University of Basel

Characteristics of the Metropolis-Hastings algorithm

Marcel Lüthi, Departement of Mathematics and Computer Science, University of Basel

The Metropolis-Hastings Algorithm

Generalization of Metropolis algorithm to asymmetric proposal distribution

$$Q(\boldsymbol{x}'|\boldsymbol{x}) \neq Q(\boldsymbol{x}|\boldsymbol{x}')$$

$$Q(\boldsymbol{x}'|\boldsymbol{x}) > 0 \Leftrightarrow Q(\boldsymbol{x}|\boldsymbol{x}') > 0$$

Initialize with sample $\boldsymbol{x}$

Generate next sample, with current sample $\boldsymbol{x}$

1. Draw a sample $\boldsymbol{x}'$ from $Q(\boldsymbol{x}'|\boldsymbol{x})$ ("proposal")
2. With *probability* $\alpha = \min\left\{\frac{P(x')}{P(x)}\frac{Q(x|x')}{Q(x'|x)}, 1\right\}$ accept $\boldsymbol{x}'$ as new state $\boldsymbol{x}$
3. Emit current state $\boldsymbol{x}$ as sample

Example: 2D Gaussian

Target: $P(\boldsymbol{x}) = \frac{1}{2\pi\sqrt{|\Sigma|}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x}-\boldsymbol{\mu})}$

Proposal: $Q(\boldsymbol{x}'|\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}'|\boldsymbol{x}, \sigma^2 I_2)$ — Random walk

Target

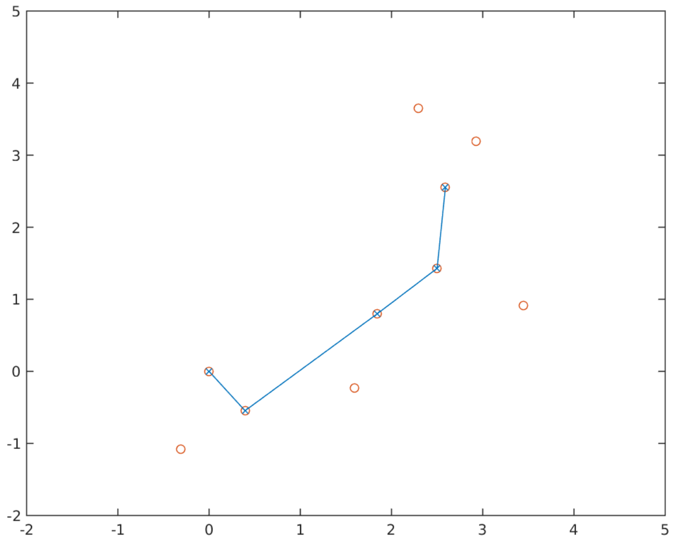
$$\mu = \begin{bmatrix} 1.5 \\ 1.5 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 1.25 & 0.75 \\ 0.75 & 1.25 \end{bmatrix}$$

Sampled Estimate

$$\hat{\mu} = \begin{bmatrix} 1.56 \\ 1.68 \end{bmatrix}$$

$$\hat{\Sigma} = \begin{bmatrix} 1.09 & 0.63 \\ 0.63 & 1.07 \end{bmatrix}$$

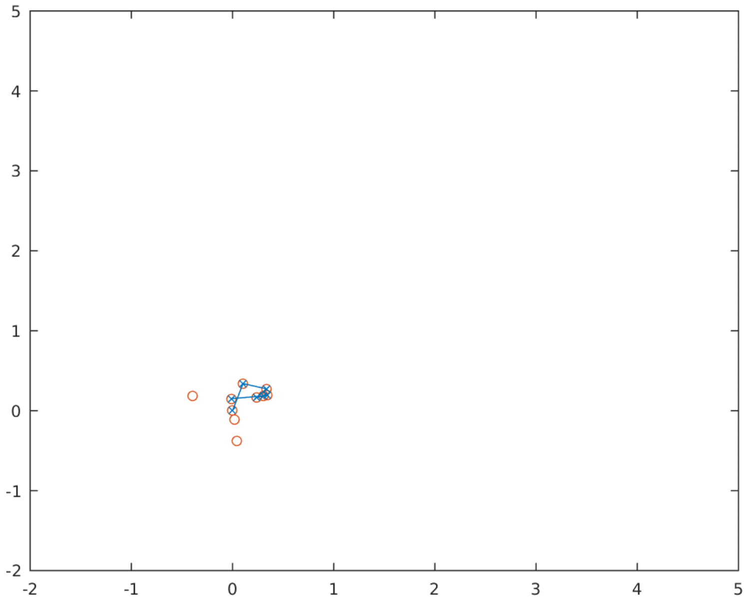


2D Gaussian: Different Proposals

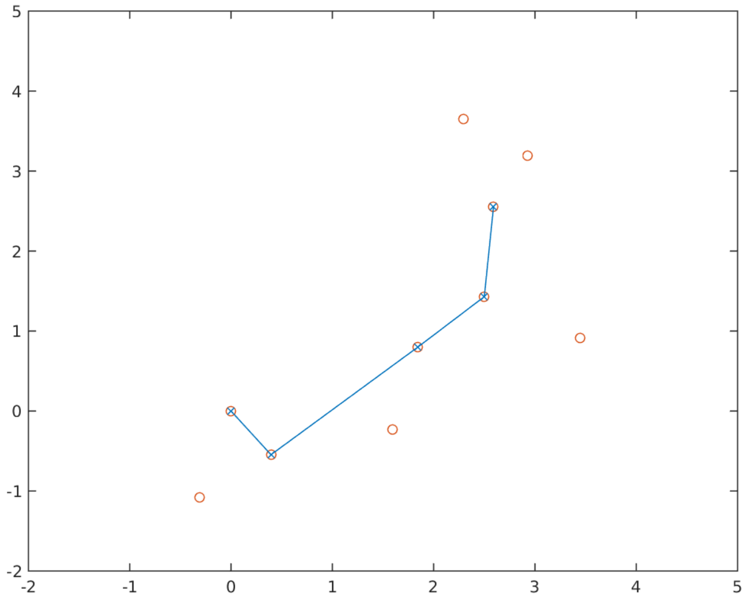
Target: $P(\boldsymbol{x}) = \frac{1}{2\pi\sqrt{|\Sigma|}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}$

Proposal: $Q(\boldsymbol{x}'|\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}'|\boldsymbol{x}, \sigma^2 I_2)$

$\sigma = 0.2$



$\sigma = 1.0$



Samples are unbiased, but not uncorrelated

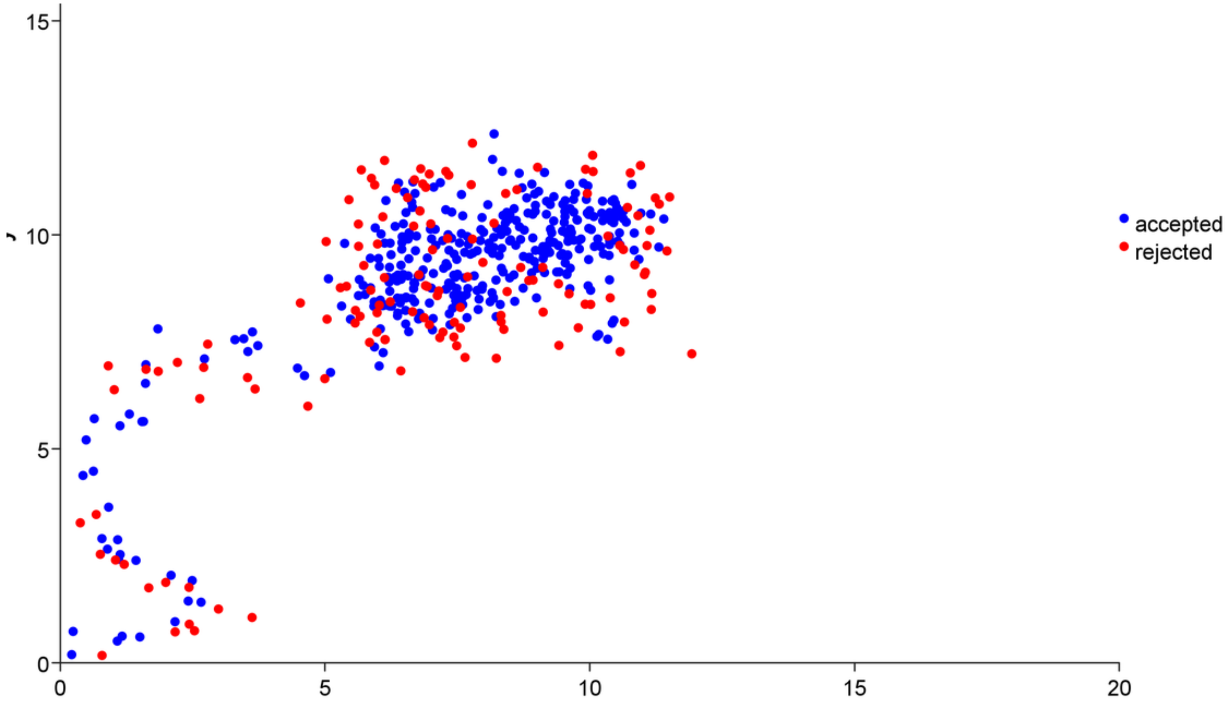
Burn in Phase

Might start far away from high-probability area

- Needs time until chain converges

Samples from *burn in phase* needs to be discarded

- *Length of burn in phase unclear*



Metropolis-Hastings as a Propose-and-Verify Architecture

Propose
Draw a sample x' from $Q(x'|x)$

Verify
With *probability* $\alpha = \min\left\{\frac{P(x')}{P(x)}\frac{Q(x|x')}{Q(x'|x)}, 1\right\}$ accept $\boldsymbol{x}'$ as new sample

Decouples the steps of finding the solution from validating a solution

- Natural to integrate uncertain proposals Q (e.g. automatically detected landmarks, ...)
- Possibility to include “local optimization” (e.g. a ICP or ASM updates, gradient step, …) as proposal

A way to structure complex probabilistic fitting applications