ENM 360: Introduction to Data-driven Modeling

Lecture #3: Approximation of functions and data

Paris Perdikaris
September 5, 2019

Penn
UNIVERSITY *of* PENNSYLVANIA

Numerical differentiation with finite differences

$$\mathbf{A} = \begin{bmatrix} -2 & 1 & & & & & \\ 1 & -2 & 1 & & & & \\ & 1 & -2 & 1 & & & \\ & & 1 & -2 & 1 & & \\ & & & 1 & -2 & 1 & \\ & & & & \ddots & \ddots & \ddots \\ & & & & & 1 & -2 \end{bmatrix}$$

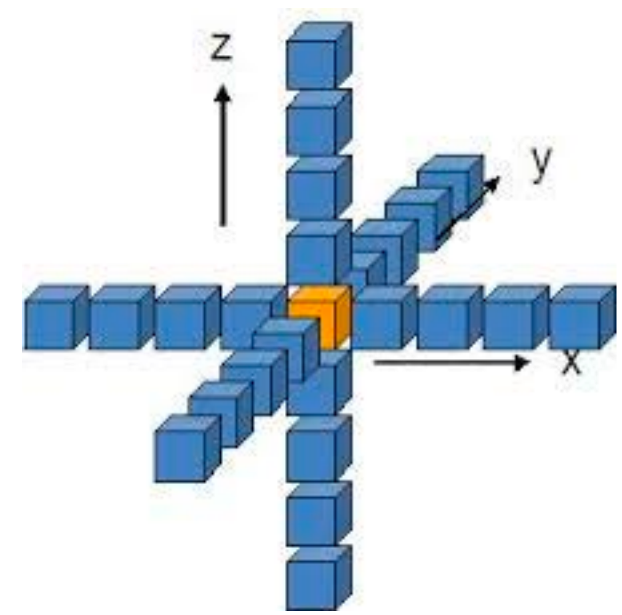
Central difference stencil for second derivative approximation in 1D

$$\left[\begin{array}{cccc|cccc|cccc|cccc}
\mathbf{4} & -1 & & & -1 & & & & & & & & & & & \\
-1 & \mathbf{4} & -1 & & & -1 & & & & & & & & & & \\
 & -1 & \mathbf{4} & -1 & & & -1 & & & & & & & & & \\
 & & -1 & \mathbf{4} & & & & -1 & & & & & & & & \\
\hline
-1 & & & & \mathbf{4} & -1 & & & -1 & & & & & & & \\
 & -1 & & & -1 & \mathbf{4} & -1 & & & -1 & & & & & & \\
 & & -1 & & & -1 & \mathbf{4} & -1 & & & -1 & & & & & \\
 & & & -1 & & & -1 & \mathbf{4} & & & & -1 & & & & \\
\hline
 & & & & -1 & & & & \mathbf{4} & -1 & & & -1 & & & \\
 & & & & & -1 & & & -1 & \mathbf{4} & -1 & & & -1 & & \\
 & & & & & & -1 & & & -1 & \mathbf{4} & -1 & & & -1 & \\
 & & & & & & & -1 & & & -1 & \mathbf{4} & & & & -1 \\
\hline
 & & & & & & & & -1 & & & & \mathbf{4} & -1 & & \\
 & & & & & & & & & -1 & & & -1 & \mathbf{4} & -1 & \\
 & & & & & & & & & & -1 & & & -1 & \mathbf{4} & -1 \\
 & & & & & & & & & & & -1 & & & -1 & \mathbf{4}
\end{array}\right]$$

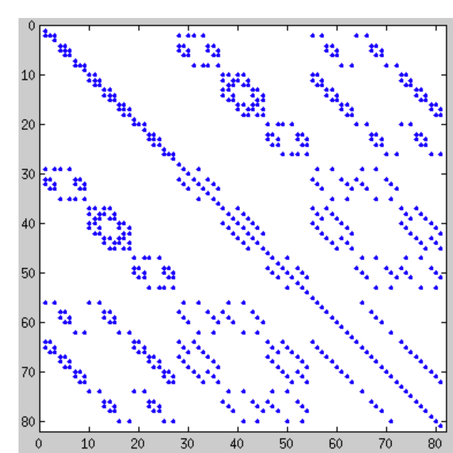
Central difference stencil for second derivative approximation in 2D

1D

2D



3D



Sparsity pattern of a realistic FD stencil matrix

Numerical differentiation with spectral methods

For *continuous* periodic function $f(x)$, $f(x+2\pi) = f(x)$, represented by a Fourier series:

$$f(x) = \sum_{n=-\infty}^{\infty} \hat{f}_n\, e^{inx}$$

The differentiation of $f(x)$ can then be evaluated by:

$$\frac{df(x)}{dx} = \sum_{n=-\infty}^{\infty} \underbrace{(in\hat{f}_n)}_{\text{Fourier coefficient of } f'(x)}\, e^{inx}$$

∴ Once the coefficients of the Fourier series $\hat{f}_n$ is obtained, the differentiation can be evaluated by summing the Fourier series with new coefficients $(in\hat{f}_n)$

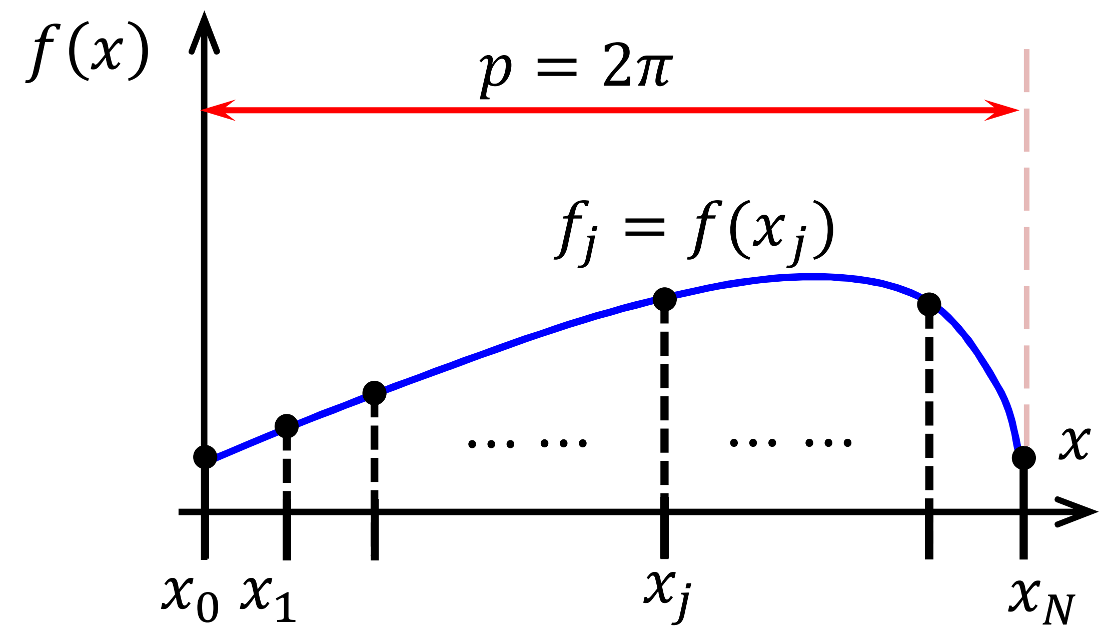
Numerical differentiation with spectral methods

- Now, for discrete periodic function f_j defined at x_j, $j = 0, 1, \ldots, N-1$:

$$f_j = \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} \hat{f}_n\, e^{inx_j} = \sum_{n=0}^{N-1} \hat{f}_n\, e^{inx_j}$$

where $\hat{f}_n$ is the discrete Fourier transform:

$$\hat{f}_n = \frac{1}{N} \sum_{j=0}^{N-1} f(x_j) e^{-inx_j}$$



- Can the differentiation of f defined at x_j, i.e., $\left.\frac{df}{dx}\right|_j$ be evaluated by:

$$\left.\frac{df}{dx}\right|_j = \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} (in) \hat{f}_n\, e^{inx_j}$$

Similarly,

$$\left.\frac{d^2 f}{dx^2}\right|_j = \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} (in)^2 \hat{f}_n\, e^{inx_j} = -\sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} n^2 \hat{f}_n\, e^{inx_j}$$

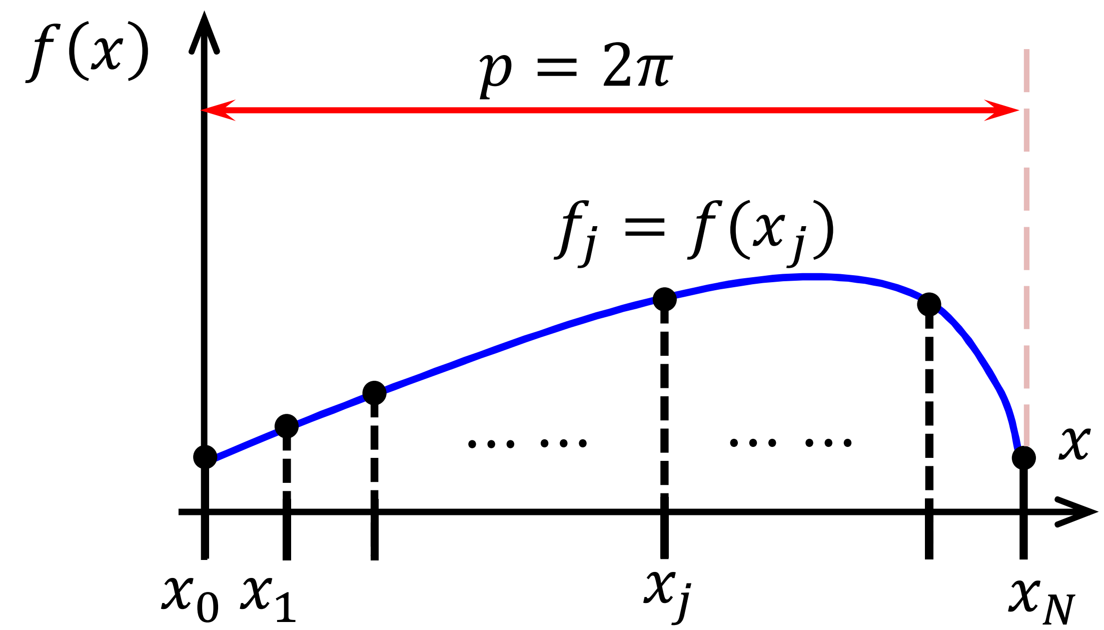
Numerical differentiation with spectral methods

- Now, for discrete periodic function f_j defined at x_j, $j = 0, 1, \dots, N-1$:

$$f_j = \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} \hat{f}_n\, e^{inx_j} = \sum_{n=0}^{N-1} \hat{f}_n\, e^{inx_j}$$

where $\hat{f}_n$ is the discrete Fourier transform:

$$\hat{f}_n = \frac{1}{N}\sum_{j=0}^{N-1} f(x_j) e^{-inx_j}$$



- Can the differentiation of f defined at x_j, i.e., $\left.\frac{df}{dx}\right|_j$ be evaluated by:

$$\left.\frac{df}{dx}\right|_j = \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} (in)\hat{f}_n\, e^{inx_j}$$

Similarly,

$$\left.\frac{d^2 f}{dx^2}\right|_j = \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} (in)^2 \hat{f}_n\, e^{inx_j} = -\sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} n^2 \hat{f}_n\, e^{inx_j}$$

Numerical differentiation with spectral methods

- The spectral derivative is much more accurate than any finite-difference schemes for *periodic functions*.
- The major cost involved is the use of fast Fourier transform.
- However, it is inaccurate and does not converge when the derivative is discontinuous.

Numerical integration: The midpoint rule

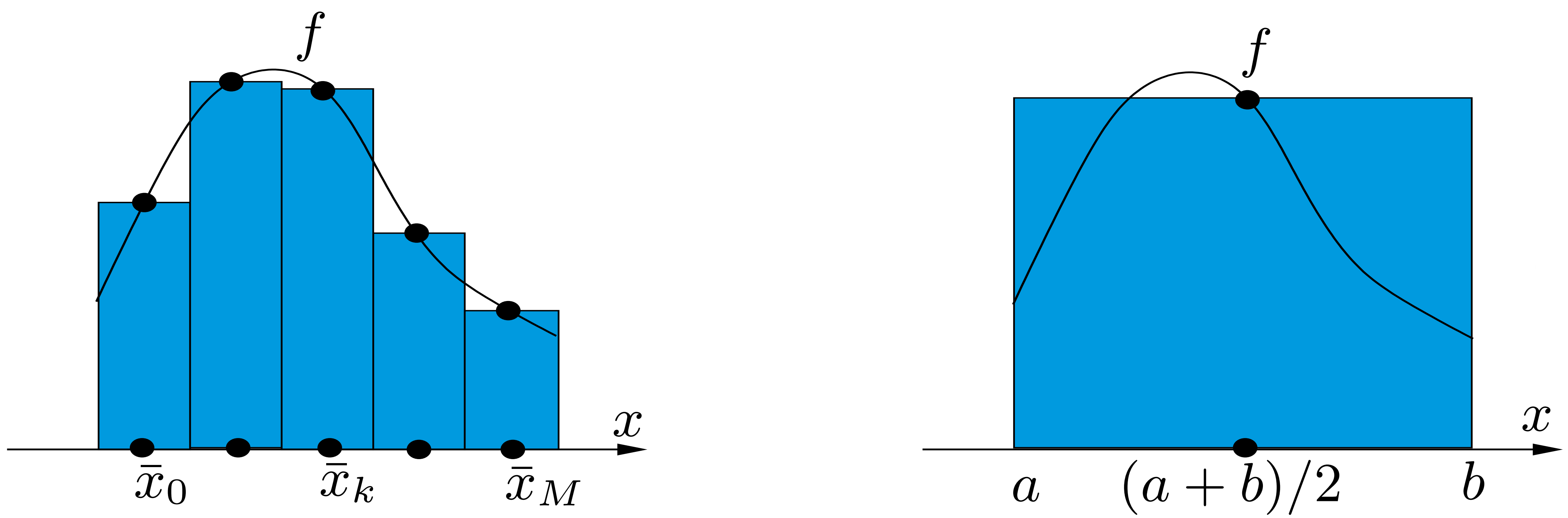


Fig. 4.3. The composite midpoint formula (*left*); the midpoint formula (*right*)

$$I^c_{mp}(f) = H \sum_{k=1}^{M} f(\bar{x}_k)$$

$$I_{mp}(f) = (b-a)f[(a+b)/2]$$

Numerical integration: The trapezoidal rule

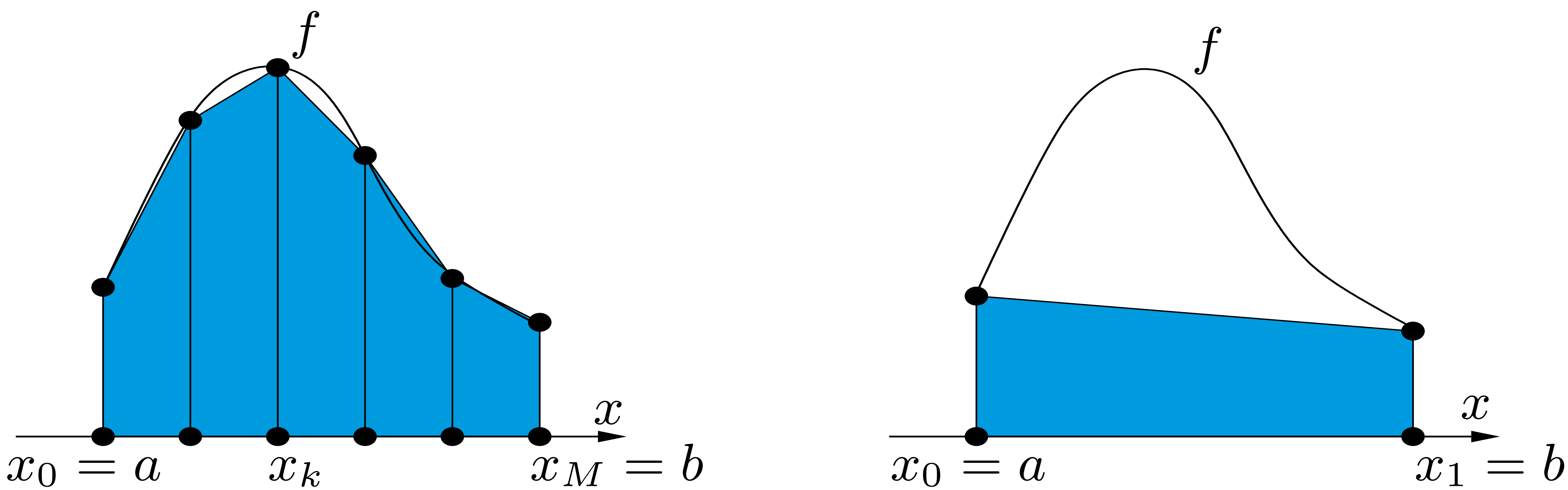


Fig. 4.4. Composite trapezoidal formula (*left*); trapezoidal formula (*right*)

$$\begin{aligned} I_t^c(f) &= \frac{H}{2}\sum_{k=1}^{M}[f(x_k) + f(x_{k-1})] \\ &= \frac{H}{2}[f(a) + f(b)] + H\sum_{k=1}^{M-1} f(x_k) \end{aligned}$$

$$I_t(f) = \frac{b-a}{2}[f(a) + f(b)]$$

Numerical integration: Simpson's rule

$$I_s(f) = \frac{b-a}{6}\left[f(a) + 4f((a+b)/2) + f(b)\right]$$

Simpson's formula

$$I_s^c(f) = \frac{H}{6}\sum_{k=1}^{M}\left[f(x_{k-1}) + 4f(\bar{x}_k) + f(x_k)\right]$$

The composite Simpson's rule

Gauss-Legendre quadrature

$$I_s(f) = \sum_{j=1}^{n} w_j f(x_j)$$

w_j *weights*

x_j *nodes*

n	x_j	w_j
1	$\{\pm 1/\sqrt{3}\}$	$\{1\}$
2	$\{\pm\sqrt{15}/5, 0\}$	$\{5/9, 8/9\}$
3	$\Big\{\pm(1/35)\sqrt{525 - 70\sqrt{30}},$	$\{(1/36)(18 + \sqrt{30}),$
	$\pm(1/35)\sqrt{525 + 70\sqrt{30}}\Big\}$	$(1/36)(18 - \sqrt{30})\}$
4	$\Big\{0, \pm(1/21)\sqrt{245 - 14\sqrt{70}}$	$\{128/225, (1/900)(322 + 13\sqrt{70})$
	$\pm(1/21)\sqrt{245 + 14\sqrt{70}}\Big\}$	$(1/900)(322 - 13\sqrt{70})\}$

Table 4.1. Nodes and weights for some quadrature formulae of Gauss-Legendre on the interval $(-1, 1)$. Weights corresponding to symmetric couples of nodes are reported only once

Monte Carlo approximation

$$\mathbb{E}_{x\sim p(x)}[f(x)] = \int f(x)p(x)dx \approx \frac{1}{n}\sum_{i=1}^{n} f(x_i),$$

where x_i are drawn iid from $p(x)$

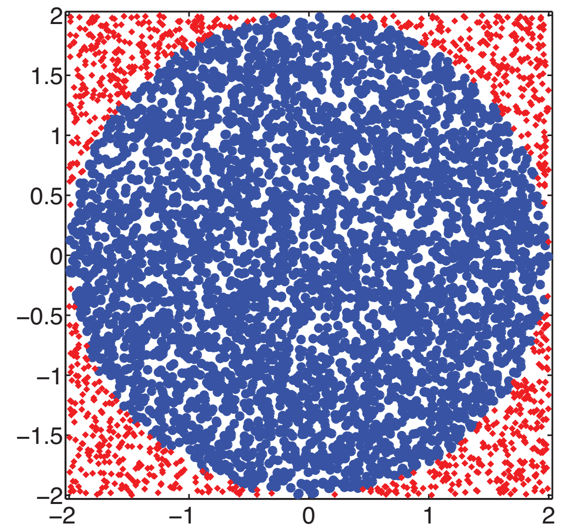
Monte Carlo approximation

Example: estimating π by Monte Carlo integration

MC approximation can be used for many applications, not just statistical ones. Suppose we want to estimate π. We know that the area of a circle with radius r is πr^2, but it is also equal to the following definite integral:

$$I = \int_{-r}^{r} \int_{-r}^{r} \mathbb{I}(x^2 + y^2 \leq r^2) dx dy \tag{2.99}$$

Hence $\pi = I/(r^2)$. Let us approximate this by Monte Carlo integration. Let $f(x,y) = \mathbb{I}(x^2 + y^2 \leq r^2)$ be an indicator function that is 1 for points inside the circle, and 0 outside, and let $p(x)$ and $p(y)$ be uniform distributions on $[-r, r]$, so $p(x) = p(y) = 1/(2r)$. Then



$$\begin{aligned} I &= (2r)(2r) \int\int f(x,y)p(x)p(y)dxdy && (2.100) \\ &= 4r^2 \int\int f(x,y)p(x)p(y)dxdy && (2.101) \\ &\approx 4r^2 \frac{1}{S} \sum_{s=1}^{S} f(x_s, y_s) && (2.102) \end{aligned}$$

We find $\hat{\pi} = 3.1416$ with standard error 0.09 (see Section 2.7.3 for a discussion of standard errors). We can plot the points that are accepted/ rejected as in Figure 2.19.