## 72. Полный факторный эксперимент (ПФЭ). Предпосылки

активный

Любое экспериментальное исследование содержит 3 этапа:

1. Этап постановки задачи

2. Этап планирования и проведения эксперимента

3. Анализ и интерпретация результатов.

Пассивный эксперимент – результат исследований, при которых исследователь лишен возможности управлять факторами, а вынужден только фиксировать сложившуюся ситуацию.

Модель-это, упрощённая система, отражающая отдельные стороны, свойства изучаемого объекта. Модели бывают физическими и математическими.

Физическая модель представляет собой уменьшенную копию реального объекта (очень полезная вещь).

Математическая модель является методом качественного и количественного описания объектов или процессов. При этом реальный объект, процесс или явление схематизируются, упрощаются и описываются уравнениями.

Мат. модель – это описание исследуемого объекта, которое нельзя получить в виде точной формулы функции, справедливой во всем диапазоне существования аргументов. Модель может быть лишь приближенной и на небольшом участке в окрестностях выбранной базовой точке.



Эффективность экспериментального исследо-вания будет значительно выше, если эксперимент будет математически спланирован. Следует определить минимальное количество точек N – факторного пространства и их координаты в которых следует поставить эксперимент, чтобы получить максимальное количество информации. Система из N – нормальных уравнений может быть представлена в виде матрицы, чтобы матрица имела одно единственное решение матрица должна быть невырожденной, то есть векторы столбцы линейно не зависимы, а чтоб коэффициент уравнения регрессии были независимы друг от друга на нее можно наложить дополнительное условие ортогональности. Этим всем условиям отвечает матрица планирования ПФЭ.

 Условия построения матрицы ПФЭ:

Симметричность относительно центра.(1)

Условие нормировки (2)

Условие ортогональности (3)

Рототабельность, т.е. точки в матрице планирования подбираются так, чтоб точность предсказания значения параметров оптимизации была одинакова на равных расстояниях от центра эксперимента.

Математическое описание объекта в окрестностях базового режима получается путем варьирования каждого фактора Х\* на двух уровнях отличающихся друг от друга на величину шага варьирования ΔXi.



Шаг варьирования выбирается таким, чтоб приращение выходного параметра ΔY к базовому значению у\* было различимо на фоне «шума» при небольшом числе параллельных опытов. Первоначальный шаг варьирования выбирается ΔXi ≈ 0,15Xi\*. В случае получения неадекватной модели ШВ уменьшается, и опять проводится эксперимент. ПФЭ – называется эксперимент реализующий все возможные комбинации уровней независимых факторов, каждый из которых принудительно варьируется на двух уровнях. Xi ± ΔX, N=2n (все возможные комбинации факторов), т.к. все факторы имеют разную размерность и даже разные порядки, то лучше перейти к от абсолютных единиц измерения к относительным, то есть xi = (Xi – Xi\*)/ΔXi. Таким образом, мы переносим эксперимент в центр базового режима и условные единицы измерения по каждому фактору становится ШВ ΔXi. Такое преобразование дает возможность построить ортогональную матрицу планирования эксперимента и значительно облегчить расчеты, т.к. в этом случае Xiв и Xiн по каждому фактору становится относительными хiв = 1, хiн = -1.

При реализации плана необходимо:



1. В каждой точке этого плана необходимо получить выборку.

2. Надо воспользоваться рандомизацией, т.к. существует влияние неучтённых факторов.

Для каждой точки: ; . Для сравнения дисперсий можно воспользоваться критерием Кохрена или Бартлетта. Дисперсию эксперимента нужно считать по формуле: а)  — для Кохрена; б) Средневзвешенная дисперсия по Бартлетту.

Если дисперсии неразличимы, то, стало быть, есть грубый промах: 1. Выбросить грубый промах; 2. На его место становится среднее арифметическое.

Теперь находим коэффициенты модели, проверяем их на значимость, проверяем адекватность модели.

;; т.к. 

введём фиктивную переменную , тогда 

Получаем исходные коэффициенты модели, т.е. нашли не истинные коэффициенты, а их оценки. Сравним их с НУЛЁМ – проверим на значимость.

.

Если неравенство справедливо, то данный коэффициент признаём незначимым, а если нет, то оставляем его в нашей модели. Количество значимых коэффициентов модели обозначим за d.

 — Это наша модель.  — Предсказанное значение выходной величины. Мерой отклонения экспериментальных данных от предсказанных служит дисперсия адекватности.

, где N>d. Если N = d, то мы не можем оценить адекватность нашей модели.

Для оценки адекватности модели пользуются критерием Фишера: . Если неравенство выполняется, то модель адекватна, если нет, то неадекватна.

**Предпосылки полного факторного эксперимента**

Математическая модель является методом качественного и количественного описания объектов или процессов. При этом реальный объект, процесс или явление схематизируются, упрощаются и описываются уравнениями.



Предпосылки факторного анализа

1. Результаты наблюдений  выходной величины Y в N точках факторного пространства, представляют собой независимые, нормально распределённые случайные величины.

2. Выборочные дисперсии опытов  однородны, т.е. статистически неразличимы. Это требование означает независимость выборочной дисперсии от места факторного пространства, в котором проводится эксперимент. (Свойство – Ротатабельность.)

3. независимые переменные  измеряются с ошибкой много меньшей, чем величина возможного отклонения выходного параметра Y от действия неучтённых факторов.

(=0 при идеальном случае.)

 d- число членов модели.

Для получения единственного решения необходимо чтобы матрица была невырожденной (определитель ? 0).

Величины коэффициентов регрессии не зависели от числа членов матрицы, на неё нужно наложить опр. условие ортогональности вектор-столбцов.

## 73. Модифицированный метод случайного баланса (ММСБ). Идеология метода

пассивный

Одним из наиболее удобных методов моделирования исследуемого технологического процесса по пассивным данным является ММСБ. Целью эксперимента в МСБ является упорядочение факторов по степени их влияния на целевую функцию и оценка модели. Под результатом пассивного эксперимента будем понимать таблицу, каждая строка которой представляет собой числовое значение целевой функции при некоторых условиях Матрицу планирования переделываем путем перестановки строк таким образом, чтобы на первое место стала строка с минимальным y и т.д. На последнее место и т.д. На последнее место  (ранжировка строк по y). Тогда, взяв первые и последние несколько строк матрицы, путем просмотра столбцов факторов и их парных взаимодействий можем на глаз определить имеют ли мах и min значения у одинаковые знаки в столбцах x , если знаки в столбцах одинаковы, то такой фактор (или парное сочетание) не рассматривается. Если знаки разные, то такие факторы надо рассмотреть (с помощью формул), потому что среди них могут быть значимые.



Такая модель не может быть проверена на адекватность.

Надо перейти от насыщенного к обыкновенному плану.

На II этапе расчетов после выделения значимых факторов и их парных взаимодействий, необходимо перестроить исходную матрицу планирования, оставив в ней только столбцы значимых факторов. При этом некоторые строчки новой матрицы будут совпадать по комбинации знаков x . Такие строки необходимо совместить, а из величин y получить выборку малого объема, которая даст нам дисперсию строки, а через нее по критерию Бартлетта

С помощью свернутой таблицы - новой матрицы планирования заново определяем оценки коэффициентов bk и заново определяем их значимость.

Если в строчных выборках нет грубых промахов, то пересчитывать коэффициенты не надо. В большинстве случаев, в строчных выборках появляются грубые промахи.

 - объем выборки в g строк



Среди активных экспериментов особое место занимают МСБ и ММСБ, которые позволяют при большом числе факторов и при малом числе опытов получить разделение всех факторов на значимые и незначимые. Значимые факторы могут быть оценены по степени влияния на выходную величину через эмпирические коэффициенты регрессии bk которые войдут в регрессионную модель. Модель может быть оценена на адекватность исходным данным, и тем самым, работа будет завершена.

 Оба метода построены по одному принципу - сверхнасыщенному плану, но отличаются методами обработки результатов эксперимента.

## 74. Планирование второго порядка. Идеология методов

Полный факторный экстримент типа N=2n дает неполную квадратичную модель, которая хорошо описывает исследуемый объект (процесс) в области, далекой от экстремума, но которая становится непригодной по мере приближения базовой точки к экстремуму. В этой экстремальной области в модели необходимо наличие членов при квадратах факторов , т.е. оценки коэффициентов bii.Если иметь полную квадратичную модель

(19.1)

то для нахождения оптимальной точки достаточно приравнять нулю значения компонентов градиента, вычисленных по формуле:



Решая систему линейных уравнений, получаем координаты оптимальной точки в относительных единицах, так как начало координат при таком планировании находится в базовой точке. После перехода от нормированных величин xi к физическим Xi можно определить точность предсказания точки экстремума по квадрату расстояния от найденной точки до действительной точки экстремума 



Практикуется также оценка по разности значений выходной величины в точках действительного и предсказанного экстремума.



а также относительная оценка точности

Эти оценки используются при исследовании методов планирования на модели, когда известна точка действительного оптимума. В реальных условиях точность предсказания оценивается по величине дисперсии σ2{Y}.Попытка получить полную квадратичную модель методами ПФЭ типа N=2n нереальна. Задача может решена так: для получения раздельных оценок b0 и bii к ПФЭ типа 2n добавляется центральная точка с координатами /0,0,…,0/ и так называемые звездные точки с координатами /0,0,…,±α ,…,0/, лежащие на сфере диаметра 2α. Эти дополнительные точки для трехфакторной задачи приведены в табл.

Относительные координаты звездных и центральных точек.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| q | x0 | x1 | x2 | x3 |
| 9101112131415 | +1+1+1+1+1+1+1 | −α+α00000 | 00−α+α000 | 0000−α+α0 |

Особое внимание следует обратить на центральную точку, роль которой очень велика. Проиллюстрируем это на следующем примере. Пусть в трех разных ситуациях отклики, (см.рис.), равны в вершинах квадратов, но отличаются в центральной точке.

Результаты ПФЭ с измерением в центральной точке

а) bii<0 б) bii=0 в) bii>0

Тогда линии равного уровня (пунктир) дадут разные картины, разными будут коэффициенты bii .

В теории планирования второго порядка в зависимости от критерия оптимальности плана различают ортогональное центральное композиционное планирование и рототабельное центральное композиционное планирование.

В ортогональном центральном композиционном планировании (ОЦКП) критерием оптимальности плана является ортогональность столбцов матрицы планирования. В силу этой ортогональности все оценки коэффициентов регрессии определяются независимо друг от друга.

Критерий оптимальности плана ставит задачу построения матрицы планирования с ортогональными вектор-столбцами. Для обеспечения этого необходимо преобразовать модель (19.1) следующим образом:

 где ; N – общее число точек в плане. Где



В этом случае выполняется условие ортогональности.

Ротатабельное центральное композиционное планирование(РЦКП)

Критерием оптимальности в ротатабельном плане является условие , то есть, требование симметричности информационных контуров. Для планов второго порядка его можно достичь, если все нечетные моменты, вплоть до 4-го порядка будут равны нулю, а для четных моментов будет выполняться соотношение:  где  и  выбираются из условия невырожденности матрицы нормальных уравнений (Х т Х) и удовлетворяют неравенству:



Ротатабельные планы оптимальны также и в том смысле, что они позволяют минимизировать систематические ошибки, связанные с неадекватностью представления результатов исследования полиномом второго порядка.

РЦКП строится аналогично ортогональному плану. В качестве ядра используется ПФЭ (2n). К нему добавляются центральные точки и звездные точки. Число опытов N0 в центре плана выбирается из следующих соображений: выдвигается требование, чтобы информация о значении выходной переменной оставалась неизменной (или почти неизменной) для точек внутри сферы единичного радиуса с центром в центре плана. Иными словами, требуется, чтобы информационный профиль ротатабельного плана мало изменялся при значениях радиуса сферы от 0 до 1. Оказывается, что такие планы можно получить, меняя число точек в центре ротатабельного плана. Пример плана РЦКП в общем виде для трех факторов представлен в табл. 6.9.

Расчет плеча звездных точек можно осуществить по формулам:

Расчет плеча звездных точек можно осуществить по формулам:

 

Значения α и других вспомогательных величин для РЦКП представлены обычно в таблицах.

## 75. Метод наименьших квадратов с предварительной ортогонализацией факторов (МНКО). Идеология метода

пассивный

Один из самых старых и разработанных методов моделирования по пассивным данным – метод наименьших квадратов, который базируется на подборе такого уравнения регрессии, чтобы сумма квадратов разности между уравнением и экспериментальными данными была наименьшей из всех возможных.

Задача заключается в подборе таких значений x, чтобы значения этих функций были максимально близки к некоторым значениям . По существу речь идет о «решении» переопределенной системы уравнений ,  в указанном смысле максимальной близости левой и правой частей системы. Сущность МНК заключается в выборе в качестве «меры близости» суммы квадратов отклонений левых и правых частей  . Таким образом, сущность МНК может быть выражена следующим образом: .

 Для произвольной системы факторов задача нахождения обратной матрицы является довольно громоздкой даже для ЭВМ, причем трудоемкость стремительно возрастает с увеличением числа факторов. Одновременно существует еще одна проблема – при признании какого–либо из найденных коэффициентов bk незначимым следует, исключив фактор Xk, всю вычислительную процедуру проделать заново с самого начала. Проблема существенного упрощения процедуры определения коэффициентов регрессии и отсеивания незначимых факторов может быть решена путем предварительной ортогонализации факторов.

1. Анализ особенностей МНКО как в теоретическом плане, так и в плане практического применения позволяет обратить внимание на следующее:

2. В условиях пассивного эксперимента оценки коэффициентов bk в отличие от Ak являются смешанными. Однако по сравнению с МНК предложенный метод позволяет точнее оценить независимый вклад каждого эффекта в соответствующий коэффициент bk. Это обстоятельство обуславливает более высокую чувствительность МНКО по сравнению с МНК, которая тем выше, чем больше количество исследуемых факторов, причем в этот список могут входить как сильно–, так и слабодействующие факторы.

3. Эффективность метода зависит от порядка следования факторов (эффектов) друг за другом при расчете коэффициентов модели. В случае расположения их в порядке убывания значимости эффективность метода возрастает. Поэтому целесообразно перед применением МНКО предварительно расположить исследуемые факторы (эффекты) в порядке убывния значимости (степени влияния) по отношению к целевой функции. Для этого можно рекомендовать воспользоваться предварительной моделью, полученной с помощью ММСБ или какого–либо другого метода.

4. Поскольку оценки МНКО получены на основе тех же предпосылок, что и оценки МНКК, то они обладают одинаковыми свойствами, то есть несмещенности, эффективности и состоятельности. Однако исследования последних лет отрицают свойство состоятельности за статистическими оценками, с чем, по-видимому, следует согласиться (см. подраздел 5.6).

5. Здесь же следует подчеркнуть, что поскольку МНКК выведен для закона нормального распределения факторов и выходной величины, то, по-видимому, эти же требования в значительной степени относятся и к МНКО.

6. Существенной особенностью и преимуществом МНКО является то обстоятельство, что в силу перехода данных в заведомо ортогональную систему координат можно получать оценки коэффициентов и для коррелированных факторов и для квадратных членов.

7. Другой существенной особенностью и преимуществом МНКО является то, что для получения модели не требуется слишком длинной таблицы исходных данных как в ММСБП, лишь бы координаты точек факторного пространства были бы достаточно далеки друг от друга.

8. Модель МНКО является обычным алгебраическим выражением, коэффициенты ее представляют собой смешанные оценки и не являются, как в ММСБ, весами соответствующих факторов.

На графиках все прекрасно видно. Красная линия – это найденная прямая y = 0.165x+2.184, синяя линия – это , розовые точки – это исходные данные.

Для чего это нужно, к чему все эти аппроксимации?

**В МНК если какой-то из найденных коэффициентов незначим, следует исключить фактор и всю процедуру проделать заново с самого начала.**

**В МНКО всё свуодится к тому, что связь между выходной величиной y и факторами x ищут в виде полинома, включающего эффекты факторов и их взаимодействия. Для этого находят вспомогательный полином и провести замену переменных.**

**Отсеивание в МНКО проводится по критерию Стьюдента.**

**Проверка адекватности по критерию Фишера (равенство дисперсий)**

## 76. Методы оптимизации. Основные понятия. Градиентные и неградиентные методы

Задача оптимизации.

Мы должны найти некую функцию ; 

Если мы знаем полную функцию то ; i=1,2,...,n. По каждой координате частный экстремум. Это можно записать так: grad y()=0.

grad y()= - второй способ нахождения экстремума когда не знаем функцию.

Частные производные это коэф. модели b1,b2,...,bn необходимо чтобы линейные коэф. пришли в нули.

Yнаблюдаемое = Yистинное + ξ.

При втором методе мы должны быть очень осторожны из-за ξ. Задача резко усложняется при наличии дрейфа объекта. Будем разбирать только стационарные модели.

Смысл всех методов заключается в том, что мы получаем функцию отклика в какой-то определенной точке, при этом получаем некое направление на экстремум. Через несколько шагов вновь ставим эксперимент, чтобы скорректировать направление.

Метод крутого восхождения (Бокса-Уилсона).

Если модель сделана в виде полинома второго порядка, то найти оптимальные точки очень просто: т.к. сама модель второго порядка символизирует, что мы находимся в области экстремума. Линейная модель свидетельствует о том, что:

1) мы находимся далеко от вершины холма

2) в силу малости известной области требуется выработка некоторых правил, согласно которым будем двигаться к экстремуму.

Существует несколько путей нахождения:

1) метод Гауса-Зейделя

2) метод градиента

3) метод случайного поиска

4) метод крутого восхождения

5) оптимизация в условиях ограничений

**Метод крутого восхождения**

При использовании алгоритма крутого восхождения пошаговое движение из точки вектора  совершается в направлении наискорейшего возрастания функции, т.е. по градиенту в этой точке. Однако, в отличии от градиентного метода, корректировка направления производится не после каждого следующего шага, а только по достижению частного экстремума.

Важной особенностью метода является также регулярный статистический анализ результатов экспериментов по мере продвижения к экстремуму.

Пусть у нас k факторов, которые независимы друг от друга.

Алгоритм.

1. В центре  проводим ПФЭ (ДФЭ) с целью получения градиента. При этом результаты эксперимента подвергаются статистическому анализу, который заключается в следующем:

а) проверка воспроизводимости эксперимента (грубые промахи)

б) проверка значимости коэффициентов линейной модели

в) проверка адекватности линейной модели

2. Вычисляются произведения  среди них находится max{} = (b - величина относительная,  - абсолютная).

3. Для базового фактора выбирается шаг крутого восхождения (только для базового). Дальше будем обозначать его .

4. Определяются размеры



5. Проводятся т.н. «мысленные опыты», которые заключаются в вычислении предсказанных значений целевой функции в определенных точках факторного пространства. . Для этого независимые переменные линейной модели изменяются с учетом  таким образом, чтобы изображающая точка совершала движение в направлении вектора grad y() полученного в пункте 1 занимая последовательно положения .

6. «Мысленные опыты» продолжаются до тех пор, пока выполняется неравенство 

7. Некоторые из «мысленных опытов» (обычно через каждые 2-3 «мысленных шага») реализуются в виде эксперимента один раз. Для проверки соответствия «мысленных опытов» и реальности.

8. Точка , где в реальном опыте получено максимальное значение целевой функции, принимается за новую начальную точку нового цикла крутого восхождения

9. Поскольку каждый цикл крутого восхождения приближает изображающую точку к области экстремума, где крутизна поверхности отклика меньше, то для каждого следующего цикла величина  может выбираться меньше предыдущей.

10. Поиск прекращается, когда градиент равен нулю, т.е. все bi оказались незначительными.

Замечания

1)Наиболее эффективно применять МКВ для симметричных функций.

2)Направление движения идет по градиенту и определяется единственным способом из центра плана.

**Метод Гаусса - Зайделя.**

При оптимизации по методу Г-З последовательное приближение к экстремуму осуществляется путем поочередного варьирования каждого параметра до достижения частного экстремума выходной величины. Другими словами изображающая точка  перемещается поочередно вдоль каждой из координатных осей факторного пространства (хi(i=1,…,n)). Причем переход к новой i+1 оси совершается по достижению частного экстремума по i-ой оси.  достигнув частный экстремум по последней n-ой оси вновь переходим к первой и цикл повторяется пока одновременно по всем осям не будет, достигнут экстремум.



Особенностью данного метода яв-ся возможность длительной стабилизации.

Когда доходим по оси Х1 до того, что У начал уменьшаться, останавливаемся и отходим назад до max У, и начинаем изменять Х2. Когда начал уменьшаться снова идем по Х1. Когда доходим до точки А и видим, что изменение любого фактора уменьшает вых. величину (можно уменьшить шаг), то в этой точке ставим эксперимент второго порядка.

Алгоритм.

определяется начальная точка  направление движения к оптимуму.

Задается шаг варьирования ∆Хi по каждой координатной оси.

Для выяснения направления движения в первом рабочем цикле, т.е. вдоль оси Х1 необходимо сделать два пробных эксперимента:  

 если “+” то вдоль оси, если “-” то против.

Осуществляется первый цикл рабочего движения: ; ;...; .

После каждого рабочего шага в пунктах  и  делается одно измерение , ,...,. 

Первый цикл движения прекращается при достижении  эта величина определяется:  как только у<у на предыдущем шаге – стоп.

Переходим на новую ось в точке  , снова делаем пробные шаги по второй оси, снова определяем величину Ψ2 : 

; ;...; .

Смотри 7, процедура продолжается до тех пор, пока движение по всем осям не приводит к ухудшению у.

Возможные решения: а. Уменьшить шаг; б. Провести эксперимент второго порядка; в. Ничего не делать, если полученной точности нам хватает.

Особенности: Делаем однократное измерение, возможно, что мы не правильно измерили

1. Длительная стабилизация всех факторов, кроме одного, который измеряем.

2. Если на очередном шаге получили величину у почти не отличающуюся от полученной на предидушем шаге необходимо сделать еще один шаг.

3. Метод не есть самый короткий путь к экстремуму.

Очень осторожный и очень правильный метод, всегда дает результаты и никогда не обманывает. Это чисто сравнительная характеристика.

**Метод градиента при поиске экстремума**

При оптимизации процесса градиентным методом рабочее движение совершается в направлении наиболее быстрого возрастания выходного параметра, т.е. в направлении градиента целевой функции. При этом, также как и в методе случайного поиска, направление движение корректируется после каждого рабочего шага, следовательно каждый раз заново вычисляется вектор  по результатам специально спланированного эксперимента.









Алгоритм: Задается шаг варьирования, единый для всех циклов работ: 

Задается параметр рабочего шага:  В начальной точке  ставится эксперимент, на основании которого определяется градиент . Рекомендуется получать градиент с помощью 2n экспериментов:

 Определяем следующую точку : 

 ……………………………..

 Процедура прекращается, если 

 Характерными особенностями метода является постоянство пробного шага и рабочего шага , следовательно получаем неприятности:

Можем проскочить мимо оптимума

Если вершина плоская (пологая) – возникают трудности

 

h – номер шага

- выбирается () 

Формы поверхности:

Наиболее благоприятные случаи а,б,г

Наиболее опасный случай – д.

Для уверенных выводов, необходимо всю процедуру сделать дважды (причем из двух абсолютно различных начальных точек).

Худший случай – в (плоская вершина на большом протяжении), следовательно метод не срабатывает.