

Планирование экспериментов

ПМИ ФКН ВШЭ, 24 ноября 2018 г.

Денис Деркач¹

¹ФКН ВШЭ

Денис Деркач

Оглавление

Постановка задачи

Заполнение пространства

- Random sampling

- Семплирование латинскими гиперкубами

- Полный факторный дизайн

- Halton sequence

- Свойства

Поверхность отклика

- Оптимальные Дизайны RSM

- Примеры

- RSM и категориальные переменные

- DoE для RSM

Адаптивный Дизайн

- Техника Адаптивного DoE

Денис Деркач

Регрессия гауссовского процесса

Постановка задачи

Мотивация

Типы задач:

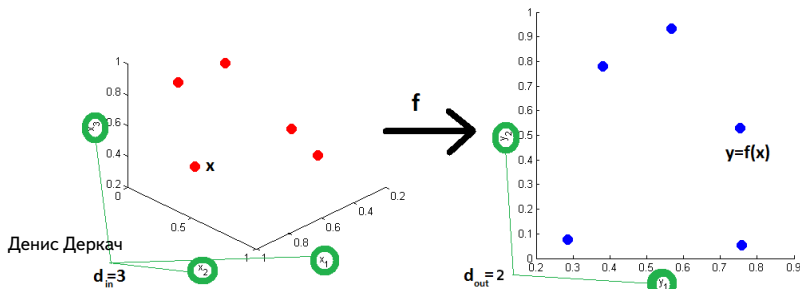
- › сравнительный эксперимент - сравнение двух разных моделей;
- › наблюдательный эксперимент - упрощение модели;
- › построение поверхности отклика;
- › регрессия.

Во всех случаях стоимость вычислений может быть достаточно высокой

Данные

Определение

- › \mathbb{X} — d_{in} -мерное пространство.
- › $\mathbf{x} \in \mathbb{X}$ — d_{in} -мерный вектор, описывающий дизайн объекта.
- › \mathbf{y} — d_{out} -мерный вектор характеристик объекта.
- › Точная зависимость $f : \mathbf{x} \mapsto \mathbf{y}$.



Суррогатное моделирование

Определение

- › $X = \{\mathbf{x}_i\}_{i=1}^N$, $\mathbf{x}_i \in X$ называется планом эксперимента.
- › $D = (X, Y = f(X)) = \{(\mathbf{x}_1, \mathbf{y}_1 = f(\mathbf{x}_1)), \dots, (\mathbf{x}_N, \mathbf{y}_N = f(\mathbf{x}_N))\}$ — обучающая выборка размера N .

Пример

	Вход \leftrightarrow компоненты \mathbf{x}			Выход (\mathbf{y})
	Angle of attack	...	Mach	Lift coeff.
Объект 1	0.5	...	0.60	0.42
Объект 2	1.4	\ddots	0.77	0.62
\vdots	...	\ddots	\ddots	...
Объект N	3.2	...	0.66	0.55

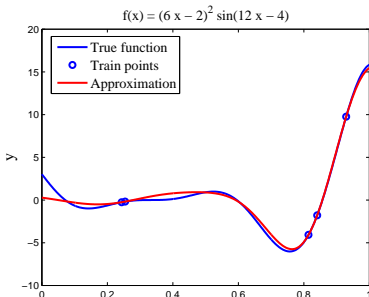
Эксперимент

Определение

$\hat{f}(x) \approx f(x)$ суррогатная модель (функция), построенная по обучающей выборке $D = (X, Y = f(X))$.

Требование к \hat{f} :

- › Легко и эффективно оценивается.
- › “Близка” к f по значениям не только из D .



Заполнение пространства

Постановка задачи

Заполнение пространства

Random sampling

Семплирование латинскими гиперкубами

Полный факторный дизайн

Halton sequence

Свойства

Поверхность отклика

Оптимальные Дизайны RSM

Примеры

RSM и категориальные переменные

DoE для RSM

Адаптивный Дизайн

Техника Адаптивного DoE

Регрессия гауссовского процесса

Выход суррогатной модели

Критерий сэмплирования

Integrated MSE Gain-Maximum Variance

Поверхность критерия максимальной дисперсии

MaxMin

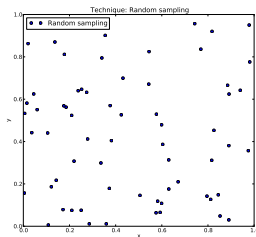
Пример

Space-filling Design of Experiments

Определение

Space-filling DoE — это равномерное сэмплирование в гиперкубе.

Предположение: мы ничего не знаем о суррогатной модели и указанных зависимостях. Поэтому будем заполнять пространство дизайна равномерно.



Random sampling

Метод состоит в равномерной генерации точек в гиперкубе.

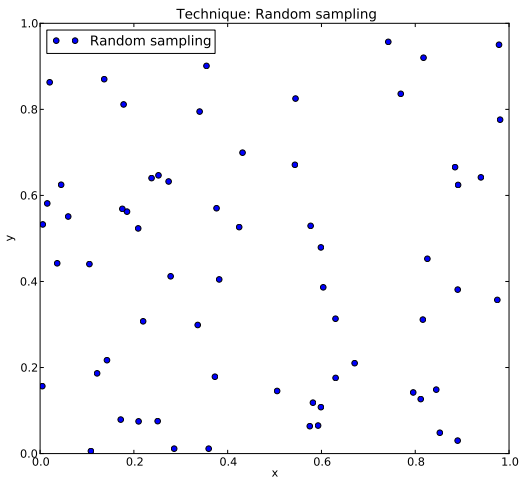
Преимущества

- › Универсальность и гибкость.
- › Всегда можно расширить добавлением точек.

Недостатки

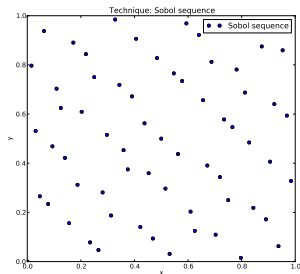
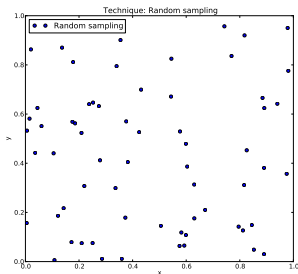
- › Нет гарантий равномерности заполнения.

Random sampling



Свойства: равномерность

Равномерность является важным свойством space-filling DoE.



Ожидаемые свойства:

- › Отсутствие "дырок" в пространстве дизайна.
- › Равномерность малоразмерных проекций.

Критерии равномерности

Пусть $\mathbb{X} = [0, 1]^d$.

› Discrepancy

$$d(X) = \sup_{0 \leq u_k < v_k < 1} \left| \frac{\#(X \cap P_{u,v})}{N} - |P_{u,v}| \right|, \quad P_{u,v} = \bigotimes_k [u_k, v_k),$$

› Минимаксное расстояние между точками

$$\rho(X) = \max_i \min_{j, j \neq i} \|\mathbf{x}_i - \mathbf{x}_j\|$$

› ϕ -метрика

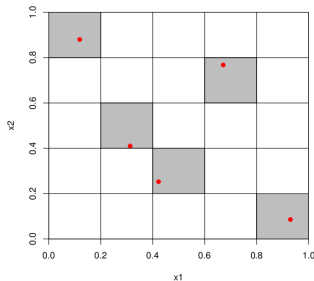
$$\phi_p(X) = \left(\sum_{i < j} \|\mathbf{x}_i - \mathbf{x}_j\|^{-p} \right)^{1/p}$$

Чем равномернее, тем ниже значение метрики.

Латинские гиперкубы

Определение

Семплирование латинским гиперкубом выполняется при помощи разделения значений каждой компоненты дизайна на N равных интервалов, в каждый из которых попадёт по одной точке.

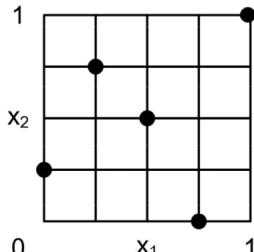
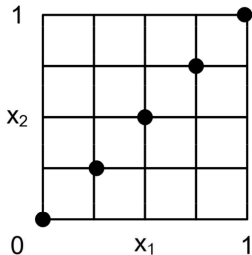


Оптимизированный LHS (OLHS)

Определение

Optimized Latin hypercube sampling это улучшенная версия LHS.

- › LHS может иногда давать нежелательный результат.
- › OLHS генерирует множество LHS дизайнов, а потом выбирает лучший по равномерной метрике.
- › Это сильно увеличивает робастность, но требует много времени.



Optimized LHS (OLHS)

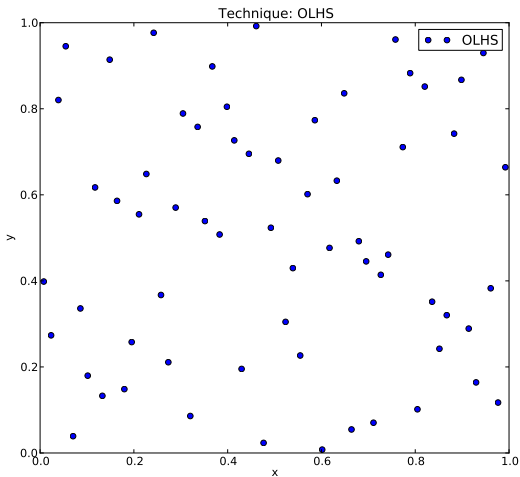
Преимущества

- › Проекции на оси достаточно равномерны.
- › Простота.

Недостатки

- › Оптимизация очень медленная.
- › Практически невозможно дополнить множество дизайна без нарушения свойств латинского гиперкуба (так как заранее заданное число N определяет число интервалов).

Optimized LHS (OLHS)



Полный факторный план эксперимента

Определение

Полный факторный план эксперимента — все возможные комбинации уровней переменной дизайна, то есть все прямые произведения $X = \prod_{j=1}^{d_{in}} L_j = L_1 \times L_2 \times \dots \times L_{d_{in}}$ конечных множеств $L_1, L_2, \dots, L_{d_{in}}$, где L_i — это конечное множество из n_i точек интервала.

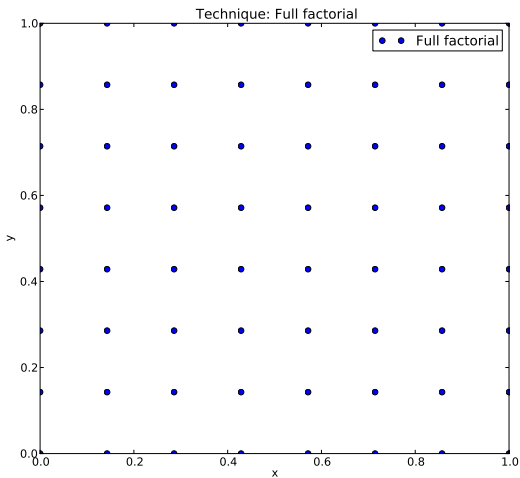
Преимущества

- Хорошо заполняет пространство.
- Очень быстро генерируется.

Недостатки

- Требуемое количество точек растёт слишком быстро.

Full—Factorial



Halton sequence

Определение

- › Для заданной константы p определим ψ_p как

$$\psi_p(n) = \sum_{i=0}^{R(n)} \frac{a_i(n)}{p^{i+1}},$$

где a_i это числа из записи n в системе счисления с основанием p , а $R(n)$ обозначает максимальный индекс, для которого $a_i(n)$ не ноль.

- › n -й элемент последовательности Холтона это

$$\mathbf{x}_n = (\psi_{p_1}(n), \dots, \psi_{p_{d_{in}}}(n)),$$

где p_i это i -е простое число.

Halton sequence

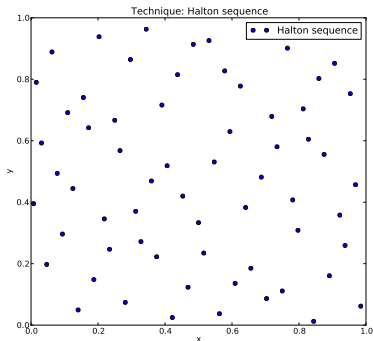
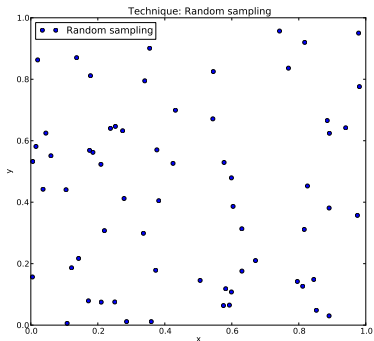
Преимущества

- › Эффективность в малых размерностях.
- › Всегда можно расширить добавлением новых точек.
- › Быстрая генерация.

Недостатки

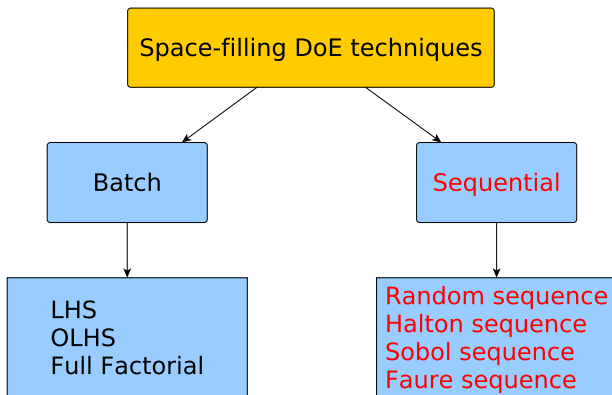
- › Проекции с высокой размерностью ($d_{in} > 6$) сильно коррелированы.

Halton sequence



$$d_{in} = 2, N = 64$$

Техники Space-filling DoE



Свойства: Случайность или предопределённость

Некоторые техники DoE используют псевдо-случайные числа. Преимущество заключается в том, что результаты оказываются более разнообразными, но с другой стороны их труднее воспроизводить.

- › Включающие псевдо-случайность: Random sampling, LHS, OLHS.
- › Полностью детерминированные: Full Factorial, Halton sequence, Sobol sequence, Faure sequence.

Свойства: Добавление новых точек

Часть техник Space-filling поддерживает добавление новых точек к сгенерированному дизайну, а остальные не поддерживают из-за своих особенностей.

- › **Поддерживающие добавление точек:** Random sampling, Halton sequence, Sobol sequence, Faure sequence. Эти техники могут быть использованы для построения дизайна от массива точек, так и в режиме последовательного добавления.
- › **Сгенерированный дизайн не может быть расширен:** Full Factorial, LHS, OLHS. Эти техники позволяют только использование всех точек сразу.

Поверхность отклика

Постановка задачи

Заполнение пространства

Random sampling

Семплирование латинскими гиперкубами

Полный факторный дизайн

Halton sequence

Свойства

Поверхность отклика

Оптимальные Дизайны RSM

Примеры

RSM и категориальные переменные

DoE для RSM

Адаптивный Дизайн

Техника Адаптивного DoE

Регрессия гауссовского процесса

Выход суррогатной модели

Критерий сэмплирования

Integrated MSE Gain-Maximum Variance

Поверхность критерия максимальной дисперсии

MaxMin

Пример

RSM

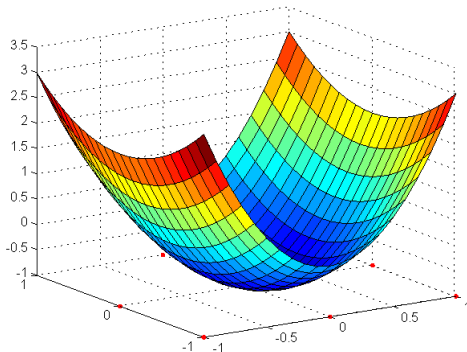
Определение

Response Surface Model (RSM):

$$\hat{f}(\mathbf{x}) = \alpha_0 + \sum_{i=1}^{d_{in}} \alpha_i x_i + \sum_{i,j=1, i \leq j}^{d_{in}} \beta_{ij} x_i x_j \longleftrightarrow$$
$$\mathbf{x} = (x_1, \dots, x_{d_{in}}) \in \mathbb{X}.$$

Параметры RSM настраиваются по обучающей выборке $D = (X, Y = f(X))$ с заранее заданным $X = (\mathbf{x}_i)_{i=1}^N$.

RSM



Модели RSM:

- › линейная (все $\beta_{ij} = 0$)
- › RSM с перекрёстными членами ($\beta_{ii} = 0$ для всех i)
- › квадратичная

Оптимальные Дизайны RSM: введение

Response Surface Model (RSM):

$$\hat{f}(\mathbf{x}) = \alpha_0 + \sum_{i=1}^{d_{in}} \alpha_i x_i + \sum_{i,j=1, i \leq j}^{d_{in}} \beta_{ij} x_i x_j \longleftrightarrow$$
$$\mathbf{x} = (x_1, \dots, x_{d_{in}}) \in \mathbb{X}.$$

Определение

Оптимальный Дизайн RSM это оптимальный $X = (\mathbf{x}_i)_{i=1}^N$, минимизирующий:

1. дисперсию оценки параметров RSM,
2. дисперсию прогноза модели,
3. другие оценки.

Оптимальные Дизайны для RSM

Оптимальные Дизайны минимизируют для RSM одну из величин:

- дисперсию оценки параметров RSM (D-optimality),
или
- дисперсию прогноза модели. (IV-optimality).

Преимущества

Оптимальные дизайны уменьшают стоимость эксперимента посредством того, что статистическая модель может быть оценена через меньшее число запусков.

Критерий оптимальности: Обозначения

- › RSM может быть записан в виде

$$\hat{f}(\mathbf{x}) = \psi(\mathbf{x})\mathbf{c},$$

где $\mathbf{c} = (\mathbf{a}, \mathbf{b})$ и ψ — соответствующее отображение.

Например

$\psi(\mathbf{x}) = (1, x_1, \dots, x_{d_{in}}, x_1x_2, x_1x_3, \dots, x_{d_{in}-1}x_{d_{in}})$ для RSM с перекрёстными членами.

- › $\psi(X)$ будет обозначать $(\psi(\mathbf{x}_1), \dots, \psi(\mathbf{x}_N))$.

Критерий оптимальности: Мотивация

- › Предположим, что шум нормальный: $\mathbf{c} \sim \mathcal{N}(\hat{\mathbf{c}}, \text{Cov}(\hat{\mathbf{c}}|X))$.
- › Метод наименьших квадратов даёт оценку ковариации $\hat{\mathbf{c}}$.

$$\text{Cov}(\hat{\mathbf{c}}|X) \sim (\psi(X)^T \psi(X))^{-1}.$$

- › Аналогично для прогноза ответа
 $f(\mathbf{x}_0) \sim \mathcal{N}(\hat{f}(\mathbf{x}_0), \text{Var}(\hat{f}|X))$.
- › Дисперсия прогноза $\hat{f}(\mathbf{x}_0) = \psi(\mathbf{x}_0)\hat{\mathbf{c}}$ в точке $\mathbf{x}_0 \in \mathbb{X}$.

$$\text{Var}(\hat{f}|X) \sim \psi(\mathbf{x}_0) (\psi(X)^T \psi(X))^{-1} \psi(\mathbf{x}_0)^T.$$

Критерий оптимальности:

D-оптимальность

- › Критерий D-оптимальности даёт такой дизайн, что детерминант должен быть минимален:

$$\det \left[\left(\psi(X)^T \psi(X) \right)^{-1} \right] \rightarrow \min_X .$$

- › D-оптимальный дизайн минимизирует дисперсию оценки параметров.

Критерий оптимальности:

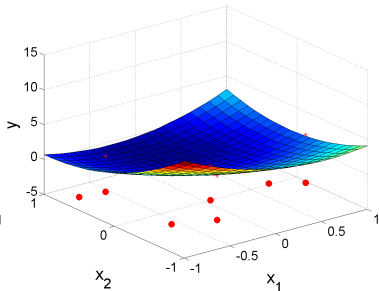
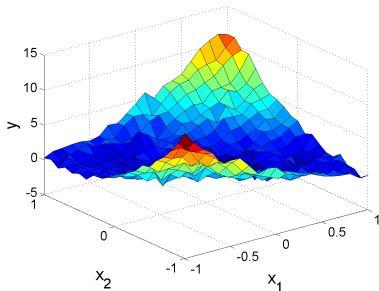
IV-оптимальность

- › IV-оптимальный дизайн минимизирует дисперсия итогового прогноза модели:

$$\int_{\mathbb{X}} \psi(\mathbf{x}) \left(\psi(X)^T \psi(X) \right)^{-1} \psi(\mathbf{x})^T d\mathbf{x} \rightarrow \min_X .$$

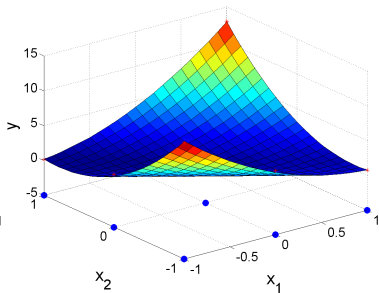
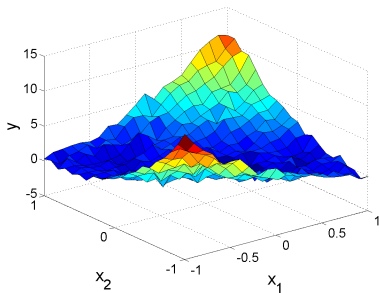
RSM: Пример 1

Настоящая функция ответа с шумом(слева) и RSM обученный со случайным DoE:



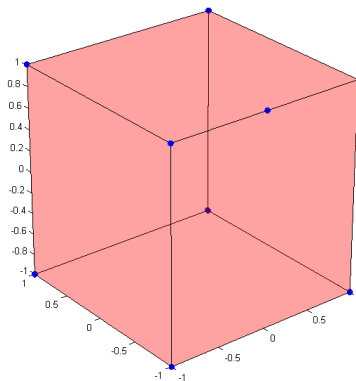
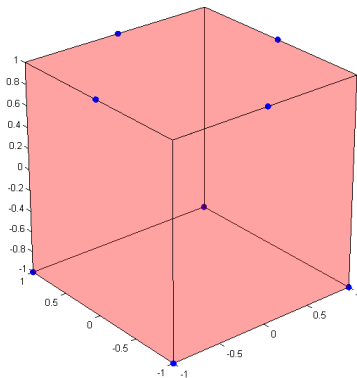
RSM: Пример 1

Настоящая зависимость (слева) и RSM обученный на D-оптимальном дизайне:



RSM: Пример 2

Оптимальный Дизайн для $\mathbb{X} \subset \mathbb{R}^3$:



RSM и категориальные переменные

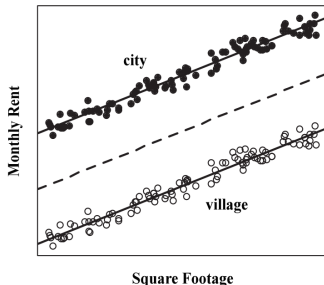
Определение

Категориальные переменные принимают одно из конечного числа значений. Они не имеют численного смысла или порядка.

Пример

Например, “цвет” (черный, зелёный, красный), “пол” (мужской, женский) и т.д.

Рисунок: линейная регрессия между месячной платой, типом жилья и квадратной площадью.



Категориальные переменные в RSM

Пусть $x \in \{v_1, \dots, v_m\}$

1. Для каждого значения v_i строить свою модель RSM.
2. Dummy-variables. Заменяем x на $y = (y_1, \dots, y_m)$, где

$$y_i = \begin{cases} 1, & \text{если } x = v_i \\ 0 & \text{в противном случае.} \end{cases}$$

Критерий оптимальности: Алгоритм оптимизации

Оптимизационная процедура для поиска D- и IV- оптимальных дизайнов опирается на алгоритм Федорова:

Алгоритм

1. Установить число уровней для каждой переменной.
2. Сгенерировать полный факторный дизайн состоящий из всех комбинаций численных значений уровней. Это множество называется **множеством кандидатов** для Оптимального Дизайна.

Критерий оптимальности: Алгоритм оптимизации

Продолжение алгоритма

3. Взять необходимое число точек из множества кандидатов случайным образом и посчитать для них оптимальный критерий.
4. Продолжить оптимизацию добавлением и исключением точек из дизайна с целью минимизировать значение функционала. Новые точки берутся из множества кандидатов.

DoE для RSM

Преимущества

- › Даёт наилучшие возможные оценки для Response Surface Models.
- › Требуется меньшее число запусков эксперимента, чем классические дизайны с той же точностью.
- › Позволяет использовать категориальные переменные.

Недостатки

- › Эффективность доказана только для Response Surface Model с априори заданной структурой.

Адаптивный Дизайн

Постановка задачи

Заполнение пространства

Random sampling

Семплирование латинскими гиперкубами

Полный факторный дизайн

Halton sequence

Свойства

Поверхность отклика

Оптимальные Дизайны RSM

Примеры

RSM и категориальные переменные

DoE для RSM

Адаптивный Дизайн

Техника Адаптивного DoE

Регрессия гауссовского процесса

Выход суррогатной модели

Критерий сэмплирования

Integrated MSE Gain-Maximum Variance

Поверхность критерия максимальной дисперсии

MaxMin

Пример

Адаптивный Дизайн

Определение

Адаптивный DoE — это метод, который итеративно добавляет точки к обучающей выборке, минимизируя ошибку модели.

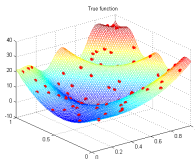


Рис.: Неизвестная функция

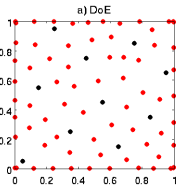


Рис.: Изначальные и
добавленные точки

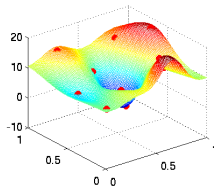


Рис.: Изначальная модель

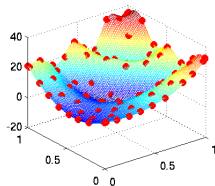


Рис.: Конечная модель

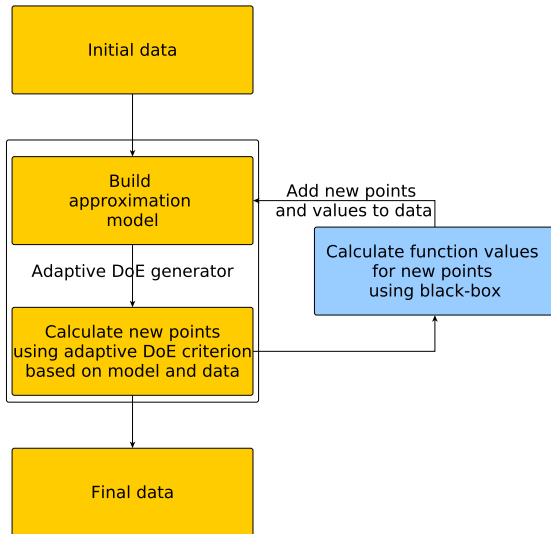
Адаптивный Дизайн

Адаптивный Дизайн эксперимента позволяет контролировать процесс суррогатного моделирования при помощи выборочного семплирования, улучшая тем самым качество аппроксимации.

Техника Адаптивного DoE:

- › Сгенерировать начальный дизайн одной из техник space-filling;
- › Построить аппроксимацию на полученном множестве;
- › Обновить обучающую выборку, итеративно добавляя точки, с целью обеспечить наибольшее возможное улучшение аппроксимации.

Процесс Адаптивного DoE



Адаптивный DoE

Причины такого подхода:

1. Аппроксимация содержит в себе информацию о поведении функции.



Аппроксимация может подсказать, как выбрать новый элемент, чтобы увеличить качество больше всего.

2. Адаптивное обновление обучающей выборки даёт гибкие возможности для изменения числа точек.



Если желаемое качество аппроксимации достигнуто, процесс DoE может быть закончен.

Регрессия гауссовского процесса

Постановка задачи

Заполнение пространства

Random sampling

Семплирование латинскими гиперкубами

Полный факторный дизайн

Halton sequence

Свойства

Поверхность отклика

Оптимальные Дизайны RSM

Примеры

RSM и категориальные переменные

DoE для RSM

Адаптивный Дизайн

Техника Адаптивного DoE

Регрессия гауссовского процесса

Выход суррогатной модели

Критерий сэмплирования

Integrated MSE Gain-Maximum Variance

Поверхность критерия максимальной дисперсии

MaxMin

Пример

Регрессия гауссовского процесса

- › Предположим, что $f(\mathbf{x})$ является реализацией гауссовского процесса (GP) со средним $\mu(\mathbf{x})$ и ковариационной функцией $k(\mathbf{x}, \mathbf{x}')$.
- › GP это стохастический процесс, все конечные сечения которого $f(\mathbf{x}_1), f(\mathbf{x}_2), \dots, f(\mathbf{x}_l)$ — гауссовские.
- › Гауссовский процесс полностью определяется своим средним и ковариационной функцией.
- › Предположим, что $\mu(\mathbf{x}) \equiv 0$ и ковариационная функция $k(\mathbf{x}, \mathbf{x}')$ известна.

Регрессия гауссовского процесса

Предположим, данные зашумлены: $y(\mathbf{x}) = f(\mathbf{x}) + \varepsilon(\mathbf{x})$, где $\varepsilon(\mathbf{x})$ — гауссовский белый шум, то есть $\forall \mathbf{x}_1, \dots, \mathbf{x}_l \in \mathbb{X} \subseteq \mathbb{R}^{d_{in}} : \varepsilon(\mathbf{x}_1), \dots, \varepsilon(\mathbf{x}_l)$ удовлетворяют свойствам:

- › Независимо и одинаково распределены.
- › Распределение нормально.
- › С нулевым средним и дисперсией $\tilde{\sigma}^2$.

Регрессия гауссовского процесса

$D = (X, \mathbf{y}) = \{(\mathbf{x}_i, y_i = f(\mathbf{x}_i) + \varepsilon(\mathbf{x}_i))\}_{i=1}^N$ — обучающее множество.

Здесь и далее для простоты возьмём $d_{out} = 1$, то есть $Y = \mathbf{y} \in \mathbb{R}^1$.
Апостериорное распределение $f(\mathbf{x})$:

$$\text{Law}(f(\mathbf{x})|D) = \mathcal{N}(\hat{f}(\mathbf{x}), \hat{\sigma}^2(\mathbf{x})).$$

Выход суррогатной модели

- Апостериорное среднее, используемое как выход суррогатной модели $\hat{f}(\mathbf{x})$, имеет вид:

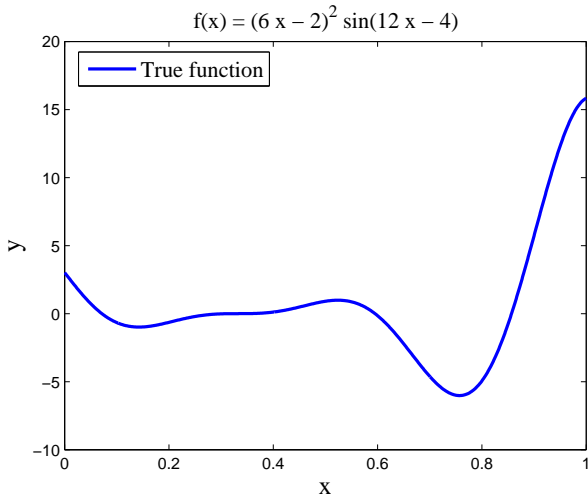
$$\hat{f}(\mathbf{x}) = \mathbf{k}(\mathbf{x})^T (\mathbf{K} + \tilde{\sigma}^2 \mathbf{I})^{-1} \mathbf{y},$$

где $\mathbf{K} = \{k(\mathbf{x}_i, \mathbf{x}_j)\}_{i,j=1}^N$, $\mathbf{k}(\mathbf{x}) = \{k(\mathbf{x}, \mathbf{x}_i)\}_{i=1}^N$.

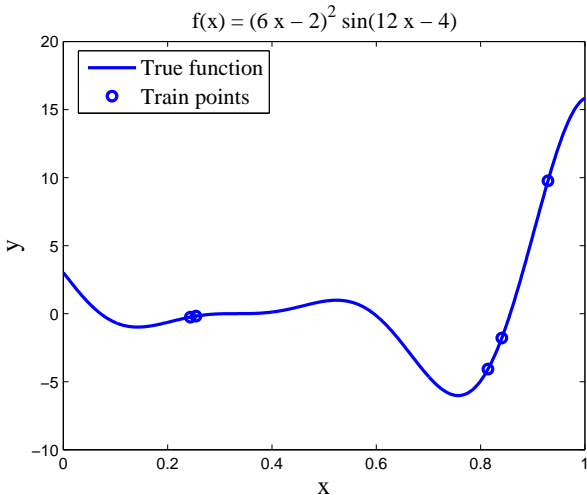
- Апостериорная ковариация, используемая как оценка точности, имеет вид:

$$\hat{\sigma}^2(\mathbf{x}) = k(\mathbf{x}, \mathbf{x}) - \mathbf{k}(\mathbf{x})^T (\mathbf{K} + \tilde{\sigma}^2 \mathbf{I})^{-1} \mathbf{k}(\mathbf{x}).$$

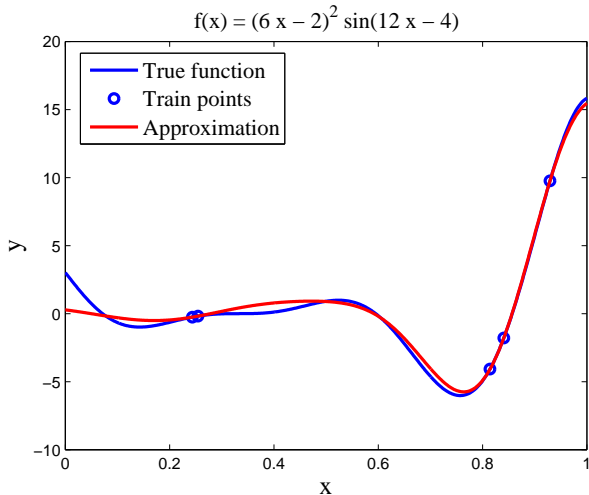
Выход суррогатной модели



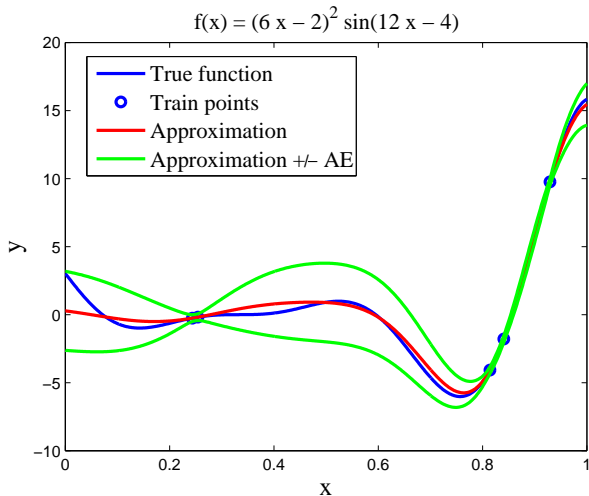
Выход суррогатной модели



Выход суррогатной модели



Выход суррогатной модели



Ковариационная функция

В реальных задачах ковариационная функция обычно неизвестна, и нужно её оценить по данным.

Для оценки ковариации используются два семейства функций:

Определение

- › Взвешенное Евклидово расстояние:

$$k(\mathbf{x}, \mathbf{x}') = \sigma^2 \exp \left(- \sum_{i=1}^{d_{in}} \theta_i^2 (x_i - x'_i)^s \right), s \in [1, 2],$$

- › Расстояние Махаланобиса:

$$k(\mathbf{x}, \mathbf{x}') = \sigma^2 \exp \left(- (\mathbf{x} - \mathbf{x}')^T \Theta (\mathbf{x} - \mathbf{x}') \right),$$

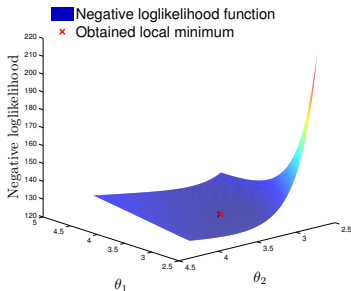
где Θ это $d_{in} \times d_{in}$ матрица.

Оценивание параметров

Параметры ковариационной функции $k(\mathbf{x}, \mathbf{x}')$ оцениваются при помощи максимизации правдоподобия:

$$-\frac{1}{2} \ln |\mathbf{K}| - \frac{\mathbf{y}^T \mathbf{K}^{-1} \mathbf{y}}{2} \rightarrow \max_{\mathbf{a}},$$

где $\mathbf{a} = (\theta_1, \dots, \theta_{d_{in}}; \sigma)$ или $\mathbf{a} = (\Theta; \sigma)$ в зависимости от типа используемой ковариационной функции.



Критерий сэмплирования

Постановка задачи

Заполнение пространства

Random sampling

Семплирование латинскими гиперкубами

Полный факторный дизайн

Halton sequence

Свойства

Поверхность отклика

Оптимальные Дизайны RSM

Примеры

RSM и категориальные переменные

DoE для RSM

Адаптивный Дизайн

Техника Адаптивного DoE

Регрессия гауссовского процесса

Выход суррогатной модели

Критерий сэмплирования

Integrated MSE Gain-Maximum Variance

Поверхность критерия максимальной дисперсии

MaxMin

Пример

Критерий сэмплирования

Процедура выбора новой точки для дизайна может быть сформулирована как задача максимизации некоторого критерия:

$$\mathbf{x}_{new} = \arg \max_{\mathbf{x} \in \mathbb{X}} \mathcal{I}(\mathbf{x}|D, \hat{f}, \hat{\sigma}^2),$$

где $\mathcal{I}(\mathbf{x}|D, \hat{f}, \hat{\sigma}^2)$ - это критерий выбора точек, основанный на обучающем множестве D , текущей аппроксимации \hat{f} и текущей ошибке $\hat{\sigma}^2$.

Критерий максимальной дисперсии

Определение

Критерий максимальной дисперсии это

$$\mathcal{I}_{MV}(\mathbf{x}) = \hat{\sigma}^2(\mathbf{x}|X),$$

где $\hat{\sigma}^2(\mathbf{x}|X)$ - ошибка обученной на $D = (X, \mathbf{y})$ аппроксимации в точке \mathbf{x} .

Преимущества

- › Легко вычислим.
- › Включает информацию о поведении функции.

Недостатки

- › Учитывается только локальное поведение.
- › Склонна давать точки ближе к границе X .

Минимизация средней ошибки предсказания

Определение

Оптимальный критерий для минимизация ожидаемой среднего квадрата ошибки аппроксимации на следующей итерации:

$$\mathcal{I}_{\rho_2}(\mathbf{x}) = \frac{1}{|\mathbb{X}|} \int_{\mathbb{X}} (\hat{\sigma}^2(\mathbf{v}|X) - \hat{\sigma}^2(\mathbf{v}|X \cup \mathbf{x})) d\mathbf{v},$$

где

- › $\hat{\sigma}^2(\mathbf{v}|X)$ - ошибка предсказания аппроксимации построенной на множестве $D = (X, \mathbf{y})$.
- › $\hat{\sigma}^2(\mathbf{v}|X \cup \mathbf{x})$ - ошибка предсказания аппроксимации построенной на множестве $D^{ext} = D \cup (\mathbf{x}, y(\mathbf{x}))$.

Integrated MSE Gain-Maximum Variance

Определение

Integrated MSE Gain-Maximum Variance критерий — это

$$\mathcal{I}_{IGMV}(\mathbf{x}) = \mathcal{I}_{\rho_2}(\mathbf{x}) * \mathcal{I}_{MV}(\mathbf{x}).$$

Преимущества

- › Включает информацию о поведении функции на всём пространстве дизайна.

Недостатки

- › Относительно сложная вычислимость.

Взгляд на критерий ImseGain-MaxVar изнутри

Некоторые матричные вычисления позволяют выразить этот критерий следующим образом:

$$\mathcal{I}_{\rho_2}(\mathbf{x}) = \frac{1}{|\mathbb{X}|} \int_{\mathbb{X}} \frac{\hat{K}^2(\mathbf{x}, \mathbf{v})}{\hat{\sigma}^2(\mathbf{x}|X)} d\mathbf{v},$$

где $\hat{K}(\mathbf{x}, \mathbf{v}) = K(\mathbf{x}, \mathbf{v}) - \mathbf{k}(\mathbf{x})K^{-1}\mathbf{k}(\mathbf{v})^T$ — апостериорная ковариация между значениями гауссовского процесса в точках \mathbf{x} и \mathbf{v} .

Критерий может быть вычислительно нестабильным, если $\hat{\sigma}^2(\mathbf{x}|X)$ мало.

Денис Деркач



Взгляд на критерий ImseGain-MaxVar изнутри

Можно решить эту проблему, совместив критерии $\mathcal{I}_{\rho_2}(\mathbf{x})$ и $\mathcal{I}_{MV}(\mathbf{x})$ в один с меньшим числом недостатков:

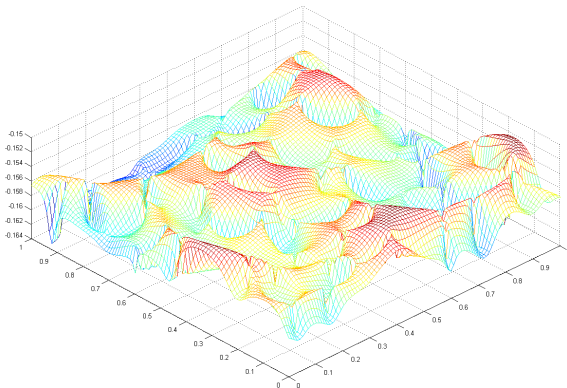
$$\mathcal{I}_{IGMV}(\mathbf{x}) = \mathcal{I}_{\rho_2}(\mathbf{x}) * \mathcal{I}_{MV}(\mathbf{x}) = \frac{1}{|\mathbb{X}|} \int_{\mathbb{X}} \hat{K}^2(\mathbf{x}, \mathbf{v}) d\mathbf{v}.$$

Преимущества

- › Нет проблемных членов в знаменателе.
- › По-прежнему ведёт к аппроксимации по всему пространству дизайна.

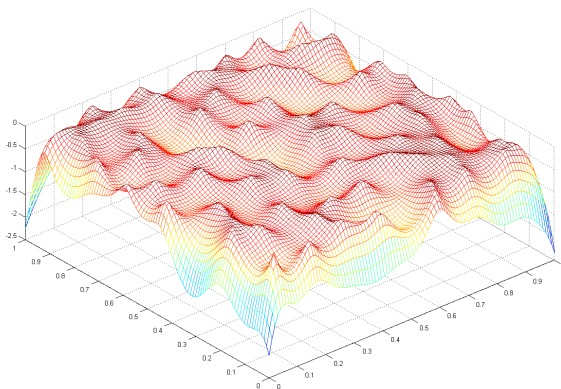
\mathcal{I}_{ρ_2} поверхность

- › Высокая мультимодальность.
- › Узкие и не робастные локальные минимумы.



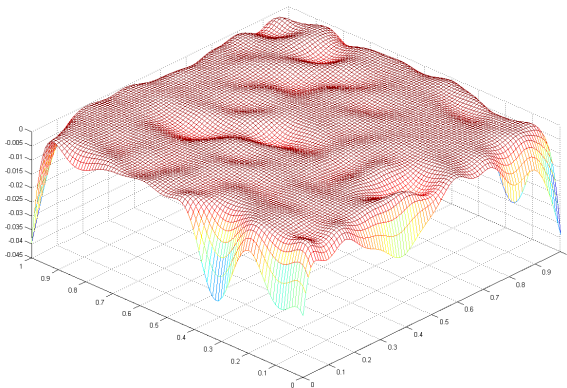
Поверхность критерия максимальной дисперсии

- › Высокая мультимодальность.
- › Поведение более регулярное.



Поверхность ImseGain-MaxVar

- › Меньше локальных минимумов.
- › Регулярность поведения.



MaxMin

Определение

MaxMin критерий это:

$$\mathcal{I}_{MM}(\mathbf{x}) = \min_{\mathbf{v} \in X} d^2(\mathbf{v}, \mathbf{x}),$$

где $d(\mathbf{v}, \mathbf{x})$ — это Евклидово расстояние между точками \mathbf{v} и \mathbf{x} :
 $(d^2(\mathbf{v}, \mathbf{x}) = \sum_{i=1}^{d_{in}} (v_i - x_i)^2).$

Преимущества

- › Очень быстро вычисляется.
- › Не опирается на тип аппроксимируемой модели.
- › Берёт точки равномерно по пространству дизайна.

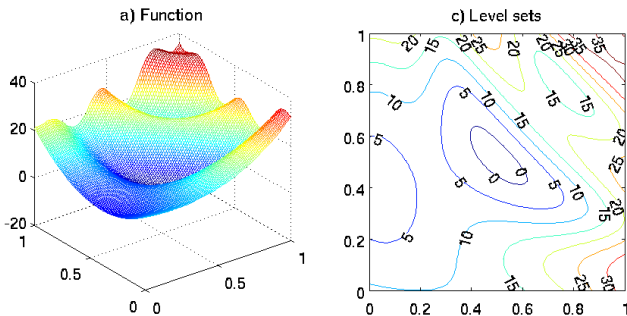
Недостатки

Денис Деркач

Пример: Неизвестная функция

Рассмотрим задачу аппроксимации следующей функции:

$$y = 2 + 0.25(x_2 - 5x_1^2)^2 + (1 - 5x_1)^2 + 2(2 - 5x_2)^2 + 7\sin(2.5x_1)$$



Пример: начальная аппроксимация

Изначальный экспериментальный дизайн содержит 10 точек.
Строим аппроксимацию регрессией гауссовского процесса.

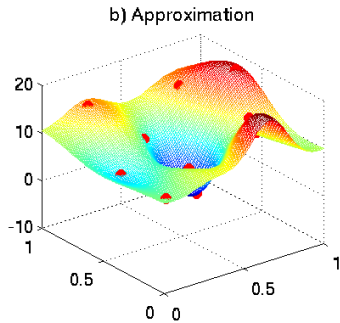
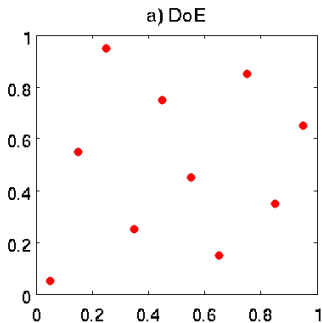
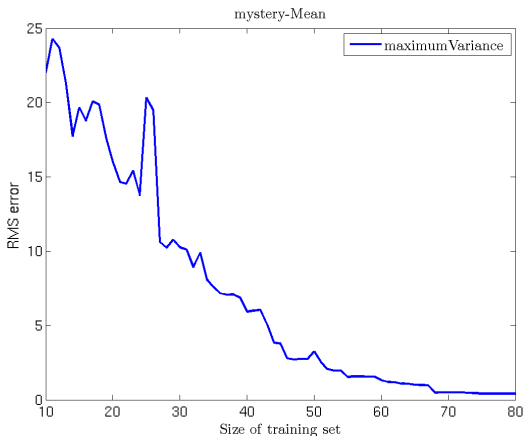


Рис.: Начальная обучающая выборка и аппроксимация

Пример: адаптивный процесс DoE

Добавлено 70 точек при помощи критерия Максимальной Дисперсии.

Изменение ошибки показано на рисунке ниже.



Пример: финальная аппроксимация

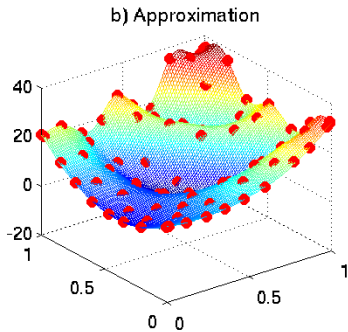
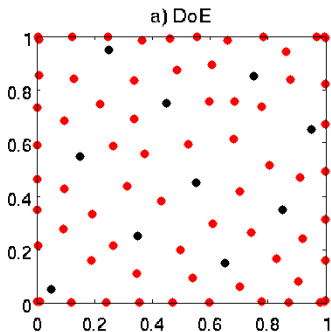


Рис.: Финальная обучающая выборка и аппроксимация

Пример: результаты

Адаптивный процесс DoE даёт очевидное снижение ошибки, и итоговая модель очень точная.

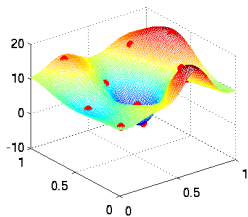


Рис.: Изначальная модель

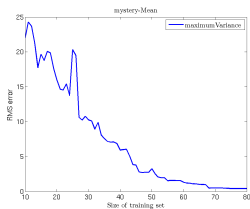


Рис.: Зависимость ошибки от размера обучающей выборки

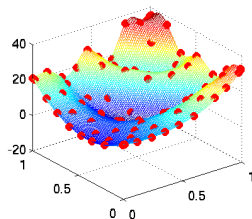


Рис.: Финальная модель