SPR Day 23

Probabilistic Programming

Reminders:

- Quiz 4 due Fri
- HW5 due Sun 4/26
- Quiz 5 due Wed 4/29
- CPS due Mon 5/4
- Final due Fri 5/8 noon

Today: Final Exam policy

ProbProg! High-level intro

Walk thru of 2 notebooks — PyStan, PyMC3

Not planning to drop hw, qz, quiz

Final Exam OPTIONAL

don't take:

final score = max(midterm score, lowest quiz)

take:

final score = max(real exam score, don't take score)

Pass/fail deadline last Day Mon 4/27

grade reports by Fri:

don't take score

Office Hours

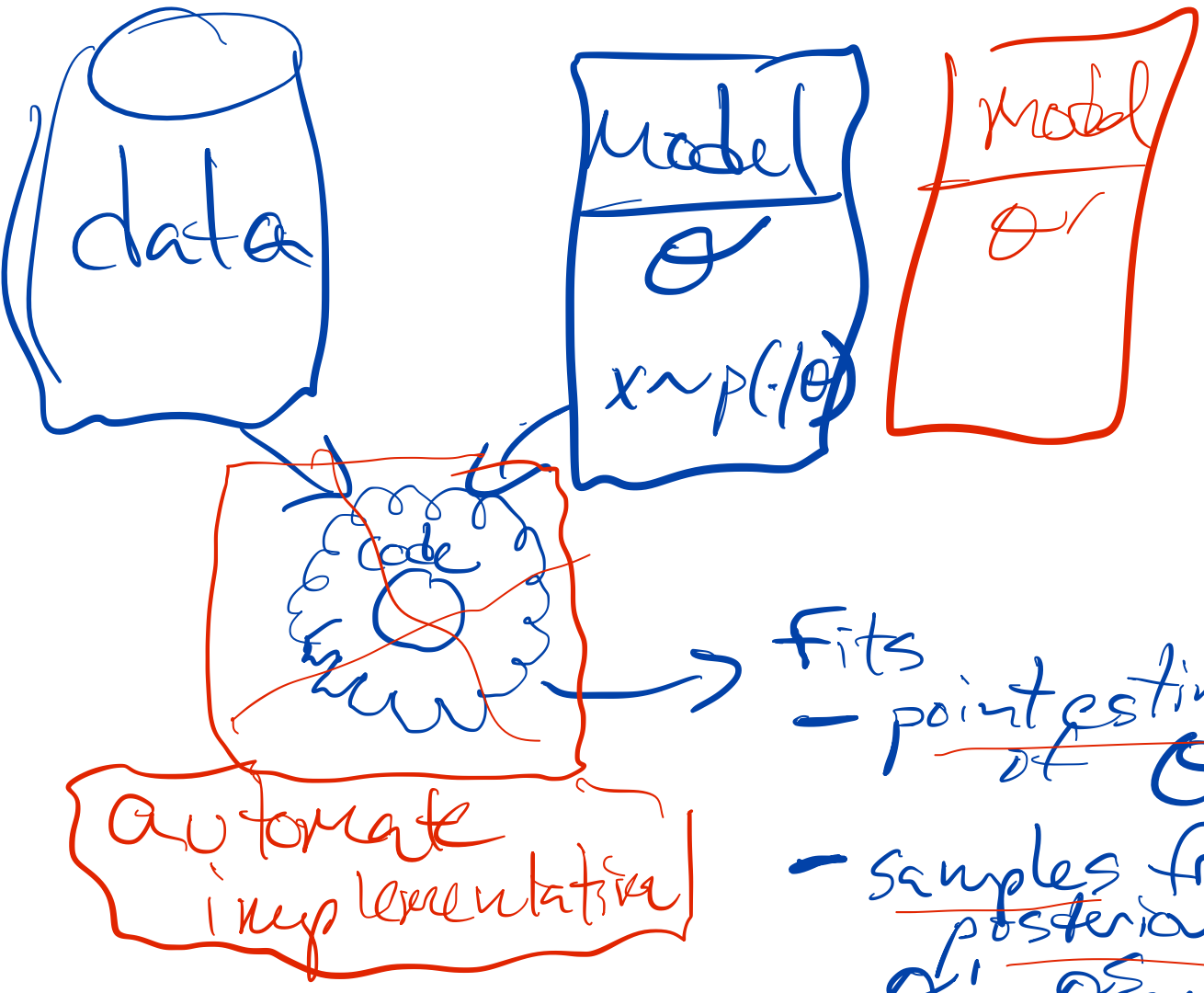
yes as scheduled

every M/W until W 5/6

not today!

vigilant on Piazza

Probabilistic Programming



→ Fits
- point estimates of θ ✓
- samples from posterior $\theta_1 \ldots \theta_S \sim p(\theta|x)$ ✓

CP3: heroic effort to impl EM

even more effort

reduce M step

CP4: effort for HMM

Goal: Reduce implementation manual effort

Challenges:

- How to do sampling for general family of models?

- How to do points? (MAP) for general family.

$$\max_{\theta} \log p(x|\theta) + \log p(\theta)$$

θ is $\mathbb{R}$

$\log p(x|\theta)$: can take gradients

$\log p(\theta)$: can take gradients

apply L-BFGS to optimize!

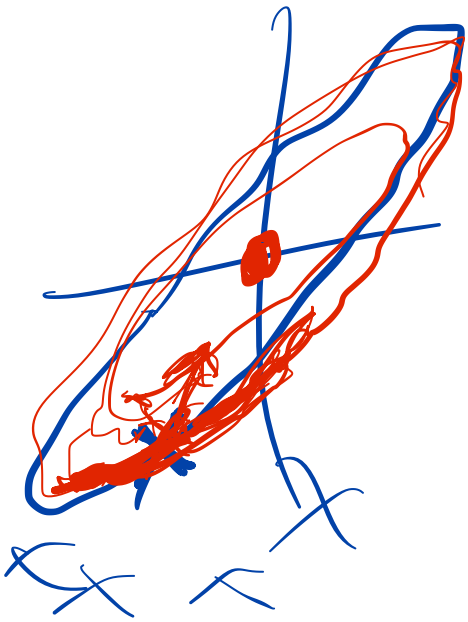
MCMC for generic models?

$p^*(\theta_1 \dots \theta_M)$

- Gibbs — $p^*(\theta_m \mid \theta_{\setminus m})$ hard in general
- Random Walk — possible if $\theta \in \mathbb{R}^M$ but SLOW to converge
- **Hamiltonian MCMC**

take BDL COMP 150

HMC is M-H algorithm where proposals informed by gradients of log pdf



θ all continuous gradients possible

Linear Regr

- What would it take to do Gibbs sampling?

$w \sim N(0, 10)$

weight $w \in \mathbb{R}$

$\tau \sim$ 

$\tau > 0$

$$y_n | x_n \sim N(w \cdot x_n, \tau^2)$$

$$p(\tau) = \text{HalfNormPDF}(\tau) = \begin{cases} 2\,\text{NormPDF}(\tau | 0, 1) & \tau > 0 \\ 0 & \text{o.w.} \end{cases}$$

$p^*(\tau | w)$ ← x, y HARD

b/c HalfNormal + Normal is not standard posterior

Need:
- distrib
- how to sample

$p^*(w | \tau)$ POSSIBLE

Rest of time:

Review same LinReg model fit using PyStan & PyMC3

Key Idea: Can get samples from posterior, not just point estimate

---

Limitations

to use LBFGS for point estimates or HamiltonianMCMC for samples,

we need model to:

1) have continuous real θ params
2) have "unimodal" posterior (NOT well-separated multiple peaks)