Title

halton() — Generate a Halton or Hammersley sequence

Syntax

real matrix `halton(`*real scalar n*`,` *d* `[,` *real scalar burn* `[,` *real scalar method*`]])`
void `_halton(`*real matrix x*`,` `[,` *real scalar burn* `[,` *real scalar method*`]])`

where inputs are

n: length of the sequence
d: dimension of the sequence (≤ 20)
x: an $n \times d$ real matrix
burn: starting index of the sequence, default 0
method: type of sequence, 1=Halton (default), 2=Hammersley

Description

This routine generates either a Halton, the default, or a Hammersley sequence of length n and dimension $d \leq 20$. These sequences have a more uniform scatter over the unit hypercube than sequences generated from pseudo-uniform random numbers.

Remarks

The Halton sequence of dimension d is generated from the first d primes, p_k, so that on draw i we have $\mathbf{h}_i = (r_{p_1}(i), r_{p_2}(i), \ldots, r_{p_d}(i))$, where

$$r_{p_k}(i) = \sum_{j=0}^{q} b_{jk}(i) p_k^{-j-1} \in (0,1)$$

is the radical inverse function of i with base p_k so that $\sum_{j=0}^{q} b_{jk}(i) p_k^j = i$ where $p_k^q \leq i < p_k^{q+1}$ (Fan and Wang, 1994).

We demonstrate with base $p_3 = 5$, $i = 33$, and we will generate $r_5(33)$. Here, $q = 2$, $b_{0,3}(33) = 3$, $b_{1,5}(33) = 1$ and $b_{2,5}(33) = 1$, so that $r_5(33) = \frac{3}{5} + \frac{1}{25} + \frac{1}{625}$.

The Hammersley sequence uses an evenly spaced set of points with the first d-1 components the Halton sequence, $\mathbf{h}_i = \left(\frac{2i-1}{2n}, r_{p_1}(i), r_{p_2}(i), \ldots, r_{p_{d-1}}(i)\right)$, $i = 1, \ldots, n$.

The discrepancy of a sequence, denoted $D(n,d)$, is a measure of its nonuniformity over the unit hypercube. For all rectangles on the unit hypercube $(\mathbf{0}, \mathbf{t}), t = (t_1, \ldots, t_d), t_j \leq 1$, it measures the maximum of the absolute value of the difference between the ratio of the number of points in $(\mathbf{0}, \mathbf{t})$ and the total number of points, nd, and the volume of $(\mathbf{0}, \mathbf{t})$. See Section 1.4 of Fang and Wang (1994) for a more rigorous definition. The discrepancy of the uniform pseudo-random sequence has a convergence rate that is independent of d, $D_{uni}(n) \leq K\sqrt{\frac{\log\log n}{n}}$ for some $K > 0$. For the Halton sequence we have a faster rate of convergence, $D_{hal}(n,d) \leq \dfrac{K(d)(\log n)^d}{n}$, while the Hammersley sequence has an improvement over the Halton, $D_{ham}(n,d) \leq \dfrac{K(d)(\log n)^{d-1}}{n}$.

Conformability

`halton(`*n*, *d*, *burn*, *method*`)`

input:

n: *1 x 1*
d: *1 x 1*
burn: *1 x 1*
method: *1 x 1*

output:

result: *n x d*

`_halton(`*x*, *burn*, *method*`)`

input:

x: *n x d*
burn: *1 x 1*
method: *1 x 1*

output:

x: *n x d*

Diagnostics

`halton(`*n*, *d*`)` returns an $n \times d$ matrix containing *d* sequences of length *n*. The first row of the matrix will have the sequence at index *burn*.

`_halton(`*x*`)` modifies the $n \times d$ matrix *x* so that it contains *d* sequences of length *n*.

The maximim dimension, *d*, is 20.

References

Fang, K.T. and Y. Wang 1994. *Number-Theoretic Methods in Statistics* London:Chapman & Hall.