θ from $\mathcal{D}$, $|\mathcal{D}| = n$

error $= \hat{\theta}_n - \theta$

random

$$\text{MSE}(\hat{\theta}_n) = \text{MSE}(\hat{\theta}_n, \theta) = \boxed{\mathbb{E}\left((\hat{\theta}_n - \theta)^2\right)}.$$

$E[\hat{\theta}_n]$

$Var[\hat{\theta}_n]$

By simple algebra, the MSE of $\hat{\theta}_n$ equals

$$\begin{aligned}
\text{MSE}(\hat{\theta}_n, \theta) &= \mathbb{E}\left(\overbrace{(\hat{\theta}_n - \theta)}^{\text{err}}{}^2\right) \\
&= \mathbb{E}\left((\underbrace{\hat{\theta}_n - \mathbb{E}(\hat{\theta}_n)}_{\text{var}} + \underbrace{\mathbb{E}(\hat{\theta}_n) - \theta}_{\text{bias}})^2\right) \\
&= \underbrace{\mathbb{E}\left((\hat{\theta}_n - \mathbb{E}(\hat{\theta}_n))^2\right)}_{=\text{Var}(\hat{\theta}_n)} + 2\underbrace{\mathbb{E}\left(\hat{\theta}_n - \mathbb{E}(\hat{\theta}_n)\right)}_{=0} \cdot (\mathbb{E}(\hat{\theta}_n) - \theta) + \left(\underbrace{\mathbb{E}(\hat{\theta}_n) - \theta}_{=\text{bias}(\hat{\theta}_n)}\right)^2 \\
&= \text{Var}(\hat{\theta}_n) + \text{bias}^2(\hat{\theta}_n).
\end{aligned}$$

$E[\hat{\theta}_n] - E[\hat{\theta}_n]$

Namely, the MSE of an estimator is the variance plus the square of bias. This decomposition is also known as the *bias-variance tradeoff* (or bias-variance decomposition).

Lemma

If an estimator has MSE converging to 0*, then it is a consistent estimator.*

$$\text{MSE}(\hat{\theta}_n, \theta) \to 0 \implies \hat{\theta}_n \xrightarrow{P} \theta.$$

Markov Chain Monte Carlo

goal: answer statistical question by MC
(= estimate θ)

when: large systems
multivariate p, or P

Markov Chain

$X_1 - X_2 - X_3 - X_4$

no edge
$\Rightarrow$ no "direct" dependence
$\equiv$ conditional independence

$X_1 \perp X_3 \mid X_2$

$X_2 \perp X_4 \mid X_3$

$X_1 \perp X_4 \mid X_2 X_3$ or
X_2 or X_3

M. Blanket

past

$X_1 - X_2 - \cdots - X_{t-1} - X_t - X_{t+1} - \cdots X_T$

future

Independence properties

Markov properties

$X_t \perp X_{1:t-2} \mid X_{t-1}$

$X_t \perp X_{t+2:T} \mid X_{t+1}$

$X_t \perp$ all other $X_{t'}$ $\mid \underbrace{X_{t-1}, X_{t+1}}$

$t' \neq t, t\pm 1$

Markov blanket of X_t

Transition Matrix $\underline{\mathbf{P}}$

$X_t \in S$ discrete finite for $t = 1:T$

- $P[X_1]$ initial state probability

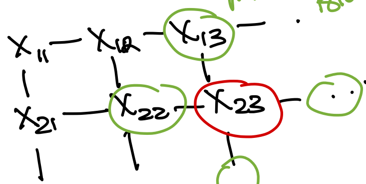
$P = [P_{ij}]_{i,j=1:m}$

$|S| = m$

$P_{ij} = [X_{t+1} = j \mid X_t = i]$

$\sum_{j=1}^{m} P_{ij} = 1 \Leftrightarrow P$ is stochastic matrix

Markov Blanket



images

D

Distribution on S^T

$\tilde{x} = (x_1, \dots x_T) \in S^T$

$\tilde{p} : S^T \to \mathbb{R}$ distribution of $\tilde{x}$

Ex: $S = \{0,1\}$

$\tilde{x} = (0,1,1,0) \in S^4$

$T=4$

Ex: $S = \{\text{aminoacids}\}$

$|S| = 20$

$T = 10$

$\tilde{x}$ = peptides of length T