## Previously

- Simulation, sampling
- Monte Carlo Simulations
- Inverse cdf method
- Rejection sampling

- Today: sampling cont.,
- Bayesian inference via sampling
- Eigenvalues and Eigenvectors
- Markov processes, PageRank alg.

1

## Previously Monte Carlo Integration

- Compute $\int_0^1 f(x)dx$

1. Draw $u_1, \cdots, u_n$ from U(0,1)
2. Approximate integral as

$$\hat{I} = \frac{1}{n}(f(u_1) + \cdots + f(u_n))$$

Now to arbitrary integral from on interval (a,b)

$$\int_a^b f(x)dx \qquad \hat{I} = \frac{(b-a)}{n}(f(u_1) + \cdots + f(u_n))$$

This is interpreted as expected value of RV

$$\hat{I} = E[(b-a)f(x)])$$

2

## How to sample from $f(x)$

- Inversion Method, where $F$ is cdf of $f(x)$

$$X = F^{-1}(U)$$

3

## Rejection sampling

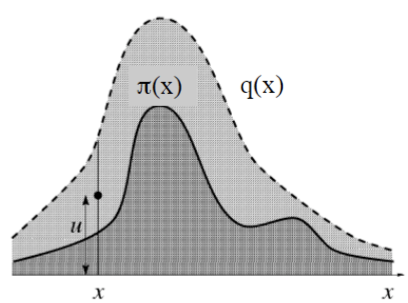



Image by MacKay

1. Generate $x$ from $g(x)$
2. Draw $u$ from $unif(0,1)$
3. Accept if $u < f(x)/Mg(x)$
4. The accepted $x$ follows $f(x)$

Problems many samples can get rejected if $g(x)$ is too different from $f(x)$

## Importance sampling

- Suppose you have $g(x)$ and $f(x)$
- But do not know the scale M
- Sample from $g(x)$ $x_1, \ldots, x_n$
- Calculate weight for each sample

$$w_i = f(x_i)/g(x_i)$$

- Mass for each sample is

$$q_i = w_i / \sum_{i=1}^{n} w_i$$

5

## Importance Sampling

- The weights are biggest when the distributions agree

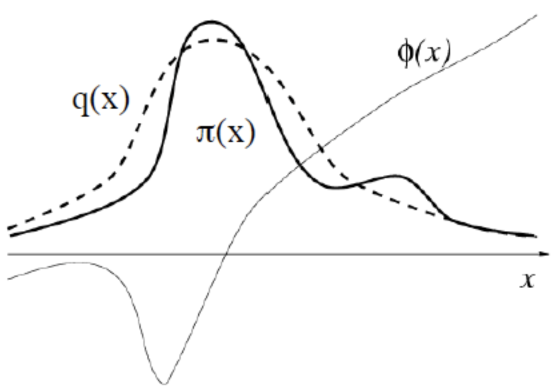



Image by MacKay

6

## Importance Sampling in 1D

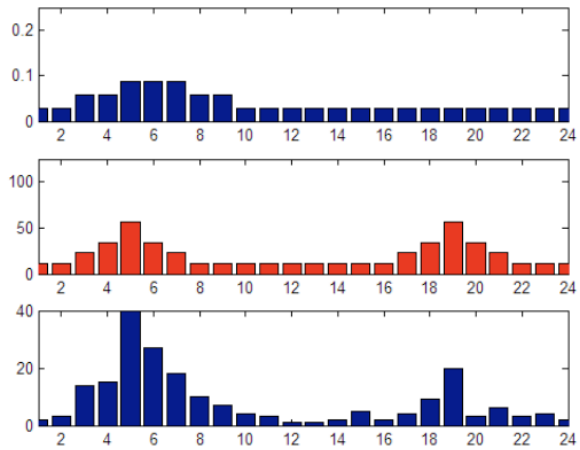



Image Frank Deallert

7

## Bayes calculations

- Recall $p(x \mid z) \propto l(x;z)\, p(x)$
- We can do rejection sampling from posterior, but no guarantee that

$$p(x) >= p(x \mid z)$$

- Idea, sample from $p(x)$ , and give each sample importance weight $w_i = p(z \mid x_i)$

$$q_i = w_i / \sum_{i=1}^{n} w_i$$

8

## Bayesian inference via sampling

- Robot example
- Probability of robot’s position in 2D
- Different representations of probability
- Histograms
- Mixture of Gaussians
- Does not scale up to higher dimensions

9

## Monte Carlo Expected Value

- Prior – sample from mixture of Gaussians
- Compute expected value of the bearing (angle)
- to origin

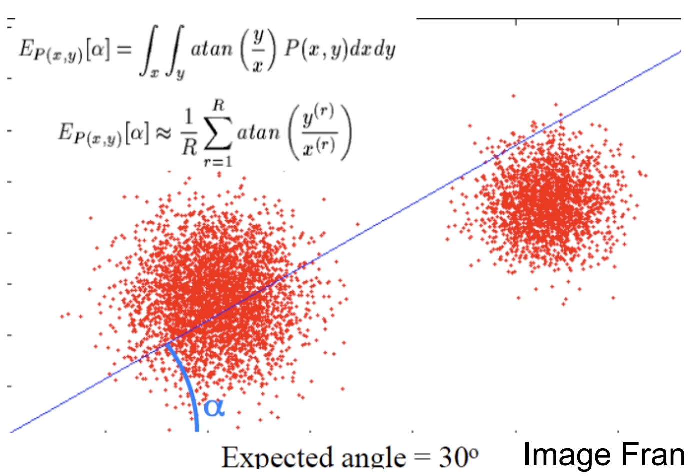
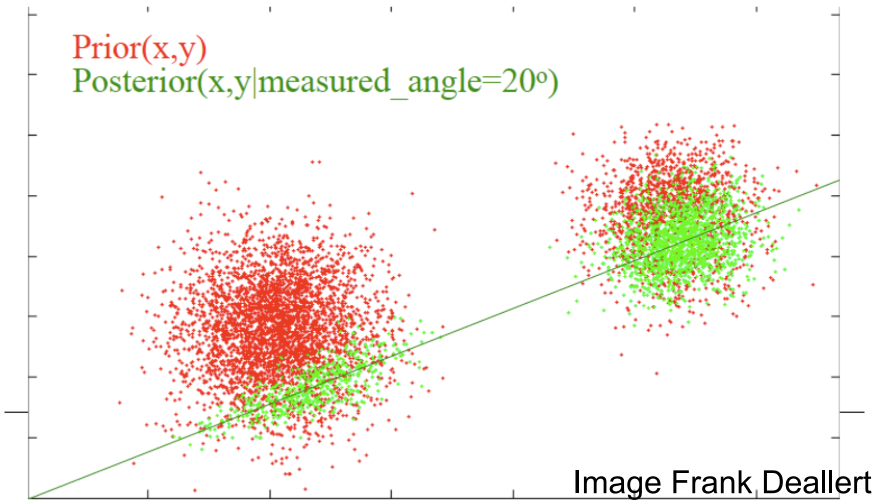



Image Frank Deallert

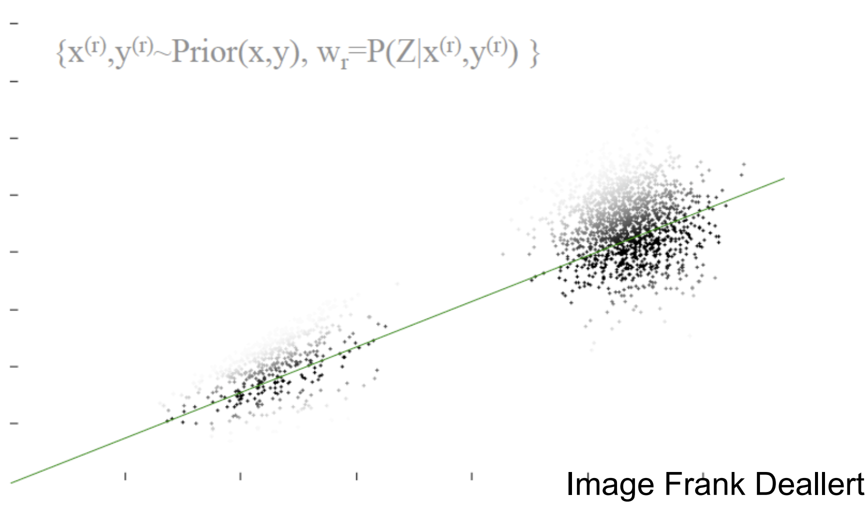
10

## Rejection sampling example

- P(measured_angle|x,y) = N(predicted_angle,3 degrees)




## Importance Sampling Example



## Linear algebra - digression

- Previously
- Rank of a matrix
- Determinant of a matrix
- Matrix inverse
- Matrix pseudo-inverse
- Solving system of equations
- Solving linear least squares problems

13

## Determinants

- If $\det(A) \neq 0$ then matrix is invertible
- If $\det(A) = 0$ then matrix is not invertible
- A is rank deficient
- Focus was on solving matrix inversion problems
- Now we look at other properties of matrices
- Useful when A represents a transformations

$$y = Ax$$

- Or A simply represents data

14

## Eigenvalues and Eigenvectors

- Motivated by solution to differential equations
- For square matrices $A \in \Re^{n \times n}$ $\quad \dot{\mathbf{u}} = A\mathbf{u}$ $\qquad A = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix}$

For scalar ODE's

$$\dot{u} = au$$

$$u(t) = e^{at}u(0)$$

Behavior varies
If a > 0 unstable
a = 0 neutrally stable
a < 0 stable

- We look for the solutions
- of the following type exponentials

$$v(t) = e^{\lambda t} y \qquad u(t) = [v(t)\ w(t)]^T$$

$$w(t) = e^{\lambda t} z$$

Substitute back to the equation

$$\lambda \cancel{e^{\lambda t}} y = 4 \cancel{e^{\lambda t}} y - 5 \cancel{e^{\lambda t}} z$$

$$\lambda \cancel{e^{\lambda t}} z = 2 \cancel{e^{\lambda t}} y - 3 \cancel{e^{\lambda t}} z$$

$$\mathbf{x} = \begin{bmatrix} y \\ z \end{bmatrix} \qquad \lambda \mathbf{x} = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix} \mathbf{x}$$

## Eigenvalues and Eigenvectors

$$\lambda \mathbf{x} = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix} \mathbf{x} \qquad A\mathbf{x} = \lambda \mathbf{x}$$

eigenvalue ($\lambda$) eigenvector ($\mathbf{x}$)

Solve the equation: $(A - \lambda I)\mathbf{x} = 0$ (1)

x – is in the null space of $(A - \lambda I)$
λ is chosen such that $(A - \lambda I)$ has a null space

Computation of eigenvalues and eigenvectors (for dim 2,3)

1. Compute determinant
2. Find roots (eigenvalues) of the polynomial such that determinant = 0
3. For each eigenvalue solve the equation (1)

For larger matrices – alternative ways of computation

## Eigenvalues and Eigenvectors

For the previous example

$$\lambda_1 = -1, x_1 = [1,1]^T \qquad \lambda_2 = -2, x_2 = [5,2]^T$$

We will get special solutions to ODE $\dot{\mathbf{u}} = A\mathbf{u}$

$$A = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix} \quad \mathbf{u} = e^{\lambda_1 t}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \mathbf{u} = e^{\lambda_2 t}\begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

Their linear combination is also a solution (due to the linearity of $\dot{\mathbf{u}} = A\mathbf{u}$

$$\mathbf{u} = c_1 e^{\lambda_1 t}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 e^{\lambda_1 t}\begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

In the context of diff. equations – special meaning
Any solution can be expressed as linear combination
Individual solutions correspond to modes

## Eigenvalues and Eigenvectors

$$A\mathbf{x} = \lambda\mathbf{x}$$

Only special vectors are eigenvectors
- such vectors whose direction will not be changed by the transformation A (only scale)
- they correspond to normal modes of the system act independently

| Examples | eigenvalues | eigenvectors |
|---|---|---|
| $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ | 2, 3 | $\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ |

Whatever A does to an arbitrary vector is fully determined by its eigenvalues and eigenvectors

$$A\mathbf{x} = 2\lambda_1 v_1 + 5\lambda_2 v_2$$

## Eigenvalues and Eigenvectors - Diagonalization

- Given a square matrix A and its eigenvalues and eigenvectors – matrix can be diagonalized

$$A = S\Lambda S^{-1}$$

$$A = S\Lambda S^{-1}$$

Matrix of eigenvectors ↙ ↘ Diagonal matrix of eigenvalues

$$AS = S\Lambda$$

$$A \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} = \begin{bmatrix} \lambda_1 x_1 & \lambda_2 x_2 & \dots & \lambda_n x_n \end{bmatrix} \qquad A\mathbf{x} = \lambda \mathbf{x}$$

$$\begin{bmatrix} \lambda_1 x_1 & \lambda_2 x_2 & \dots & \lambda_n x_n \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 \ \dots & \\ & & \lambda_n \end{bmatrix}$$

$$A = S\Lambda S^{-1}$$

- If some of the eigenvalues are the same, eigenvectors are not independent

## Diagonalization

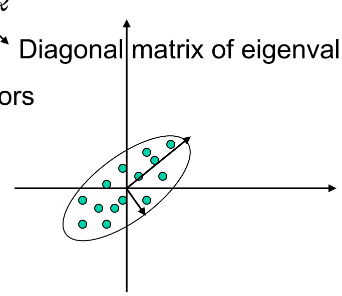
- If there are no zero eigenvalues – matrix is invertible
- If there are no repeated eigenvalues – matrix is diagonalizable
- If all the eigenvalues are different then eigenvectors are linearly independent

For Symmetric Matrices

If A is symmetric $A = Q\Lambda Q^T$

↙ orthonormal matrix of eigenvectors ↘ Diagonal matrix of eigenvalues

i.e. for a covariance matrix
or some matrix B = A^TA



## Dimensionality Reduction

- Next:
- Many dimensions are often interdependent (correlated);

We can:

- Reduce the dimensionality of problems;
- Transform interdependent coordinates into significant and independent ones;