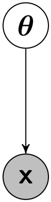
Neural density estimation for likelihood-free inference

George Papamakarios
g.papamakarios@ed.ac.uk

University of Edinburgh

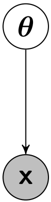
Bayesian inference

$$\underset{\text{posterior}}{p(\boldsymbol{\theta} \mid \mathbf{x})} \propto \overset{\text{likelihood}}{p(\mathbf{x} \mid \boldsymbol{\theta})}\ \underset{\text{prior}}{p(\boldsymbol{\theta})}$$



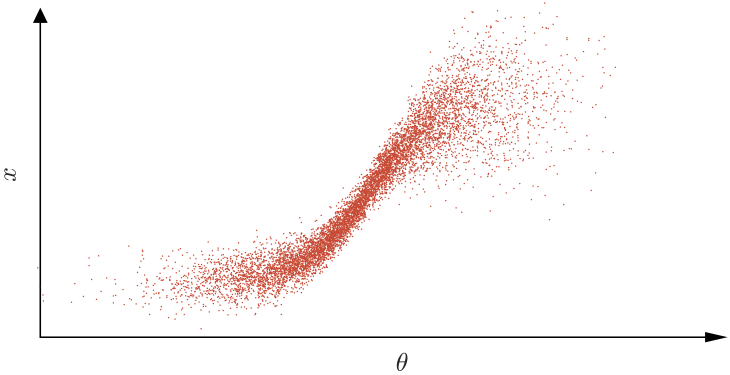
Likelihood-free inference

$$\underbrace{p(\boldsymbol{\theta} \mid \mathbf{x})}_{\text{posterior}} \propto \overset{\cancel{\text{likelihood}}}{\cancel{p(\mathbf{x} \mid \boldsymbol{\theta})}} \underbrace{p(\boldsymbol{\theta})}_{\text{prior}}$$

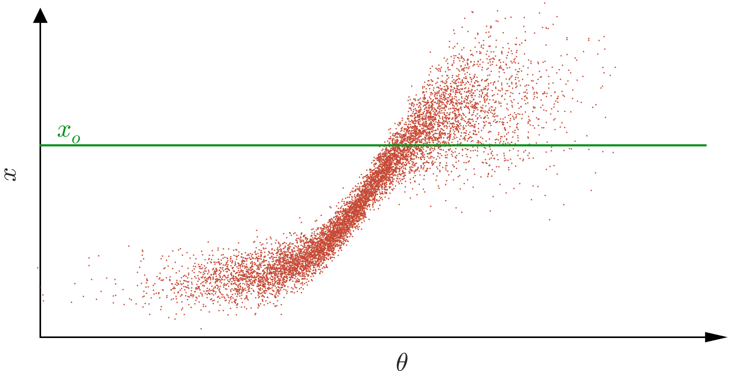


Can simulate $\boldsymbol{\theta} \to \mathbf{x}$ for any $\boldsymbol{\theta}$

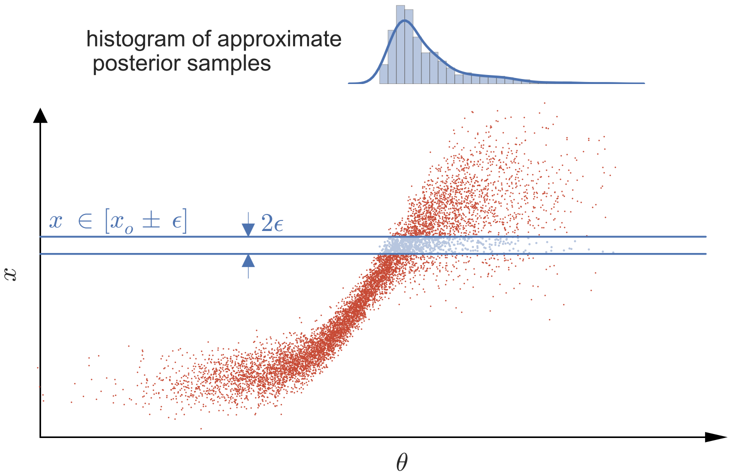
Simulations



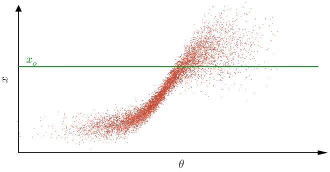
Simulations & observation

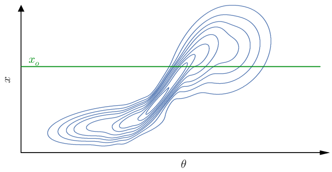


Approximate Bayesian Computation



Modelling conditional densities

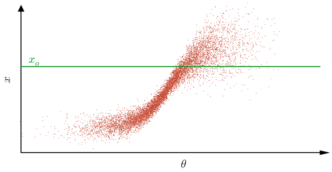




From simulated data $\{(\boldsymbol{\theta}_1, \mathbf{x}_1), \ldots, (\boldsymbol{\theta}_N, \mathbf{x}_N)\} \Rightarrow$ learn a **neural network model** of either

- ~~the joint $p(\mathbf{x}, \boldsymbol{\theta})$~~ meh
- the likelihood $p(\mathbf{x} \mid \boldsymbol{\theta})$
- the posterior $p(\boldsymbol{\theta} \mid \mathbf{x})$

Guiding simulations



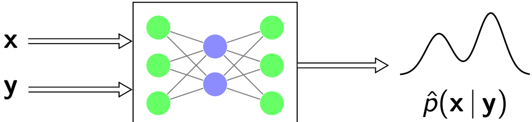
Early models of the posterior $p(\boldsymbol{\theta} \mid \mathbf{x})$ can guide future simulations

- Learn a model $\hat{p}(\boldsymbol{\theta} \mid \mathbf{x})$ using only data simulated so far
- Choose $\boldsymbol{\theta}$ to simulate next by proposing from $\hat{p}(\boldsymbol{\theta} \mid \mathbf{x})$

Papamakarios & Murray, *Fast ϵ-free Inference of Simulation Models with Bayesian Conditional Density Estimation*, NIPS 2016

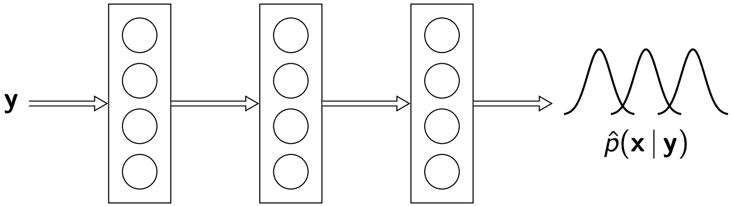
Conditional neural density estimation

Given data $\{(\mathbf{x}_1, \mathbf{y}_1), \ldots, (\mathbf{x}_N, \mathbf{y}_N)\} \Rightarrow$ learn $\hat{p}(\mathbf{x} \mid \mathbf{y})$



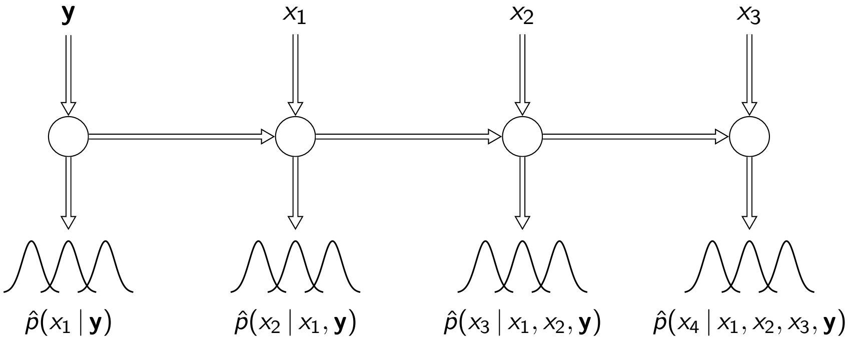
Mixture Density Network

(Bishop, 1994)



A feedforward neural net that outputs a Gaussian mixture

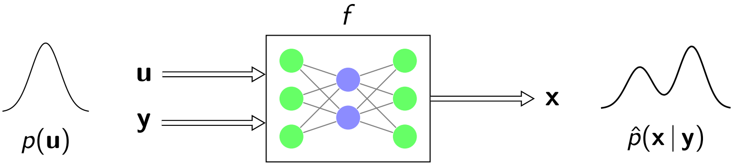
Autoregressive models



Examples

Seq to Seq (Sutskever et al., 2014)
PixelRNN / PixelCNN (van den Oord et al., 2016)
WaveNet (van den Oord et al., 2016)

Normalizing flows

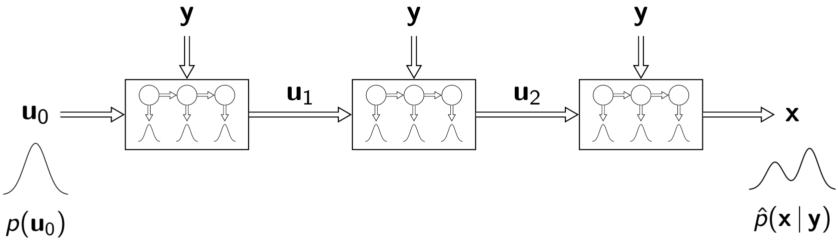


$\mathbf{x} = f(\mathbf{u}, \mathbf{y})$ where $f(\cdot, \mathbf{y})$ is easily invertible with tractable Jacobian

Examples

Planar / radial flows (Rezende & Mohamed, 2014)
Inverse Autoregressive Flow (Kingma et al., 2016)
RealNVP (Dinh et al., 2017)

Masked Autoregressive Flow



A sequence of autoregressive models, each modelling the random numbers $\mathbf{u}_i$ driving the next model in the sequence

Papamakarios et al., *Masked Autoregressive Flow for Density Estimation*, arXiv 2017

Summary

Likelihood-free inference

- Can be done by conditional density estimation
- Learning proposals can lead to large savings in simulations

Neural density estimation

- Training neural networks to learn densities from data
- Fast-growing field: autoregressive models, normalizing flows, autoregressive flows

Summary

Likelihood-free inference

- Can be done by conditional density estimation
- Learning proposals can lead to large savings in simulations

Neural density estimation

- Training neural networks to learn densities from data
- Fast-growing field: autoregressive models, normalizing flows, autoregressive flows

Thank you!