Numerical Optimization 13: Sampling Plans

Qiang Zhu

University of Nevada Las Vegas

May 20, 2020

Overview

1. Sampling
2. Full Factorial
3. Random Sampling
4. Uniform Projection Plans
5. Stratified Sampling
6. Space-Filling Metrics
7. Quasi-Random Sequences
8. Summary

Optimization with expensive function evaluations

For many optimization problems, function evaluations can be quite expensive.

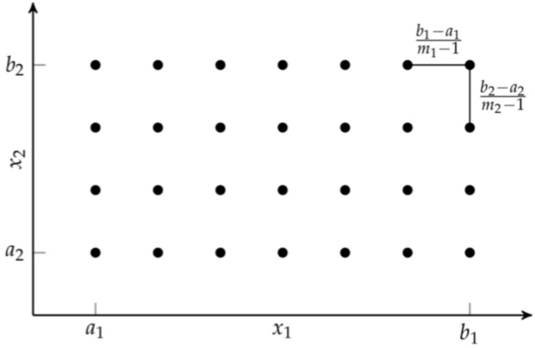
- an aircraft design may require a wind tunnel test
- deep learning hyperparameters may require a week of GPU training
- $\cdots$

A common approach for optimizing in these contexts is to build a surrogate model, Further evaluations of the true objective function can be used to improve the model. Fitting such models requires an initial set of points, ideally points that are space-filling; that is, points that cover the region as well as possible.

Full Factorial

The full factorial sampling plan places a grid of evenly spaced points over the search space.

- a lower/upper-bound vector a, b such that $a_i \leq x_i \leq b_i$
- m_i samples in each x_i separated by a distance $(b_i - a_i)/(m_i - 1)$

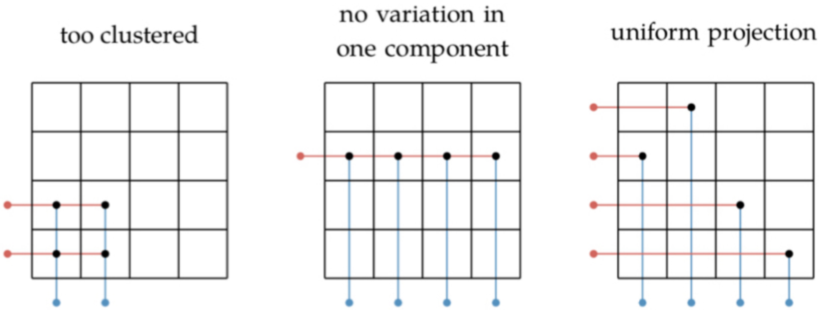


Random Sampling

In some cases, it may be possible to transform a problem so that constraints can be removed. For example, bound constraints $a \quad x \quad b$ can be removed by passing x through a transform

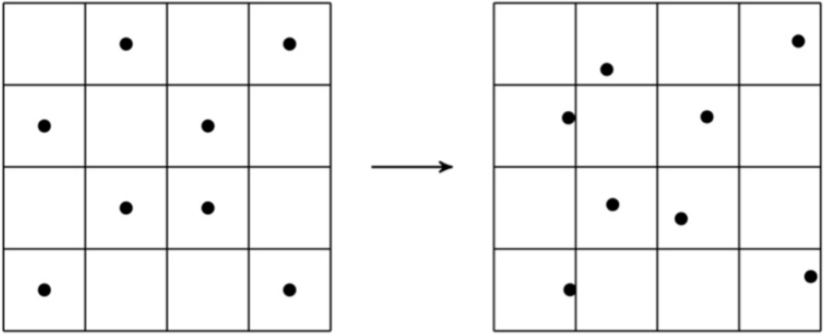
Uniform Projection Plans

A uniform projection plan with m samples on an $m \times m$ grid can be constructed using an m-element permutation. There are therefore $m!$ possible uniform projection plans.



Uniform Projection Plans

Stratified sampling modifies any grid-based sampling plan, including full factorial and uniform projection plans. Cells are sampled at a point chosen uniformly at random from within the cell rather than at the cells center



Space-Filling Metrics

A good sampling plan fills the design space since the ability for a surrogate model to generalize from samples decays with the distance from those samples. Not all plans, even uniform projection plans, are equally good at covering the search space.

- Discrepancy, the maximum difference between the fraction of samples in a hyper-rectangular subset H and that subsets volume:

$$d(X) = \sup_{H} \left| \frac{\#X \cap H}{\#X} - \lambda(H) \right|$$

  where $\#X$ and $\#X \cap H$ are the numbers of X points and X in H.
- Pairwise Distances between all points within each sampling plan

Quasi-Random Sequences

Quasi-random sequences are often used in the context of trying to approximate an integral over a multidimensional space:

$$\int_{\chi} f(\boldsymbol{x}) d\boldsymbol{x} \approx \frac{v}{m} \sum_{i=1}^{m} f(\boldsymbol{x}^i)$$

where each $\boldsymbol{x}^i$ is sampled uniformly at random over the domain X and v is the volume of χ .
Quasi-random sequences are deterministic sequences that fill the space in a systematic manner so that the integral converges as fast as possible in the number of points m. They are typically constructed for the unit n-dimensional hypercube with the following methods.

- Additive Recurrence
- Halton Sequence
- Sobol Sequence

Additive Recurrence

Quasi-random sequences are often used in the context of trying to approximate an integral over a multidimensional space:

$$x^{k+1} = x^k + c \pmod 1$$

produce space-filling sets provided that c is irrational. The value of c leading to

$$c = 1 - \Phi = \frac{\sqrt{5} - 1}{2} = 0.618$$

where Φ is the golden ratio. We can construct a space-filling set over n dimensions using an additive recurrence sequence for each coordinate, each with its own value of c. The square roots of the primes are known to be irrational, and can thus be used to obtain different sequences for each coordinate:

$$c_1 = \sqrt{2},\ c_2 = \sqrt{3},\ c_3 = \sqrt{5}, c_4 = \sqrt{7}, c_5 = \sqrt{11},$$

Halton Sequence

Radical Inversion

$$i = \sum_{k=0}^{M-1} a_k(i) b^k$$

$$\Psi_{b,C} = (b^{-1}, \cdots, b^{-M})[C(a_0(i), \cdots, a_M(i))^T]$$

where b is the base number, and C is the generator matrix. When C is the identity matrix, it is called van der Corput sequences,

- $b = 2$

$$X = \left\{ \frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{5}{8}, \frac{3}{8}, \frac{7}{8}, \frac{1}{16}, \dots \right\}$$

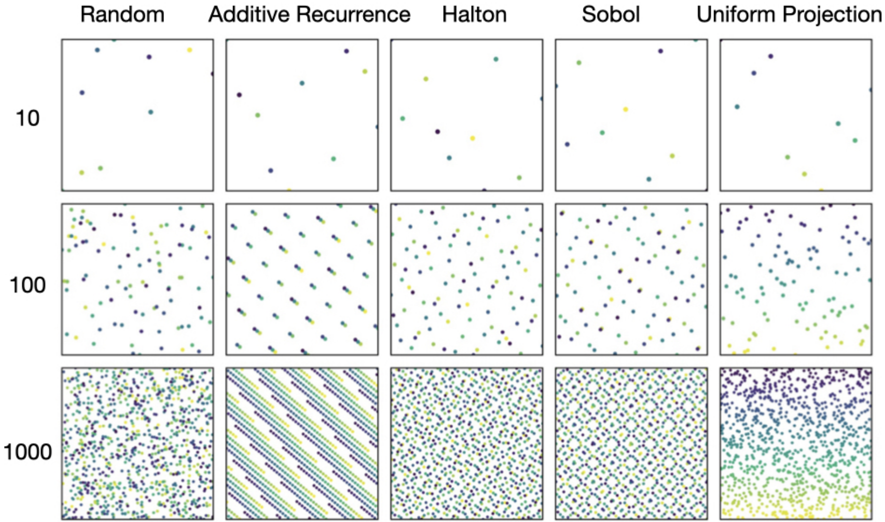
- $b = 5$

$$X = \left\{ \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \frac{1}{25}, \frac{6}{25}, \frac{11}{25}, \dots \right\}$$

Halton Sequence uses coprime numbers in order to be uncorrelated.

Sobol Sequence

In the Sobol sequence, each dimension uses the base 2 with different C.



Summary

- Sampling plans are used to cover search spaces with a limited number of points.
- Full factorial sampling, which involves sampling at the vertices of a uniformly discretized grid, requires a number of points exponential in the number of dimensions.
- Uniform projection plans,which project uniformly over each dimension,can be efficiently generated and can be optimized to be space filling.
- Quasi-random sequences are deterministic procedures by which space-filling sampling plans can be generated.