

**Московский государственный университет имени М.В.Ломоносова.**

---

**Научно-исследовательский институт ядерной физики  
имени Д.В.Скобельцына  
Кафедра нейтронографии.**

**В.Б. Злоказов**

**МАТЕМАТИЧЕСКАЯ ОБРАБОТКА  
ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ НЕЙТРОННОГО  
РАССЕЯНИЯ В ФИЗИКЕ НИЗКИХ ЭНЕРГИЙ.**

**Учебное пособие**

Москва 2007

УДК 621.315.3

Злоказов В.Б. МАТЕМАТИЧЕСКАЯ ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ НЕЙТРОННОГО РАССЕЯНИЯ В ФИЗИКЕ НИЗКИХ ЭНЕРГИЙ. Учебное пособие. - М.: 2007. - 60 с.

Учебное пособие написано на основе специального курса кафедры нейтронографии, читаемого студентам в 8 семестре. Целью курса является изучение современных математических методов анализа экспериментальных данных, освоение типового программного обеспечения для обработки данных экспериментов с конденсированными средами - гамма, нейтронных и нейтронно-дифракционных спектров.

© Злоказов В.Б., 2007

© НИИЯФ МГУ, 2007

## СОДЕРЖАНИЕ

Лекция 1. СЛУЧАЙНАЯ ВЕЛИЧИНА И РЕГРЕССИЯ.....	5
Аргумент и параметр функции. Функция распределения, ее плотность. Центральные моменты наиболее часто употребляемых распределений в физике: равномерного, нормального, Пуассона, мультибиномиального, экспоненциального, хи-квадрат. Генераторы случайных чисел.	
Лекция 2. ОЦЕНИВАНИЕ ПАРАМЕТРОВ ФУНКЦИЙ РАСПРЕДЕЛЕНИЯ.....	9
Среднее, смещение и дисперсия оценок. Несмещенные оценки. Оценки среднего и дисперсии. Средне-квадратичное отклонение. Функция правдоподобия: непрерывная, дискретная, логарифмическая. Энтропия. Оценка максимального правдоподобия (ОМП). Эффективность оценок максимального правдоподобия. ОМП для типичных функций распределения. Байесовские оценки.	
Лекция 3. ТОЧНОСТЬ ОЦЕНОК ПАРАМЕТРОВ.....	15
Неравенство Рао-Крамера. Количество информации по Фишеру. Теорема о дисперсии МП-оценок. Дисперсии МП-оценок для важнейших распределений.	
Лекция 4. ПРОВЕРКА ГИПОТЕЗ.....	17
Критерии статистического тестирования. Уровень значимости. Хи-квадрат критерий.	
Лекция 5. ОЦЕНИВАНИЕ ПАРАМЕТРОВ РЕГРЕССИЙ. ЛИНЕЙНЫЙ СЛУЧАЙ.....	19
Методы параметризации регрессий. Линейный случай. Метод наименьших квадратов (МНК). Теорема о среднем и ковариациях МНК-оценок. Теорема Гаусса-Маркова. Теорема о нормальности МНК-оценок. Теорема о распределении остаточной суммы. Проверка гипотез о регрессиях.	
Лекция 6. ОЦЕНИВАНИЕ ПАРАМЕТРОВ РЕГРЕССИЙ. НЕЛИНЕЙНЫЙ СЛУЧАЙ.....	25
Нелинейный случай. Нелинейный метод наименьших квадратов. Примеры важнейших нелинейных регрессий. Метод линеаризации. Итерационный процесс. Точность МНК-оценок, симуляция данных для оценки точности.	

Лекция 7. ОЦЕНИВАНИЕ ПАРАМЕТРОВ РЕГРЕССИЙ. РОБАСТНОЕ ОЦЕНИВАНИЕ.....	27
Альтернативные методы анализа регрессий. "Выбросы" и робастное оценивание. Метрика модулей. Задача деконволюции. Метод регуляризации Тихонова.	
Лекция 8. МЕТОДЫ НЕЛИНЕЙНОЙ МИНИМИЗАЦИИ.....	29
Геометрическая иллюстрация. Методы минимизации: Ньютона, Гаусса-Ньютона, градиентно-подобные, стохастические. Шаг, демпфер. Проблема обрыва процесса.	
Лекция 9. АПРИОРНАЯ ИНФОРМАЦИЯ И УСЛОВНАЯ МИНИМИЗАЦИЯ.....	34
Проекторы. Метод Лагранжа. Метод штрафных функций.	
Лекция 10. ФИЛЬТРАЦИЯ.....	35
Дискретное преобразование Фурье. Интерполяция, сплайны. Дифференцирование, интегрирование. Сглаживание. Подавление плавных компонент. Пико-усиливающие фильтры.	
Лекция 11. ОСНОВНЫЕ ЗАДАЧИ АНАЛИЗА ЭКСПЕРИМЕНТАЛЬНЫХ СПЕКТРОВ.....	43
Пиковый анализ. Rietveld анализ. Powder Match. Автоматическая индексация поликристалла. Синтез Фурье.	
Приложение. МАТЕМАТИЧЕСКИЙ МИНИМУМ .....	53
ЛИТЕРАТУРА.....	60

## ЛЕКЦИЯ 1. СЛУЧАЙНАЯ ВЕЛИЧИНА И РЕГРЕССИЯ.

Итак, пусть задана случайная величина  $\xi$ , которая принимает значения из  $X$  и имеет функцию распределения  $F(x)$ . Кратко мы можем сказать, что задана тройка величин  $\xi, X, F(x) = P(\xi < x)$ . Рассмотрим ряд наиболее употребительных функций распределения в физических приложениях математической статистики и посмотрим, каковы их центральные моменты.

- биномиальное - дискретный случай.

$$\xi, X = 0, 1, \quad P(\xi = 1) = p, \quad P(\xi = 0) = 1 - p.$$

Центральные моменты - среднее и дисперсия определяются легко

$$\hat{E}\xi = p, \quad \hat{V}\xi = p(1 - p) = p - p^2.$$

- мультибиномиальное - это распределение суммы независимых биномиальных одинаково распределенных случайных величин

$$\psi = \sum_{i=1}^n \xi_i = 0, 1, \dots, n$$

Функция распределения и центральные моменты имеют вид:

$$P(\psi = m) = C_n^m p^m (1 - p)^{n-m}$$

$$\hat{E}\psi = pn, \quad \hat{V}\psi = np(1 - p) = np - np^2.$$

- Пуассона - обычно распределение целочисленных сумм "редких" независимых событий

$$\xi \in [0, 1, 2, \dots, \infty)$$

Функция распределения и центральные моменты имеют вид:

$$P(\xi = m) = \frac{a^m}{m!} \exp(-a)$$

$$\hat{E}\xi = a; \quad \hat{V}\xi = a.$$

Это свойство пуассоновских случайных величин - равенство среднего и дисперсии - имеет большое значение для практики анализа экспериментальных данных.

- равномерное - распределение случайной величины, принимающей с одинаковой вероятностью значения из интервала  $[a, b]$ :  $\xi \in [a, b]$ .

Функция распределения, плотность и центральные моменты имеют вид:

$$P(\xi < t) = \begin{cases} 0 & \text{если } t < a; \\ (t - a)/(b - a) & \text{если } a \leq t \leq b; \\ 1 & \text{если } t > b. \end{cases} \quad p(t) = \frac{1}{b - a} \chi_{[a, b]}(t)$$

$$\hat{E}\xi = \frac{a + b}{2}, \quad \hat{V}\xi = \frac{(b - a)^2}{12}.$$

Преобразованием  $\xi' = \xi - a$  мы получаем случайную величину  $\xi'$ , имеющую то же самое равномерное распределение в интервале  $[0, c]$ , где  $c = b - a$ , и моменты

$$\hat{E}\xi' = \frac{c}{2}, \quad \hat{V}\xi = \frac{(c)^2}{12}.$$

- нормальное распределение  $\xi \in (-\infty, \infty)$ ,  $F = N(a, \sigma^2)$ .

Плотность функции распределения и центральные моменты имеют вид:

$$p(t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(t-a)^2}{2\sigma^2}\right) \quad \hat{E}\xi = a; \quad \hat{V}\xi = \sigma^2$$

На нормальное распределение похоже распределение Коши:

$$p(t) = \frac{1}{\pi\sigma} \frac{1}{(1 + ((t-a)/\sigma)^2)},$$

но оно обладает рядом экзотических свойств: имеет лишь условное математическое ожидание (практически это значит, что среднее по выборке будет устойчивым лишь при условии ее симметрии относительно истинного среднего), и не имеет дисперсии (т.е. выборочная дисперсия будет неограниченной величиной).

- $n$ -мерное нормальное  $\xi \in (-\infty, \infty)$ .

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \dots \\ \xi_n \end{pmatrix} \quad P \begin{pmatrix} \xi_1 < t_1 \\ \xi_2 < t_2 \\ \dots \\ \xi_n < t_n \end{pmatrix} = \int p(t_1, t_2, \dots, t_n) dt_1 dt_2 \dots dt_n.$$

$$p(t_1, t_2, \dots, t_n) = (2\pi)^{-\frac{n}{2}} \det(A)^{\frac{1}{2}} \exp\left(-\frac{1}{2}(AT, T)\right),$$

$$A > 0, \quad T = \begin{pmatrix} t_1 - a_1 \\ t_2 - a_2 \\ \dots \\ t_n - a_n \end{pmatrix} \quad \hat{V}\xi = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{pmatrix}$$

Примеры,  $n = 2$ . Случай (1) диагональной и (2) полной матрицы.

$$1. \quad A = \begin{pmatrix} \frac{1}{\sigma_1^2} & 0 \\ 0 & \frac{1}{\sigma_2^2} \end{pmatrix} \quad 2. \quad A = \frac{1}{1-r^2} \begin{pmatrix} \frac{1}{\sigma_1^2} & -\frac{r}{\sigma_1\sigma_2} \\ -\frac{r}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{pmatrix}$$

Здесь  $r$  коэффициент корреляции. Нормально-распределенные некоррелированные случайные величины являются также и статистически независимыми величинами.

- экспоненциальное распределение  $\xi \in [0, \infty)$ .

Функция распределения, плотность и центральные моменты имеют вид:

$$P(\xi < t) = \begin{cases} 1 - \exp(-\lambda t) & \text{если } t \geq 0; \\ 0 & \text{иначе} \end{cases}$$

$$p(t) = \lambda \exp(-\lambda t) \chi_{[0, \infty)}(t); \quad \hat{E}\xi = \frac{1}{\lambda}, \quad \hat{V}\xi = \frac{1}{\lambda^2}$$

Часто  $\lambda = \frac{\ln 2}{T_{1/2}}$ . Причины:  $P(\xi < T_{1/2}) = \frac{1}{2}$ . Экспоненциальное распределение обычно используется для описания радиоактивных распадов.

- $\chi_n^2$  - распределение.

$$\xi = \sum_{i=1}^n \xi_i^2, \quad \text{где } \xi_i \text{ независимы и имеют распределение } N(0, 1) . \quad \xi \in [0, \infty)$$

Центральные моменты имеют вид:  $\hat{E}\xi = n$ ,  $\hat{V}\xi = 2n$ .

$\chi_n^2$  - распределение играет большую роль при проверке гипотез.

**Псевдо-случайные числа.** На ЭВМ вместо случайных используются псевдо-случайные числа, которые, с одной стороны, формально удовлетворяют самым разнообразным критериям случайности, а, с другой, обладают точно проверяемыми характеристиками.

Обычно исходными являются псевдо-случайные числа, распределенные равномерно на  $(0, 1)$ ; Ниже приводится пример Pascal-программы для получения таких чисел.

```
FUNCTION rndm(x:REAL):REAL;
{ Purpose: to produce a uniformly distributed pseudo-random number
  in the range [0,1]. (F.James, CPC(60)3,p.329)
  Version: May 14 1991 }
VAR k,iz:LONGINT;
BEGIN k:=is1 DIV 53668;
is1:=40014*(is1-k*53668)-k*12211;
IF(is1<0)THEN is1:=is1+2147483563;
k:=is2 DIV 52774;
is2:=40692*(is2-k*52774)-k*3791;
IF(is2<0)THEN is2:=is2+2147483399;
iz:=is1-is2;
IF(iz<1)THEN iz:=iz+2147483562;
rndm:=iz*4.656613E-10;
END;
```

Здесь переменная  $x$  является фиктивной (нужна лишь для совместимости обращения к функции с фортранным обращением); в качестве ее можно взять любое число типа REAL.

Из равномерно распределенных чисел можно построить огромное количество других:

- Нормальное распределение. Если  $\xi_1$  и  $\xi_2$  р.р. (равномерно - распределены) на  $(0, 1)$ , то имеют распределение  $N(0, 1)$  величины  $a = u_1 \cdot s$ ,  $b = u_2 \cdot s$ , где

$$u_1 = 2\xi_1 - 1, \quad u_2 = 2\xi_2 - 1, \quad s = \sqrt{-2 \frac{\ln m}{m}}, \quad m = u_1^2 + u_2^2, \quad \text{если } m < 1.$$

- Экспоненциальное. Если  $\xi$  р.р. на  $(0, 1)$ , то  $a = -\ln(\xi) \cdot \lambda$  распределена экспоненциально с плотностью  $p(t) = \lambda \cdot \exp(-\lambda t)$ .

- Пуассона. Следующий алгоритм даст  $n$ , распределенную по Пуассону с параметром  $\lambda$ . Если  $\xi$  р.р. на  $(0, 1)$ , то

```
n:=0; j:=Round(a/delta);
FOR i:=1 TO j DO
  BEGIN x:=rndm(-1);
  IF(x<delta)THEN Inc(n);
  END;
```

где  $\delta$  - малое число, существенно меньшее  $\lambda$ .

Псевдо-случайные числа могут быть использованы для построения данных, имитирующих реальные данные, но имеющие (в отличие от последних) точно известные характеристики. Обработка таких данных с известным ответом позволяет проверить работу методов анализа реальных данных, излагаемых в данном курсе. Если от (псевдо) случайных статистически независимых чисел  $x_i$ ,  $i = 1, 2, \dots, n$  со средним  $m_1$  и дисперсией  $v_1$  надо перейти к случайным числам  $y$ , имеющим среднее  $m_2$  и дисперсию  $v_2$ , то сделать это можно с помощью преобразования:

$$y_i = \sqrt{v_2} \cdot \frac{x_i - m_1}{\sqrt{v_1}} + m_2$$

Далее от  $x_i$  перейти к попарно-коррелированным с коэффициентом ковариации  $k$  можно с помощью, например, такого преобразования:

$$y_i = (x_i + \frac{k}{v_1} \cdot x_{i+1})$$

Обратная процедура - переход от зависимых величин  $x_i$ ,  $i = 1, 2, \dots, n$  к независимым (некоррелированным)  $y_i$  можно осуществить с помощью, например, такого преобразования:

$$Y = A \cdot X,$$

где  $A$  - специально подобранная ортогональная матрица ранга  $n$ .

## ЛЕКЦИЯ 2. ОЦЕНИВАНИЕ ПАРАМЕТРОВ ФУНКЦИЙ РАСПРЕДЕЛЕНИЯ.

Пусть дана выборка  $T$  - набор случайных (или псевдо-случайных) чисел  $t_1, t_2, \dots, t_m$ , подчиненных функции распределения  $P(\xi < t)$ , зависящей от параметра  $a_{true}$ , т.е. имеющая вид  $P(\xi < t, a_{true})$ .

Фундаментальные задачи обработки таких данных методами математической статистики - это следующие:

- Оценивание параметров: по заданной выборке  $T$  и известному виду функциональной зависимости  $P$  от  $a_{true}$  построить оценку параметра  $a_{true}$  и определить точность этой оценки.
- Проверка гипотезы (или принятие решения): по заданной выборке  $T$  и  $a_0$  - известному значению (или предполагаемому известным)  $a_{true}$ , проверить достоверность гипотезы  $a_{true} = a_0$ .
- Планирование эксперимента: по известному виду функциональной зависимости  $P$  от  $a_{true}$  построить план проведения эксперимента: т.е. определить объем и характер выборки, гарантирующей получение оценки параметра  $a_{true}$  с заданной точностью или проверку гипотезы  $a_{true} = a_0$  с заданной достоверностью.

**Задача построения оценок.** Оценки параметров могут быть построены любым способом, лишь бы он гарантировал хорошие статистические качества этих оценок. Здесь очень важно отметить, что хорошими оценки величин по случайным данным могут быть лишь в статистическом смысле. Термин "статистический" следует понимать как ансамблевое свойство: т.е. такое, какое характеризует множество возможных значений случайной величины в целом, но не каждое отдельное значение. Это значит, что даже если оценка параметра "хороша" в статистическом смысле, она не обязана быть таковой в каждом отдельном случае. Оценивание, как и другие статистические процедуры, являются действиями, сопряженными с известным риском, и успех их в отдельных случаях есть дело в первую очередь конечно умения, но затем и счастливого шанса, т.е. везения. Неудачный расклад данных может породить научные иллюзии. Цель математики - сведение возможностей для таких иллюзий к минимуму.

Поскольку оценка строится на основе случайных данных, она сама является случайной величиной, и, следовательно, имеет свою функцию распределения, центральные 1-е, 2-е, и т.д. моменты. Если удастся построить такую функцию распределения, то с ее помощью можно вычислить среднее и дисперсию оценки, и тем самым определить качество этой оценки.

Например, естественно потребовать, чтобы математическое ожидание оценки параметра совпадало с его истинным значением, т.е. чтобы оценка была **несмещенной**; далее, желательно, чтобы дисперсия этой оценки была как можно меньше - это в какой-то степени делает ее независимой от конкретной выборки и уменьшает риск оценивания.

Рассмотрим некоторые примеры, иллюстрирующие сказанное. Пусть дана выборка  $T = (t_1, t_2, \dots, t_m)$  некоторой случайной величины  $\psi$ , подчиненной равномерному распределению на  $(0, a)$ , и пусть параметр этого распределения  $a$  неизвестен.

Мы можем рассуждать так: так как среднее по этой выборке будет очевидно близко к точке, делящей отрезок  $(0, a)$  пополам, то в качестве оценки  $a$  возьмем величину

$$\hat{a} = 2 \cdot \sum_{i=1}^m t_i / m$$

Вычислим среднее и дисперсию этой величины

$$\hat{E}\hat{a} = 2 \cdot (m \cdot a/2) / m = a$$

$$\hat{V}\hat{a} = 4 \cdot m \frac{a}{12} / m^2 = \frac{a^2}{3m}$$

Видна, что оценка неплоха в том смысле, что она является несмещенной, т.е. ее математическое ожидание равно истинному значению параметра  $a$ . И вопрос в том только, а насколько мала дисперсия.

Возьмем другую оценку параметра  $a$ :

$$\hat{a} = \max t_i, \quad i = 1, 2, \dots, m.$$

Ее функция распределения равна

$$P(\hat{a} < x) = \begin{cases} (\frac{x}{a})^m & \text{если } 0 \leq x \leq a; \\ 1 & \text{если } x > a; \\ 0 & \text{если } x < 0 \end{cases}$$

Плотность функции распределения равна  $p(x) = \frac{x^{m-1}}{a^m}$  в интервале  $(0, a)$  и нулю вне его.

$$\hat{E}(\max t_i) = \frac{m}{a^m} \int_0^a x^m dx = \frac{ma}{m+1}$$

Видно, что оценка смещена, и ее смещение  $= \frac{1}{m}a$ .

Дисперсия:

$$\hat{E}(\max t_i)^2 = \frac{m}{a^m} \int_0^a u^{m+1} du = \frac{m}{m+2}a^2$$

Откуда, используя формулу (47), получаем

$$\hat{V}(\max t_i) = \frac{m}{(m+2)(m+1)^2}a^2$$

Мы можем построить несмещенную оценку

$$\hat{a} = \frac{m+1}{m} \max t_i, \quad i = 1, 2, \dots, m.$$

Ее математическое ожидание равно  $a$ , а дисперсия

$$\hat{V}\left(\frac{m+1}{m} \max t_i\right) = \frac{(m+1)^2}{m} \frac{m}{(m+2)(m+1)^2} a^2 = \frac{1}{m(m+2)a^2}$$

Эта дисперсия не просто меньше, чем дисперсия первой оценки (среднего), а имеет даже другую асимптотику. Если первая дисперсия с ростом  $m$  была пропорциональна  $\frac{1}{m}$  (случай т.н. эффективной оценки), то дисперсия второй оценки с ростом  $m$  будет

пропорциональна  $\frac{1}{m^2}$  (это уже случай суперэффективной оценки).

Эффективность оценки означает, что для уменьшения сигмы оценки в 2 раза (увеличение статистической точности в 2 раза) требуется выборка, большая чем начальная в 4 раза; суперэффективность означает, что для уменьшения сигмы оценки в 2 раза требуется выборка, большая, чем начальная лишь в 2 раза. Суперэффективность очень важна для обработки данных с малой статистикой, т.е. выборок малого размера. Но к сожалению, суперэффективные оценки существуют лишь в редких случаях.

Но первая оценка имеет одно важное преимущество: арифметическое среднее является всегда (для любого распределения) эффективной оценкой математического ожидания, и стало быть, ей можно оценить параметры рассмотренных выше нормального, экспоненциального, Пуассона и многих других распределений.

Другой важный параметр, присутствующий в распределениях, это дисперсия. Существует универсальный, независимый от распределения, способ построения также эффективной оценки и этой величины, если, конечно, распределение имеет конечную дисперсию (некоторые, вроде рассмотренного выше распределения Коши, ее не имеют).

Возьмем в качестве оценки дисперсии средне-квадратичное отклонение выборки от среднего (или его оценки)

$$s^2 = \frac{1}{m} \sum (t_i - \bar{t})^2$$

Здесь могут быть следующие варианты:

1. Если  $\bar{a}$  есть точное известное значение  $a$ , то  $\hat{E}s^2 = \frac{1}{m}m\sigma^2 = \sigma^2$ .
2. Если  $\bar{a}$  выборочное среднее, то  $s^2 = \frac{1}{m} \sum (t_i - a)^2 - (\bar{t} - a)^2$ , и мы имеем  $\hat{E}s^2 = \frac{m-1}{m}\sigma^2$

Во втором случае оценка смещена. Можно, конечно, построить несмещенную оценку, умножив  $s^2$  на  $\frac{m}{m-1}$ , но дисперсия несмещенной оценки возрастет настолько (детальные громоздкие выкладки здесь опущены, но они есть в любом учебнике по статистике), что суммарная мера неточности оценки, определяемая как  $B^2 + V$ , где  $B, V$  - смещение и дисперсия, соответственно, окажется больше чем такая же мера у смещенной оценки. Следовательно, не всегда несмещенность означает большую точность.

Рассмотренный подход может быть распространен и на другие оценки параметров распределений, если эти параметры являются функциями их центральных моментов: для этого достаточно взять в качестве оценок соответствующие функции от выборочных моментов - это так называемый метод моментов (ММ); а именно, если  $a = f(M_i)$ , где  $M_i$  - момент  $i$ -ого порядка, то  $\hat{a} =$  корню уравнения

$$\hat{a} = f(\hat{M}_i)$$

где  $\hat{M}_i$   $i$ -ый выборочный момент.

К сожалению, качество ММ-оценок приемлемо лишь для моментов невысокого (первого и второго) порядков.

**Оценки максимального правдоподобия.** Построение эвристических оценок

является делом трудоемким и без гарантии успеха. Естественно, возникает необходимость в универсальном, а с другой стороны, достаточно точном методе построения оценок, который был бы реализуем в самом широком классе случаев на практике. Таким методом является метод максимального правдоподобия.

Введем следующие понятия. Пусть дана выборка  $T = (t_1, \dots, t_m)$ , и  $P(\xi < t, a)$  или  $p(t, a)$  - ее функция распределения или плотность этой функции, зависящие от параметра  $a$ . Тогда "правдоподобие"  $L(a)$  определяется следующим образом:

- если функция распределения (плотность) непрерывна; тогда

$$L_c(a) = \prod_{i=1}^m p(t_i, a)$$

- если функция распределения (плотность) дискретна; тогда

$$L_d(a) = \prod_{i=1}^m P(t_i, a)$$

- логарифмическое;

$$l_c(a) = \ln L_c(a) = \sum_{i=1}^m \ln(p(t_i, a))$$

$$l_d(a) = \ln L_d(a) = \sum_{i=1}^m \ln(P(t_i, a))$$

- Энтропия.

$$E(a) = \sum_{i=1}^m p(t_i, a) \ln(p(t_i, a)).$$

С помощью введенного понятия мы можем определить "Оценку Максимального Правдоподобия" (MLE) параметра выборки следующим образом: это то значение  $a$ , которое обращает в максимум правдоподобие для данной выборки

$$\hat{a} = ARG \ MAX L(a)$$

Не имеет значения, какое правдоподобие используется: прямое или логарифмическое, так как максимум логарифма функции достигается в тех же точках, что и максимум самой функции.

Оценка MLE естественным образом обобщается на случай вектора параметров: это вектор значений параметров, на котором правдоподобие достигает максимума.

Если  $L(a)$  дифференцируема по  $a$ , то MLE находится из условия

$$\frac{dL}{da} = 0$$

В случае многих параметров  $A = (a_i)$ ,  $i = 1, \dots, n$  из системы уравнений

$$\frac{\partial L}{\partial a_i} = 0, \quad i = 1, \dots, n$$

Ниже можно рассмотреть MLE-оценки для наиболее типичных функций распределения.

- мультибиномиальное  $\xi = \sum_{i=1}^n \xi_i$ : пусть мы имеем одно значение  $t_0$  ;

$$\ln L(p) = \ln(C_n^{t_0}) + t_0 \ln(p) + (n - t_0) \ln(1 - p)$$

$$\frac{\partial l}{\partial p} \equiv \frac{t_0}{p} - \frac{n - t_0}{1 - p} = 0 \quad \text{отсюда } \hat{p} = \frac{t_0}{n}$$

Если выборка содержит много значений  $t_1, \dots, t_m$

$$\hat{p} = \frac{\sum_{i=1}^k t_i}{m \cdot n}$$

- Пуассона. Одно значение  $t_0$ ;

$$\ln L(a) = \ln\left(\frac{a^{t_0}}{t_0!} \exp(-a)\right). \quad \frac{\partial l}{\partial a} = 0 \equiv t_0 \cdot \ln(a) - 1 = 0. \quad \text{отсюда } \hat{a} = t_0$$

Много значений  $t_1, \dots, t_m$

$$\hat{a} = \frac{\sum_{i=1}^k t_i}{m}$$

- равномерное  $\xi = [0, a]$

$$p(t, a) = \frac{1}{a} \chi_{[0, a]}(t). \quad L(a) = \left(\frac{1}{a}\right)^m \prod_{i=1}^m \chi_{[0, a]}(t_i). \quad \text{отсюда } \hat{a} = \max(t_i)$$

так как, если  $\hat{a} > t_{max}$ , 1й множитель мал; если  $\hat{a} < t_{max}$  некоторые  $\chi_{0, \hat{a}}(t_i) = 0$ .

- экспоненциальное

$$\ln L(\lambda) = n \ln(\lambda) - \lambda \sum_{i=1}^m t_i. \quad \frac{\partial l}{\partial \lambda} \equiv \frac{m}{\lambda} - \sum t_i = 0.$$

$$\hat{\lambda} = \frac{m}{\sum_{i=1}^m t_i}; \quad T_{1/2}^{\hat{\lambda}} = \frac{\ln 2 \sum_{i=1}^m t_i}{m}$$

- нормальное

$$\ln L(a, \sigma) = -m \ln(\sigma) - \sum_{i=1}^m \frac{(t_i - a)^2}{2\sigma^2} - m/2 \cdot \ln(2\pi)$$

$$\frac{\partial l}{\partial a} \equiv \sum_{i=1}^m \frac{t_i - a}{\sigma^2} = 0. \quad \frac{\partial l}{\partial \sigma} \equiv -\frac{m}{\sigma} + \sum_{i=1}^m \frac{(t_i - a)^2}{\sigma^3} = 0.$$

$$\hat{a} = \sum_{i=1}^m t_i / m; \quad \hat{\sigma}^2 = \frac{1}{m} \sum_{i=1}^m (t_i - \hat{a})^2$$

**Другие оценки.** Оценки максимальной энтропии (МЕЕ). Определяются из условия

$$\hat{a} = \text{ARG MAX} E(a)$$

Если  $E(a)$  дифференцируема по  $a$ , то МЕЕ находится из  $\frac{\partial E}{\partial a} = 0$ .

ММЕ имеют меньшую точность, нежели МЛЕ-оценки, но обладают большей

робастностью - устойчивостью к неадекватности указания функции распределения вероятностей истинной, породившей данную выборку.

Большой интерес представляют Байесовские оценки (ВЕ). Они основаны на философии рандомизации параметров, т.е. предположении, что параметры можно рассматривать как случайные величины, о которых нам известно так называемое априорное распределение  $\pi(a)$  с соответствующими мат.ожиданием и дисперсией. Тогда в качестве ВЕ-оценки параметра  $a$  берется его математическое ожидание по этому априорному распределению:

$$\hat{a} = \int a d(\pi(a))$$

Теорема: в достаточно широком классе случаев

$$\pi(a) = \frac{L(a)}{\int L(a) da}$$

Тогда

$$\hat{a} = \frac{\int a L(a) da}{\int L(a) da}$$

Пример. Распределение Пуассона  $= a^m/m!exp(-a)$ , Пусть выборка  $T$  состоит из одного  $t$ . ВЕ-оценка  $a$  равна

$$\hat{a} = \frac{\int_0^\infty a a^t/t!e^{-a} da}{\int a^t/t!e^{-a} da} = \frac{\int_0^\infty a^{t+1}e^{-a} da}{\int a^t e^{-a} da} = (t+1) \frac{\int_0^\infty a^t e^{-a} da}{\int a^t e^{-a} da} = t+1$$

ВЕ-оценки как правило смещены, но иногда обеспечивают большую точность, чем несмещенные оценки (например, MLE-оценки). Однако для этого необходимо знать весьма точное априорное распределение искомого параметра, чего на практике часто не бывает.

Так что из рассмотренных оценок наилучший компромиссом между точностью, надежностью и возможностью реализации на практике обладают MLE-оценки - оценки максимального правдоподобия.

### ЛЕКЦИЯ 3. ТОЧНОСТЬ ОЦЕНОК ПАРАМЕТРОВ.

Точность выше была определена как среднеквадратичное отклонение, т.е. как дисперсия + смещение в квадрате. Смещение могут иметь все оценки, в том числе и МЛЕ-оценки.

Пример смещенной МЛЕ-оценки. Пусть задана выборка  $T = (t_1, \dots, t_m)$ , подчиненная равномерному распределению на  $(0, a)$ . Как было показано выше МЛЕ-оценкой является  $\hat{a} = \max(t_i)$ . Имеем

$$P(\max t_i < t) = \prod_{i=1}^m P(t_i < t) = \frac{t^m}{a^m}$$

Откуда находим математическое ожидание этой оценки

$$\hat{E}(\max t_i) = \int_0^a t \left(\frac{t}{a}\right)^m dt = m \int_0^a \left(\frac{t}{a}\right)^m dt = \frac{am}{m+1}$$

Оценка смещена и смещение =  $\frac{a}{m+1}$ .

Другой составляющей точности оценок является дисперсия. Для весьма широкого класса оценок дисперсия может быть оценена по крайней мере снизу. Для этого вводится понятие количества информации по Фишеру для отдельного события

$$I_1 = \hat{E}\left(\frac{\partial \ln(p(a))}{\partial a}\right)^2$$

Или

$$I_1 = \hat{E}\left(\frac{p'_a}{p}\right)^2$$

Количество информации по Фишеру для выборки из  $m$  событий определяется так

$$I_m(a) = mI_1(a)$$

С помощью этого понятия формулируется неравенство Рао-Крамера, дающее искомую границу дисперсии снизу

$$\hat{V}\hat{a} \geq I_m(a)^{-1}$$

Для некоторого класса оценок (т.н. достаточных) неравенство превращается в равенство, и, следовательно, такие оценки являются эффективными.

1. ПРИМЕР. (Нормальное распределение)

$$I_1(a) = \hat{E}\left(\frac{(t_i - a)^2}{\sigma^4}\right) = \frac{1}{\sigma^2}. \quad I_m = \frac{m}{\sigma^2}$$
$$\hat{V}\hat{a} = \frac{\sigma^2}{m}$$

2. ПРИМЕР. (Пуассона)

$$I_1(a) = \hat{E}\left(\frac{m_0}{a} - 1\right)^2 = \frac{1}{a^2} \hat{E}(m_0 - a)^2 = \frac{a}{a^2} = \frac{1}{a}. \quad I_m = \frac{m}{a}$$
$$\hat{V}\hat{a} = \frac{a}{m}$$

В заключение следует отметить хорошие качества МЛЕ-оценок, по крайней мере, в асимптотике, т.е. когда объем выборки  $m$  неограниченно возрастает. В широком классе случаев зависимостей функций распределения от параметров, МЛЕ-оценки  $\hat{a}$  параметров  $a$  асимптотически:

- несмещены, т.е.  $\lim_{m \rightarrow \infty} \hat{a} = a$ .
- состоятельны, т.е.  $P(\lim_{m \rightarrow \infty} \hat{a} = a) = 1$ .
- эффективны, т.е.  $\hat{V}a \sim \frac{C}{m}$ .
- нормально распределены, т.е. выражение  $(\hat{a} - a)/\sqrt{m}$  тем ближе будет описываться нормальным распределением  $N(0, \hat{V}a)$ , чем больше  $m$ .

## ЛЕКЦИЯ 4. ПРОВЕРКА ГИПОТЕЗ И ПРИНЯТИЕ РЕШЕНИЙ.

Пусть заданы выборка  $T = t_1, \dots, t_n$  и  $p(t, a)$  - плотность функции распределения. Рассмотрим 2 задачи:

- проверить гипотезу, что параметр  $a$  равен величине  $a_0$ ;
- из 2 гипотез:  $a = a_0$  и  $a = a_1$  - выбрать ту, которая соответствует выборке  $T$ .

Простейший способ решения первой задачи состоит в следующем. Строится так называемая статистика от выборки  $T$  (т.е. некоторая функция от выборочных значений)  $f(t_1, t_2, \dots, t_m)$  при проверяемом значении параметра  $a_0$ . Иначе статистика называется критерием. Статистика  $f$  как функция от случайных величин сама будет случайной величиной, имеющей соответствующую функцию распределения  $G(f)$  или плотность  $g(f)$ .

Далее выбирается уровень значимости критерия проверки  $\alpha$  и строится доверительный интервал  $[f_1, f_2]$  так, что значения случайной величины  $f(T)$  при различных наборах  $t_i$ ,  $i = 1, 2, \dots, m$  и при  $a = a_0$  попадают в этот интервал с вероятностью  $1 - \alpha$ . Теперь, если значение критерия  $f$  для данных  $T, a_0$  попадает в интервал  $[f_1, f_2]$ , то гипотеза  $a = a_0$  принимается; иначе не принимается.

Если  $m = 1$ , статистикой  $f$  будет само значение  $t_1$ , а плотностью ее функции распределения сама  $p(t, a)$ . Действия те же самые, что и для случая произвольного  $m$ .

В случае симметричных по знаку случайных величин уровень значимости критерия выбирается для модуля случайной величины, и тогда размер интервала незначимости бывает равен не  $\alpha$ , а  $2\alpha$ .

Доверительный интервал часто строят на основе использования среднего и дисперсии  $\sigma^2$  статистики  $f$ :  $a \pm \sigma$  или  $a \pm k\sigma$ ,  $k = 1, 2, \dots$ , что соответствует различным  $\alpha$ .

Такой способ хорош для таких симметричных распределений как нормальное или равномерное. Но в случае произвольного распределения он может дать грубо неточную доверительную область.

Рассмотрим некоторые примеры использования проверки гипотез.

1. Тест для нормальной плотности  $H : a = a_0$ ,  $\sigma$  известно . Статистикой берем

$$f(t_1, \dots, t_n) = \frac{1}{n} \sum_{i=1}^n \frac{t_i - a_0}{\sigma}$$

$$g(f, a) = N\left(0, \frac{1}{n}\right)$$

2. Тест для нормальной плотности  $H : a = a_0$ ,  $\sigma$  неизвестно . Здесь

$$f(t_1, \dots, t_n) = \sqrt{n-1} \frac{b - a_0}{s}, \text{ где } b = \frac{1}{n} \sum_{i=1}^n t_i$$

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (t_i - b)^2}$$

$g(f, a) = S(n-1) =$  распределение Стьюдента с  $n-1$  степенями свободы .

3. Тест для нормальной плотности  $H : a = a_0$  и  $\sigma$  известно.

$$f(t_1, \dots, t_n) = \sum_{i=1}^n \left( \frac{t_i - a_0}{\sigma} \right)^2$$

$$g(f, a) = \chi_n^2 = \text{распределение с } n \text{ степенями свободы.}$$

Если  $a$  неизвестно, и  $H : \sigma = s_0$

$$f(t_1, \dots, t_n) = \sum_{i=1}^n \left( \frac{t_i - b}{s_0} \right)^2, \quad \text{где } b = \frac{1}{n} \sum_{i=1}^n t_i$$

$$g(f, a) = \chi_{n-1}^2$$

Рассмотрим теперь случай конкурирующих гипотез. Пусть гипотезы  $H_0 : a = a_1$  и  $H_1 : a = a_2$  конкурируют. Принимая какое-либо решение относительно этих гипотез, мы можем сделать ошибки двух типов:

Ошибка 1го типа:  $H_0$  верна, но мы отвергаем ее.

Ошибка 2го типа:  $H_1$  верна, но мы принимаем  $H_0$ .

Естественное требование - минимум обеих ошибок - не всегда может быть выполнено.

Чаще всего минимизация одной ошибки не уменьшает (а то и увеличивает) вторую ошибку. Поступают так: задавая границу для ошибки одного типа, строят критерий принятия решения так, чтобы он минимизировал ошибку 2-го типа.

Простейшим примером этого может служить критерий отношения правдоподобий, посчитанных для обеих гипотез:

$$r = \frac{L(a \in H_0)}{L(a \in H_1)}; \quad \text{или} \quad r = \frac{\ln L(a \in H_0)}{\ln L(a \in H_1)};$$

Строится интервал значимости  $I = (1 - \delta_1, 1 + \delta_2)$ .

Если  $r > 1 + \delta_2$ , тогда  $H_0$  принимается; если  $r < 1 - \delta_1$ , принимается  $H_1$ ; если же  $r \in I$ , гипотезы неразличимы.

Для заданного уровня ошибки 1го рода, этот критерий минимизирует ошибку 2-го рода.

Неразличимость (т.е. недискриминируемость) гипотез может быть, в частности, следствием малого объема данных, или их неразличимости в критических областях проверяемых распределений. Для однозначной работы критерия дискриминации гипотез надо, чтобы данные достаточно полно соответствовали своему истинному распределению и имели достаточный объем.

В приведенной процедуре доверительный интервал строится весьма упрощенно: с помощью среднего и дисперсии. Этого достаточно для ряда симметричных распределений. Но в общем случае, разумеется, необходим более сложный алгоритм построения доверительного интервала: на основе распределения  $r$ .

## ЛЕКЦИЯ 5. ОЦЕНИВАНИЕ ПАРАМЕТРОВ РЕГРЕССИЙ. ЛИНЕЙНЫЙ СЛУЧАЙ.

В экспериментальной физике низких энергий наиболее распространенной формой данных является функция  $y(x)$ ,  $x \in [x_1, x_2]$ , математической моделью которой служит регрессия, используемая несколько в ином, нежели классический, контексте, а именно:

$$y(x) = m(x) + e(x),$$

где  $e(x)$  случайная величина для каждого  $x$  и описывается с помощью семейства функций распределения ( $P(e(x) < t(x))$ ); это может быть и одна функция (функциональная зависимость одного типа), но имеющая разные центральные моменты в разных точках  $x$ .

В важнейших частных случаях имеет место

$$\hat{E}e(x_j) = 0, \quad \hat{V}e(x_j) = v(x_j), \quad \hat{E}(e(x_j)e(x_i)) = c_{ij} = 0 \text{ при } i \neq j$$

В частности, может иметь место (и часто имеет)  $c(x_i, x_j) = 0$ .

Видно, что  $\hat{V}y(x) = \hat{V}m(x) + \hat{V}e(x) = v(x)$ . Иначе  $v(x)$  = квадрат ошибки регрессии. Аргумент  $x$  это в общем случае векторная величина.

Следует различать между регрессией и случайной величиной - это не одно и то же. Разница в определенном смысле та же, что в анализе между переменной и функцией.

Типичным примером таких регрессий является экспериментальный спектр - гамма-спектр, нейтронный или нейтронно-дифракционный и т.д.

Экспериментальное распределение  $y(x)$  получается, как правило, либо суммированием (осуществляемом либо аппаратно, либо математически) отдельных сигналов - случайных величин, либо оцифровкой некоторых непрерывных кривых.

Важный частный случай дискретных распределений, получаемых суммированием, - это случай суммирования т.н. редких событий, т.е. таких, что вероятность регистрации за малый промежуток времени  $\delta t$  двух и более сигналов является бесконечно-малой величиной по сравнению с  $\delta t$ . Суммы таких сигналов распределены приближенно по Пуассону. Это дает нам возможность хотя бы приближенно оценить дисперсию  $y(x)$ , а именно, как мы видели выше, оценкой дисперсии случайной величины, распределенной по Пуассону, служит само  $y(x)$  (MLE-оценка), или  $y(x) + 1$  (BE-оценка).

Функция  $m(x)$  является детерминированной составляющей регрессии и собственно содержит интересующую физика информацию, правда, как правило, в неявном виде, т.е. для непосредственного ее использования требуется извлечение ее из  $m(x)$ . Такое извлечение сопровождается как правило компрессией этой информации, т.е. преобразованием ее к набору величин, существенно меньшему, чем множество значений  $m(x)$ .

Наиболее универсальный способ редукции регрессии - это параметризация модели  $m(x)$ , т.е. введение в нее параметров  $P$  так, чтобы они однозначно были связаны с этой функцией, и в то же время содержали в себе в компактной форме ту физическую информацию, которая в распределенном виде заключена в  $m(x)$ :  $m(x) \rightarrow m(x, P)$ , где  $P$  (в общем случае вектор размерности  $n$ ) и есть искомый параметр.

**Примеры некоторых параметризаций.** Линейная модель. Это случай, когда параметр (но не обязательно аргумент) входят в описание функции линейно.

- $m(x, P) = ax + b$ ,  $P = (a, b)$ ; Данная модель линейна как по параметру, так и по аргументу.
- $m(x, P) = ax^2 + bx + c$ ,  $P = (a, b, c)$ ; Здесь аргумент  $x$  входит в модель уже нелинейно.
- $m(x, P) = \sum_{i=0}^{n-1} a_i x^i$ ,  $P = (a_0, a_1, \dots, a_{n-1})$ ; Это пример полиномиальной модели, которая, в частности, может использоваться как приближение некоторой гладкой функции в конечной области ее определения.
- $m(x, P) = \sum_{i=0}^{n-1} a_i f_i(x)$ ,  $P = (a_0, a_1, \dots, a_{n-1})$ . Это пример более общего описания регрессии как линейной (с неизвестными коэффициентами  $a_i$ ) комбинации фиксированных функций  $f_i(x)$ ; такое описание может быть и приближенным.

Линейные модели обладают значительными математическими достоинствами, но не всегда справляются с главной задачей параметризации: дать компактное описание информации, заключенной в модели регрессии  $m(x)$ . Кроме того, часто параметризация не выбирается нами по желанию, а задается условиями физического эксперимента, и она часто оказывается нелинейной как по аргументу  $x$ , так и по параметрам  $P$ .

Рассмотрим примеры некоторых нелинейных моделей регрессий  $m(x, P)$  - нелинейных по вектору параметров  $P$  и по аргументам  $x$ .

$$A \exp\left(-\left(\frac{x-C}{W}\right)^2\right), \quad A e^{-\frac{\ln 2t}{T}}, \quad \frac{A}{1 + ((x-C)^2/(W^2))}, \quad A \sin(Wt).$$

Нелинейные модели дают как правило более реалистичное, т.е. более точное описание наблюдаемых данных, чем линейные, но и они могут оказаться неприемлемыми во многих случаях, главным образом изза того, что используют небольшой набор весьма идеальных гладких кривых, таких как  $\exp$ ,  $\sin$ , и т.д. Реально наблюдаемые эффекты могут быть описаны идеальными кривыми точно лишь в редких случаях.

В качестве примера нелинейной модели регрессии рассмотрим пример линейчатого спектра излучения (или поглощения), порождаемого воздействием на мишени потоком частиц или квантов электромагнитного поля.

Линии в этих спектрах представляют собой фигуры типа резонансов, расположенные на некотором пологом образовании. Идеальный резонанс (или пик) очень удобно описывать функцией Гаусса типа  $m(x) = \exp(-\frac{x^2}{2})$  или функцией Коши  $m(x) = \frac{1}{1+x^2}$ , которые геометрически достаточно подобны контуру этих резонансов (пиков). Остается ввести параметры в эти модели и регрессионная модель пика будет готова. Эти параметры следующие:  $m(x) = A \exp(-\frac{(x-C)^2}{2W^2})$  и  $m(x) = \frac{A}{1+(x-C)^2/W^2}$ . В этих описаниях пиков параметры имеют четкий геометрический смысл:  $A$  - амплитуда,  $C$  - центр,  $W$  - "ширина" пика. На основе  $A, W$  можно вычислить площадь пика. Эти величины счастливым образом совпадают с тем, что интересует физика в его экспериментальных спектрах.

Однако геометрические формы пиков могут быть и не столь идеальными как вышеперечисленные и могут иметь особенности, не описываемые ими, например, асимметрию, локальные нерегулярности, и т.д. В этом случае, можно действовать следующим образом.

- Первый путь - увеличение количества параметров. В общем случае чем больше параметров, тем разнообразнее кривые, которые параметрическая функция может описать. Например, функция  $m(x) = A \exp(-\frac{(x-C)^2}{(Lx+W)^2})$ , зависящая от 4 параметров может описать не только симметричный пик (при  $L = 0$ ), но и асимметричный (при  $L \neq 0$ ). Кроме того, следует иметь в виду, что функция типа

$$\sum_{i=0}^n a_i x^i \exp(-\frac{(x-C)^2}{2W^2}),$$

зависящая от  $n + 2$  параметров ( $a_i, C, W, i = 0, 1, \dots, n$ ), представляет собой отрезок ряда Эрмита, который даже при фиксированных  $C, W$  способен описать любую аналитическую финитную (хотя бы приближенно, как гауссиан) функцию при  $n \rightarrow \infty$ ; а полином  $\sum_{i=0}^n a_i x^i$  - функция от  $n$  параметров, как часть степенного ряда может описать любую аналитическую функцию на всей оси.

- Однако увеличение числа параметров имеет также и негативные последствия:
  1. уменьшение в среднем точности каждого отдельного параметра при их определении; если при этом дополнительные параметры не имели физического смысла, а вводились лишь для большей похожести модели на данные. то это означает потери физической информации;
  2. возникновение чисто математических трудностей (см.ниже) при обработке данных;
  3. часто целью анализа физических данных является разложение суммарной линии на составляющие ее компоненты (декомпозиция регрессии); если число параметров, описывающих каждую компоненту, слишком велико, то надежность разложения резко падает: суммарная линия описывается хорошо, а отдельные компоненты плохо - изза дополнительных степеней свободы у модели регрессии.

Поэтому, требование минимума числа параметров является необходимым условием для хорошего анализа данных в дальнейшем. Другими словами, число параметров должно быть оптимальным: не слишком маленьким, но и не слишком большим.

Одним из способов поддержания числа параметров на минимальном уровне при одновременно достаточной точности соответствия модели данным служит следующий метод (метод реальной формы).

Выбирается типичная форма линии регрессии (или отдельной ее компоненты, если она есть сумма последних). Пусть это будет функция  $u(x)$ . Она может быть построена из теоретических соображений как аналитическая функция, или взята как гистограмма, дополненная до полной функции, например, с помощью интерполяции. Далее, параметрическим семейством, построенным на базе этой реальной модели, будет следующее:

$$O_k(x)u(P_n(x)), \tag{1}$$

где  $Q_k(x), P_n(x)$  суть полиномы степеней  $k, n$  соответственно. Множество (1) есть множество функций, зависящих от  $k + n + 2$  параметров. Можно доказать, что каждая функция из этого множества будет сходиться к моделируемой функции в смысле самых различных метрик при  $k, n \rightarrow \infty$ . Но самое замечательное это то, что приближенной сходимости можно добиться при фиксированных степенях полиномов (при этом небольших) за счет подгонки самой модели  $u(x)$ . Все экзотические, необычные свойства моделируемой функции следует учитывать в модели  $u(x)$ , а на параметры оставлять лишь основные, представляющие главный интерес для физика особенности моделируемой функции. Практика показала, что при анализе экспериментальных спектров хорошее качество аппроксимации дает уже модель

$$Au\left(\frac{x-C}{W}\right), \quad (2)$$

представляющая "физически осмысленную" модификацию (1), где

$$Q_k(x) = A, \quad P_n(x) = \frac{1}{W}x + \frac{C}{W},$$

т.е.  $k = 0, n = 1$ , при условии, что  $u(x)$  хорошо отображает общий контур линии, представляющей интерес.

Обобщением (1) является выражение, в котором вместо полиномов  $Q_k(x), P_n(x)$  берутся любые функции, обладающие аппроксимационными свойствами: тригонометрические полиномы, дробно-рациональные функции, сплайны и т.д. Например, практически важным выражением будет  $Au\left(\frac{x-C}{W_1x+W_0}\right)$ , обладающие более эффективными аппроксимационными свойствами, чем рассмотренное выше (2).

Такая параметризация используется в описываемой ниже программе VMRIA.

**Регрессионный анализ. Линейный случай.** Итак, пусть задана регрессия  $y(x_j) \quad j = 1, \dots, m$ , ее параметрическая модель  $f(x, P)$ , где  $P$  —  $n$ -мерный вектор, и функция дисперсии  $d(x)$  в каждой точке  $x$ .

Фундаментальные задачи обработки таких данных методами математической статистики - это в общем-то те же, что и при анализе распределений случайной величины:

- Оценивание параметров  $P$  и определение точности этих оценок.
- Проверка гипотезы (или принятие решения): по заданной регрессии  $y(x)$  и  $P_0$  - известному (или предполагаемому известным) значению параметров, проверить достоверность гипотезы  $P_{true} = P_0$ .
- Планирование эксперимента: по известному виду функциональной зависимости  $P$  от  $P_{true}$  построить план проведения эксперимента: т.е. определить точность и характер регрессии, гарантирующей получение оценки параметров  $P$  с заданной точностью или проверку гипотезы  $P_{true} = P_0$  с заданной достоверностью.

**Задача построения оценок.** Оценки параметров регрессий могут быть построены тоже любым способом, лишь бы он гарантировал хорошие статистические качества

этих оценок. Оценки параметров регрессий тоже являются случайными величинами, и, следовательно, имеют свои функции распределения, центральные 1-е, 2-е, и т.д. моменты. Если такая функция распределения известна, то с ее помощью можно вычислить средние и дисперсии оценок, и тем самым определить качество этих оценок.

Как и в случае оценок параметров функций распределения случайных величин, естественно потребовать, чтобы математическое ожидание оценок параметров совпадало с их истинным значением, т.е. чтобы оценки были несмещенными; и конечно, чтобы дисперсии этих оценок были как можно меньше.

Самым известным и популярным методом получения таких оценок является метод наименьших квадратов, кратко МНК, или LSE. Строгое математическое обоснование его существует лишь в линейном случае, т.е. для линейной зависимости моделей регрессий от параметров, поэтому, изложение данного метода мы начнем с линейного случая.

Итак задана регрессия  $y(x_j)$ ,  $j = 1, \dots, m$ , ее параметрическая модель  $f(x, P)$ , где  $P$  -  $n$ -мерный вектор, и дисперсия  $d(x)$ . Оценки МНК это значения параметров, минимизирующие по  $P$  выражение

$$S(P)^2 = \sum_{j=1}^m w(x_j)(y(x_j) - f(x_j, P))^2, \quad \text{где } w(x) = \frac{1}{d(x)}. \quad (3)$$

Если рассматривать только гладкие (дифференцируемые) по параметрам функции  $f(x, P)$ , то для нахождения этих значений необходимо решить систему  $n$  следующих уравнений (МНК-уравнений)

$$\frac{\partial S(P)^2}{\partial p_i} = 0, \quad i = 1, \dots, n$$

или

$$\sum_x w(x)(y(x) - f(x, P)) \frac{\partial f}{\partial p_i} = 0$$

Для линейной модели  $f(x, P) = \sum_{i=1}^n p_i g_i(x)$  эти уравнения имеют вид

$$\sum_x w(x) g_i(x) \sum_{k=1}^n g_k(x) p_k = \sum_x w(x) y(x) g_i(x)$$

или в матричной записи  $AP = B$ , где элементы матрицы и вектора

$$a_{ik} = \sum_x w(x) g_i(x) g_k(x), \quad b_i = \sum_x w(x) y(x) g_i(x)$$

Если Адиагональная, то  $a_{ii} p_i = b_i$  и  $p_{ii} = \frac{b_i}{a_{ii}}$  при условии что  $a_{ii} \neq 0$ . В общем случае  $P = A^{-1}B$ , при условии, что  $A$  имеет обратную.

Будем обозначать оценки параметров  $\hat{P}$ , а истинные значения  $P$ . Справедливы следующие теоремы, дающие представление о свойствах и качестве МНК-оценок.

1. Если модель  $f(x, P)$  и дисперсия  $d(x)$  верны, то оценки параметров несмещены. Это означает

$$\hat{E} \hat{P} = P$$

2. Матрица  $A^{-1}$  является матрицей ковариаций оценок параметров, ее диагональные элементы - дисперсиями оценок параметров.
3. В классе линейных несмещенных оценок мнк-оценки имеют минимальные дисперсии. (теорема Гаусса-Маркова);
4. Если все  $y(x_j)$  нормально распределены с  $(N(f(x, P), d(x)))$ , то мнк-оценки параметров имеют нормальное распределение  $N(P, A^{-1})$ , а величина  $S(P)^2$  имеет  $\chi_{m-n}^2$  распределение.

Доказательства этих теорем нетривиальны и для ознакомления с ними следует обратиться к специальной математической литературе.

**Задача проверки гипотез.** В регрессионном анализе в основном проверяется гипотеза: не противоречит ли регрессия  $y(x)$  модели  $f(x, P_{min})$  при найденных параметрах  $P_{min}$ . Исключительно удобный и при этом надежный инструмент для этого дает вышеприведенная теорема о том, что  $S(P_{min})^2$  имеет  $\chi_{m-n}^2$ -распределение при нормально распределенных ошибках регрессии  $y(x)$ . Но в действительности  $S(P_{min})^2$  будет иметь приближенно  $\chi_{m-n}^2$ -распределение и при всевозможных других ошибках, лишь бы выполнялось качественное условие, что эти ошибки являются результатом действия случайных факторов с равномерно малыми дисперсиями.

Другим условием правильности действия критерия является точное задание весов  $\frac{1}{d(x)}$ , т.е. дисперсии  $d(x)$ . В противном случае  $S(P_{min})^2$  не будет подчиняться  $\chi_{m-n}^2$ -распределению.

Очень важным дополнительным критерием является  $\chi_{m-n}^{2+}$ : отношение суммы тех членов в (3), где  $y(x_j) - f(x_j, P)$  положительно, к  $S(P_{min})^2$ . Математическое ожидание этого отношения равно  $\frac{1}{2}$ , а дисперсия равна приближенно  $\frac{0.85}{\sqrt{(m-n)}}$ .

Критерий  $\chi_{m-n}^{2+}$  устойчив к неточности задания весов  $\frac{1}{d(x)}$ , и является очень надежной характеристикой качества подгонки регрессии  $y(x)$  моделью  $f(x, P_{min})$ .

## ЛЕКЦИЯ 6. ОЦЕНИВАНИЕ ПАРАМЕТРОВ РЕГРЕССИЙ. НЕЛИНЕЙНЫЙ СЛУЧАЙ.

Как было сказано ранее, наиболее интересным и важным для практических целей является случай нелинейной параметризации регрессий. Нелинейная LS-оценка параметров находится из условия минимума по  $P$  выражения

$$S(P)^2 = \sum_{j=1}^m w(x_j)(y(x_j) - f(x_j, P))^2 \quad (4)$$

совершенно аналогично тому, как это было в линейном случае. Но система уравнений для нахождения решения будет нелинейной

$$\frac{\partial S^2}{\partial p_i} = \sum_{j=1}^m w(x_j)(y(x_j) - f(x_j, P)) \frac{\partial f}{\partial p_i} = 0, \quad (5)$$

и решение этой системы уже не будет столь простым как в линейном случае. Здесь используется прием линеаризации, заключающийся в следующем.

Предполагается, что известны грубо приближенные значения параметров  $P_0$ . Эти значения могут быть известны из различных, в частности, теоретических, соображений, или получены формальными методами, которые будут рассмотрены позже. Воспользовавшись гладкостью функций  $f$ , мы можем записать такое приближенное представление

$$f(x, P) \sim f(x, P_0) + \sum_{i=1}^n \frac{\partial f(x, P_0)}{\partial p_i} \Delta p_i$$

После подстановки его в уравнение (5) оно в матричной записи примет следующий вид:

$$B = A \Delta P,$$

где

$$B = \sum_{j=1}^m w(x_j)(y(x_j) - f_0) \frac{\partial f_0}{\partial p_i}, \quad A = \sum_{j=1}^m w(x_j) \frac{\partial f_0}{\partial p_i} \frac{\partial f_0}{\partial p_j},$$

$\Delta P$  - вектор  $\Delta p_i$ , а  $f_0 = f(x_j, P_0)$ .

Отсюда имеем

$$\Delta P = A^{-1} B$$

А из него мы можем получить уточненное значение вектора  $P$ :

$$P_1 = P_0 + \Delta P.$$

Его мы можем использовать в качестве нового начального приближения, и получить новое уточненное значение вектора параметров  $P_2$  и т.д. Другими словами мы имеем итерационный процесс

$$P_{k+1} = P_k + A(k)^{-1} B(k),$$

где  $k$  - номер итерации, а  $A(k)$  и  $B(k)$  - матрицы  $A$  и  $B$ , сосчитанные при  $P = P_k$ .

Если итерационный процесс сходится, мы получаем решение  $P_x$ . К сожалению, свойства линейных МНК-оценок не переносятся автоматически на нелинейные

оценки; в общем случае, они не будут несмещенными и их дисперсии не должны обладать свойством минимальности в том или ином смысле. Тем не менее, если выполняются такие (качественные) условия:

- 1/ адекватность модели данным;
- 2/ малость дисперсии ошибок  $e(x)$ ;
- 3/ правильная стратегия минимизации функционала (5,)

то нелинейные МНК-оценки будут обладать вышеприведенными свойствами линейных МНК-оценок по крайней мере в асимптотике (т.е. по мере роста точности данных), и соответственно, для оценки их точности можно также использовать диагональные элементы обратной МНК-матрицы  $A_1^{-1}$ .

Тем не менее в нелинейном случае необходимы дополнительные методы для оценки точности получаемых оценок параметров. Весьма эффективный метод для этого - обработка симулированных данных, моделирующих реальные данные. В нелинейном случае это обязательная рекомендация.

## ЛЕКЦИЯ 7. ОЦЕНИВАНИЕ ПАРАМЕТРОВ РЕГРЕССИЙ. РОБАСТНЫЕ ОЦЕНКИ.

МНК-оценки при всех их достоинствах обладают одним существенным недостатком - неустойчивостью к грубым неадекватностям данных и модели. Особенно этим страдают нелинейные МНК-оценки. Эта неустойчивость имеет принципиальный характер: она есть следствие повышенной чувствительности интегральной квадратичной метрики к резко неоднородным разностям значений функций, взаимное расстояние которых эта метрика оценивает: сумма квадратов разностей двух функций будет минимальной при более или менее одинаковых значениях этих разностей. Если же в какой-либо точке разность двух функций будет иметь большое значение, то при сближении этих функций сближающая процедура будет стремиться уменьшить именно эту разность даже если это приведет к увеличению других разностей.

Рассмотрим такой пример:

$$f(x_i) = 1, 1, 1, 1, \dots, 1, 101, 1, \dots, 1, \quad n = 100$$

Пусть мы хотим подобрать приближающую ее функцию-константу  $c(x) = C$  при  $x = 1, 2, \dots, 100$ . Если мы будем оценивать близость в квадратичной метрике, то  $c(x)$  будет ближайшей к  $f(x)$  при  $c(x) = \frac{\sum_{i=1}^n f_i}{n} = 2$ ; погрешность приближения исходной функции  $f(x)$  будет превосходить 100% в каждой точке кроме одной.

Но если оценивать близость в модульной метрике  $\rho$ , т.е. как

$$\rho(f, c) = \sum_{i=1}^n |f(x_i) - c(x_i)|,$$

то ближайшей функцией будет, как легко видеть,  $c(x) = 1$ ; т.е. погрешность приближения будет равна 0 во всех точках кроме одной, где  $f(x) = 101$ . Эта точка впрочем производит впечатление грубой ошибки (выброса), случайно попавшей в данные, и таким образом, приближение в модульной метрике проигнорировало эту ошибку, тогда как приближение в квадратичной метрике испытало ее пагубное влияние.

Такое свойство - игнорировать присутствие грубых ошибок (или, по крайней мере, слабо зависеть от них) называется робастностью (от английского robust).

Соотношение между МНК- и робастной оценками хорошо иллюстрирует сравнение арифметического среднего и медианы некоторого множества чисел. Если числа более-менее равномерны, среднее и медиана совпадают (точно или приближенно). Но для сильно разнородных чисел обе величины будут отличаться. Медиана тяготеет к "большинству" чисел, среднее ищет некий компромисс между числами.

Итак, подгонка в модульной метрике даст робастные оценки. К сожалению, превосходя МНК-оценки в отношении устойчивости к неадекватностям моделей регрессии, робастные оценки не обладают их эффективностью - они не обладают свойством несмещенности и дисперсии их (в случае "нормальных" данных) будут больше, чем у МНК-оценок. Другими словами, использование робастных оценок оправдано только в случаях, когда МНК-оценки заведомо имеют плохое качество.

Другой подход в построении робастных оценок состоит в использовании адаптивных весов  $w(x)$ , а именно включение в них элементов, позволяющих

автоматически корректировать дисперсии, увеличивая их там, где наиболее явно проявляются неадекватности между данными и моделью. Существует математический подход, позволяющий делать это так, что оценки параметров окажутся несмещенными и будут иметь меньшую дисперсию, чем аналогичные МНК-оценки.

Пусть  $P_0$  задано. Тогда способами введения адаптивных весов являются следующие

$$S(P)^2 = \sum_{j=1}^m w(x_j, P_0)(y(x_j) - f(x_j, P))^2 \Rightarrow \text{Minimum по } P, \quad \text{где}$$

$$w(x, P_0) = \begin{cases} 1/d(x) & \text{если } \|y(x_j) - f(x_j, P_0)\| \leq k\sqrt{d(x)}; \\ 1/|y(x_j) - f(x_j, P_0)|^2 & \text{иначе} \end{cases}$$

Или другой подход, позволяющий строить промежуточные алгоритмы между робастными и МНК:

$$w(x, P_0) = \begin{cases} 1/d(x) & \text{если } \|h(j)\| \leq const; \\ c(j)/d(x) & \text{иначе} \end{cases}$$

где

$$h(j) = \frac{y(x_j) - f(x_j, P_0)}{\sqrt{d(x)}}; \quad c(j) = \frac{1 + \beta}{(h(j)/const)^2 + \beta}.$$

При  $\beta = \infty$  мы имеем обычную LS-оценку; при малых  $\beta$  оценка будет робастной, причем случай  $\beta = 0$  будет соответствовать максимальной робастности.

## ЛЕКЦИЯ 8. МЕТОДЫ НЕЛИНЕЙНОЙ МИНИМИЗАЦИИ.

Решение уравнения (5) не может быть осуществлено непосредственно. Единственный способ, как уже говорилось, это итерационный процесс, заключающийся в нахождении начального вектора параметров  $P_0$  и итерационном уточнении его, т.е. получение  $P_i$ ,  $i = 1, 2, \dots$ , которые, в случае сходимости итерационного процесса, будут стремиться к  $P_x$ , который мы и будем рассматривать как решение. Ниже мы рассмотрим основные методы для решения уравнения (5), используемые на практике.

- Стохастический, или метод испытаний. Он требует задания не  $P_0$ , а области (которая может быть и бесконечной), содержащей решение. Метод может использоваться для нахождения как самого решения, которое он впрочем находит редко, так и более или менее приемлемого приближения к нему, которое можно использовать как  $P_0$ . Метод состоит в проверке различных векторов из заданной области, получаемых с помощью датчика псевдослучайных чисел, и запоминании того из них, который обеспечивает минимум (4).
- Градиентно-подобные методы. Начинают с  $P_0$ , точность которого может быть и не очень большой. Итерационный процесс строится по схеме

$$\Delta P_{k+1} = \lambda_k G(k) B(k), \quad k = \text{номер итерации};$$

$\lambda$  = шаг - скаляр из полу-интервала  $(0, 1]$ ;

Матрица  $G$  может быть произвольной, но обязательно положительно - определенной; в простейшем случае в качестве  $G$  можно взять диагональную единичную матрицу - в этом случае метод выглядит особенно просто

$$\Delta P_{k+1} = \lambda_k B(k).$$

Вектор  $B$  - вектор антиградиента функционала (5).

Данные методы очень просты в реализации, но обладают одним сильным недостатком: медленной и притом не монотонной сходимостью, причем чем ближе к минимуму, тем сходимость медленнее. Шаг приходится брать очень маленький, но и он не гарантирует монотонной сходимости.

- Метод Гаусс-Ньютона, являющийся упрощенным вариантом ньютоновского метода минимизации нелинейных функционалов. Начинает с  $P_0$

$$\Delta P_{k+1} = \lambda(\hat{R}A)^{-1}(k)B(k),$$

где  $\lambda$  - шаг в итерациях,  $0 < \lambda \leq 1$ , а  $k$  - номер итераций.

$\hat{R}$  - оператор демпфирования матрицы  $A$  - т.е. улучшения ее обусловленности. Плохо-обусловленная матрица или вообще необратима, или обращается с недопустимо большой погрешностью. Вектор  $B$  - это антиградиент функционала (4), он указывает в направления убывания его, и если обратная матрица указана неверно, она может повернуть спуск в ложном направлении. Напомним, что близость положительно - определенной матрицы

к вырожденной выражается в наличии у нее собственных значений, близких к нулю. Поэтому, простейшим примером демпфированной матрицы будет такой

$$\hat{R}(A) = (A + \alpha I)^{-1}$$

что соответствует увеличению собственных чисел на  $\alpha$ .

Шаг  $\lambda$  обязан ограничивать норму изменения параметров, так как линеаризованный функционал представляет исходный (4) лишь приближенно, и делать полные шаги (т.е. равные 1) можно не всегда. Особенно ограничивать шаги нужно в случае плохо - обусловленной МНК-матрицы.

Метод Гаусс-Ньютона является в настоящее время самым широко используемым для реализации МНК-процедур. Его главное достоинство - быстрая и при этом монотонная сходимость. Но он имеет и недостаток - требует хорошего начального приближения  $P_0$ .

Этот недостаток можно обойти, если использовать метод Гаусс-Ньютона в комбинации с сравнительно просто реализуемыми стохастическим или градиентным: первый для нахождения начального приближения  $P_0$ ; последние для уточнения  $P_0$ .

Практическая минимизация выдвигает следующие проблемы:

- как выбирать шаг и демпфер?
- когда обрывать процесс минимизации? по каким критериям? - ведь мы не знаем точного ответа, и должны руководствоваться косвенными соображениями.

К сожалению математически обоснованного ответа на эти вопросы в нелинейном случае не существует. Можно рекомендовать лишь одно универсальное средство: симуляцию нелинейных функционалов типа (4), их минимизацию и подбор при этом оптимальных шага и демпфера, а также и числа необходимых для хорошей минимизации итераций.

Критериями для оптимального обрыва итерационного процесса служат следующие.

- Норма приращения параметров  $\|\Delta P\|$  - если процесс минимизации монотонный и быстро сходящийся, то как только эта величина станет меньше некоторого уровня, процесс можно обрывать.
- Модуль приращения функционала  $\|S(P_{k+1})^2 - S(P_k)^2\|$  - если процесс минимизации монотонный и быстро сходящийся, то как только эта величина станет меньше некоторого уровня, процесс можно обрывать.
- Достижение  $S(P)^2$  - значения  $\sim \chi_{m-n}^2$ . Критерий правильно срабатывает, если веса и модель регрессии в (4) были заданы правильно.
- Число итераций - для немонотонных и медленно сходящихся процессов это единственно надежный критерий, но вообще максимальное число итераций, которое может быть использовано для минимизации, должно быть ограничено для любого процесса (во избежание бесконечного цикла).

- Положительная-определенность МНК-матрицы - можно использовать только в Ньютонском процессе. Он, правда, гарантирует лишь достижение локального минимума.
- Отношение  $\frac{S^{2+}}{S^2}$ . Должно быть равно 1/2 (идеальный случай).
- Графика совмещенных исходного распределения и подогнанного.

**Упрощение нелинейной параметризации.** Необходимость проведения различных манипуляций с мнк-матрицей ставят вопрос: если матрица  $A$  преобразуется (например, демпфируется), сохраняет ли итерационный мнк-процесс тем не менее свое свойство сходимости?

Запишем этот процесс так:

$$P_{k+1} = P_k - \beta_k M G(P_k) \quad (6)$$

Здесь скаляр  $\beta$  это шаг,  $G(P_k)$  - градиент функционала (4) и  $M$  - произвольная матрица.

Мы можем здесь доказать весьма общую теорему, что процесс типа (6) при соответствующей стратегии минимизации сходится с любой положительно-определенной матрицей  $M$ .

Качественно справедливость этого утверждения следует из того, что само направление минимизации (геометрически, направление спуска) определяется мнк-вектором, который является антиградиентом минимизируемого функционала (4), и, следовательно, если преобразованная матрица не поворачивает антиградиент на угол от 90 градусов и выше, достаточно малый шаг в направлении этого антиградиента ведет к уменьшению функционала. Положительно - определенная матрица как раз и осуществляет такой поворот, не превышающий 90 градусов.

Пусть  $D(P)$  является матрицей производных градиента  $G$  в точке  $P$ , а  $P_0$  точка, где (4) достигает своего минимума. Очевидно, что  $G(P_0) = 0$ .

**Теорема** Если, начиная с некоторого  $k$ ,  $P_k$  оказывается в малой окрестности  $P_0$ , где матрицы  $M, D$  являются положительно-определенными, то, выбирая соответствующие шаги  $\beta$ , мы можем добиться сходимости процесса (6) к  $P_0$ .

Доказательство. Мы имеем

$$\begin{aligned} \|P_{k+1} - P_0\| &= \|P_{k+1} - P_0 - \beta_k M G(P_k)\| = \|P_{k+1} - P_0 - \beta_k M (G(P_k) - G(P_0))\| = \\ &= \|P_{k+1} - P_0 - \beta_k M D(P_k)(P_k - P_0)\| = \|(I - \beta_k M D(P_k))(P_k - P_0)\| \end{aligned}$$

Здесь  $P_{tk} = P_0 + t(P_k - P_0)$ ,  $0 \leq t \leq 1$ , и мы использовали теорему о среднем значении, а  $I$  является единичной матрицей. Далее

$$\|(I - \beta_k M D)(P_k - P_0)\| \leq \|P_k - P_0\| \|I - \beta_k M D\|$$

$$\|I - \beta_k M D\| = \max_i |1 - \beta_k \mu_{ik}| = q_k,$$

где  $\mu_{ik}$  являются собственными значениями произведения  $M$  и  $D$ .

Очевидно, что  $q_k$  будет минимальным, если

$$1 - \beta_k \mu_{min} = -(1 - \beta_k \mu_{max})$$

или, если  $\beta_k = \frac{2}{\mu_x + \mu_n}$ , где  $\mu_n = \min_i \mu_{ik}$ , и  $\mu_x = \max_i \mu_{ik}$ .

В этом случае, поскольку матрицы  $M$  и  $D$  являются положительно-определенными, мы имеем

$$\|P_{k+1} - P_0\| \leq q_k \|P_k - P_0\|, \quad q_k \leq 1.$$

Мы можем показать по индукции, что

$$\|P_{k+1} - P_0\| \leq \prod_{i=1}^{k+1} q_i \|P_1 - P_0\|, \quad q_i \leq 1.$$

Если параметризация является правильной, то  $\mu_n > 0$ , [92] и

$$\|P_{k+1} - P_0\| \leq \left(\frac{\mu_x - \mu_n}{\mu_x + \mu_n}\right)^{k+1} \|P_1 - P_0\|,$$

и процесс (6) сходится к  $P_0$ . Ergo.

**Примечание 1.** Если  $M = D$ , т.е. ,  $D = A^{-1}$ , сходимость будет наиболее быстрой. Плохая обусловленность и демпфирование замедляют ее, но они не останавливают ее (при достаточно малых  $\beta_k$ ).

**Примечание 2.** Числа  $\beta$  (шаг) и  $\lambda$  (демпирующий фактор) определяют стратегию минимизации. Если матрица хорошо обусловлена, можно использовать полный шаг ( $\beta = 1$ ) и не использовать демпфирование ( $\lambda = 0$ ); но в случае плохой обусловленности надо брать большие  $\lambda$  и меньшие  $\beta$ .

О влиянии демпфирования на матрицу ошибок качественно можно сказать лишь одно: если  $\lambda$  значительно отличается от нуля, элементы матрицы, обратной к демпфированной, могут отличаться очень существенно от элементов матрицы, обратной к исходной, и, следовательно, матрица, вычисленная на последней итерации, будет смещенной оценкой матрицы ковариаций, и смещение будет тем больше, чем хуже обусловленность исходной матрицы.

**Примечание 3.** Малостью шага можно воспользоваться также для замены антиградиента его приближением (например, более простым для вычислений, или, если необходимо и возможно использовать более простую параметризацию выражения, моделирующего данные). Естественное условие для того - малость отличий по норме истинного антиградиента от его приближения и, разумеется, положительная определенность мнк-матрицы.

Иногда точную параметризацию трудно осуществить, например, если физическими параметрами являются как элементы некоторой прямой матрицы, так и элементы матрицы, обратной к ней, или ее детерминант; при большом числе параметров задача является исключительно сложной для реализации.

Для прояснения идеи упрощения параметризации рассмотрим следующий пример. Пусть  $F(f(P), g(Q))$  - сложное выражение, зависящее от векторов параметров  $P$  и  $Q$ , и пусть аналитическое вычисление производных  $\frac{\partial g}{\partial q_i}$  является очень сложным, а численное ведет к недопустимо большому времени работы программы. Если вектор  $Q$  независим от вектора  $P$  и мы зафиксируем  $Q$ , положив  $Q = Q_0$ , где  $Q_0$  - либо априорная оценка  $Q$ , либо его значение на предыдущей итерации, мы получим просто модель регрессии, не вполне ей адекватную, и мнк-оценки параметров вектора  $P$  будут смещенными, т.е. иметь систематическую ошибку. Вектор  $Q$  естественно при этом меняться не будет. Но если среди компонент  $Q$  будут зависящие от компонент

$P$ , а тем более совпадающие, например, некоторые  $p_i = q_i$ , то тогда антиградиент от  $F$  по таким параметрам  $p_i$  будет отличаться от истинного на величину

$$\Delta B = \sum_{j=1}^m w(x_j)(y(x_j) - F(f, g)) \frac{\partial F}{\partial g} \frac{\partial g_0}{\partial q_i},$$

а мнк-матрица соответственно на  $\Delta A$  за счет неточных производных  $\frac{\partial f}{\partial p_i}$ , и "подозрительным" возмущением приращения вектора параметров  $P$  будет величина  $(A + \Delta A)\Delta B$ .

Положительная определенность матрицы  $A + \Delta A$  следует из ее построения; следовательно, если норма  $\Delta B$  мала, то при достаточно малом шаге в итерациях изменения вектора параметров  $P$  будут происходить в направлении убывания мнк-функционала.

## ЛЕКЦИЯ 9. АПРИОРНАЯ ИНФОРМАЦИЯ И УСЛОВНАЯ МИНИМИЗАЦИЯ.

Задача формулируется следующим образом: найти минимум функционала

$$S(P)^2 = \sum_{j=1}^m w(x_j)(y(x_j) - f(x_j, P))^2 \quad (7)$$

при условиях

$$g(P) \leq 0. \quad (8)$$

Условия могут быть заданы различными способами. Простейшие условия - это двусторонние ограничения на параметры или попарные линейные связи между ними:

$$p_{il} \leq p_i \leq p_{iu}; \quad p_j = k_j p_i + b_j. \quad \text{при } i = 1, 2, \dots, n \text{ и } j > i.$$

Попарные зависимости фактически сокращают число независимых параметров. Поэтому, их учет осуществляется преобразованием МНК-матрицы: если  $p_j = k_j p_i + b_j$ , то  $i$ -ые строка и столбец вычеркиваются на  $i$ -м месте и, умноженные на  $k_j$  и сложенные с  $b_j$ , добавляются к  $j$ -м строке и столбцу.

Двусторонние ограничения учитываются с помощью проектирования текущего вектора  $P$  на эти ограничения. Оператор проектирования выглядит так:

$$\hat{p}_i = \begin{cases} p_i & \text{если } p_{il} \leq p_i \leq p_{iu}; \\ p_{il} & \text{если } p_i < p_{il}; \\ p_{iu} & \text{если } p_i > p_{iu}; \end{cases}$$

где  $\hat{p}_i$  - компоненты проекции вектора  $P$  на ограничения. Доказано, что если точка минимума (7) лежит внутри симплекса  $[p_{il}, p_{iu}]$ ,  $i = 1, 2, \dots, n$ , проекции сходятся к ней.

Метод проектирования применим в случае, если (8) представляют собой выпуклые связи, т.е. выделяют в пространстве параметров выпуклую область, - такую, что все точки, лежащие на отрезке, соединяющем любые две точки из этой области, тоже принадлежат ей. Выпуклыми будут все области, ограничиваемые линейными формами и положительно - определенными квадратичными, и многие другие. Проекции на выпуклое множество (8) сходятся (при правильной стратегии минимизации) к точке минимума функционала (7).

Построение проектора часто оказывается весьма нетривиальной задачей. Поэтому, широкое применение нашли другие методы, прежде всего метод Лагранжа и метод штрафных функций.

Идея метода Лагранжа состоит в замене функционала (7) на

$$S(P)^2 = \sum_{j=1}^m w(x_j)(y(x_j) - f(x_j, P))^2 + \lambda g(P). \quad (9)$$

с дополнительной неизвестной переменной  $\lambda$  и минимизации (9) при условии (8).

Метод штрафных функций тоже состоит в минимизации функционала (9), но  $\lambda g(P)$  играет роль "штрафа", который добавляется к функционалу (7) тем в большем размере, чем больше  $g(P)$  превосходит нуль; если  $g(P)$  отрицательна, то член  $\lambda g(P)$  не добавляется.

Обычно начинают с больших  $\lambda$ ; затем по мере успеха ее уменьшают. Эффективность метода не стопроцентна, но часто бывает приемлемой.

## ЛЕКЦИЯ 10. ФИЛЬТРАЦИЯ.

Рассмотрим предварительно важное для задач фильтрации ДИСКРЕТНОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ (DFT). Пусть заданная функция  $f(t)$  непрерывна и абсолютно интегрируема в области определения. Тогда можно записать преобразование Фурье, прямое:

$$g(\lambda) = \int_{-\infty}^{\infty} f(t)e^{-i\lambda t} dt$$

и обратное:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(\lambda)e^{i\lambda t} d\lambda$$

Отметим следующее свойство этих операций  $\hat{F}(f'(t) = i\lambda\hat{F}f(t))$ .

Далее для интегрируемой функции  $R(x)$  может быть определена операция конволюции (иначе свертки)

$$y(x) = R(x) * f(t) = \int R(t-x)f(t)dt$$

В этом случае справедливо

$$\hat{F}y = \hat{F}R \cdot \hat{F}f$$

и если известны  $y(x)$  и  $R(x)$ , то  $f(t)$  можно определить формулой

$$f(t) = \hat{F}^{-1}(\hat{F}R \cdot \hat{F}y)$$

Преобразование Фурье играет в математике огромную роль, но для практических приложений часто наиболее подходящими оказываются дискретные аналоги этих операций.

Пусть задана гистограмма  $s(k), k = 0, \dots, m$ ; тогда прямое дискретное преобразование Фурье  $s(v), v = 0, \dots, m$  и обратное  $s(k), k = 0, \dots, m$  определяются следующим образом:

$$s(v) = \frac{1}{m+1} \sum_{k=0}^m s(k)e^{-qkv}, \quad v = 0, \dots, m \quad (10)$$

$$s(k) = \sum_{v=0}^m s(v)e^{qkv}, \quad k = 0, \dots, m; \quad (11)$$

где  $q = \frac{2\pi i}{m+1}$  Свойства

- $s(v+m-1) = s(v)^*$  Отсюда: если  $m$  четное,  $s(m/2+1) = 0$
- $\hat{F}(s(k+1) - s(k)) = (e^{qkv} - 1)\hat{F}s(k)$  если  $s(0) = s(m) = 0$

Примеры.

- $s(k) = 1, k = 0, \dots, m;$

$$s(v) = \frac{1}{m+1} \sum_{k=0}^m e^{-qkv} = 1 + e^{-qv} + e^{-q2v} + \dots + e^{-qmv}, \quad \text{откуда}$$

$$s(v) = \begin{cases} 1 & \text{если } v = 0; \\ \frac{1}{m+1} \frac{1-e^{-qv(m+1)}}{1-e^{-qv}} = 0 & \text{иначе} \end{cases}$$

- $s(k) = e^{qkn}$ ,  $k = 0, \dots, m$

$$s(v) = \frac{1}{m+1} \sum_{k=0}^m e^{qkn} e^{-qkv} = 1 + e^{-q(v-n)} + \dots + e^{-qm(v-n)} = \frac{1 - e^{-q(v-n)(m+1)}}{(m+1)(1 - e^{-q(v-n)})}$$

$$s(v) = \begin{cases} 1 & \text{если } v = n; \\ 0 & \text{иначе} \end{cases}$$

Наглядный смысл выражения  $s(v)$  состоит в следующем: его модуль дает разложение гистограммы  $s(k)$  на  $m/2$  частот, которые являются кратными максимальной частоте (главной, или частоте Найквиста)  $2\pi/(m+1)$ , т.е. характеризует частотный состав гистограммы  $s(k)$ . Поэтому, преобразование Фурье используется в обработке данных в тех задачах, где требуется воздействие на их частотные характеристики.

- Поиск периодичностей. Идея метода видна из последнего примера. Модуль преобразования Фурье функций  $\sin(w_i t)$  и  $\cos(w_i t)$  имеет максимум в точках  $w_i$ , следовательно, гистограмма с максимумами в точках  $w$  будет содержать периодические компоненты на частотах  $w \cdot 2\pi/(m+1)$ .
- Сглаживание гистограмм. Удаление (или подавление) членов  $s(v)$  в точках  $v_1, v_2, \dots$  и  $v_{m-v_1}, v_{m-v_2}, \dots$  и последующее применение обратного преобразования к остатку соответствует удалению (подавлению) в исходной гистограмме частот  $v_1, v_2, \dots$

Для гистограмм можно ввести дискретный аналог конволюции или свертки. Свертку можно получить следующим образом. Пусть даны гистограммы  $s_i$ ,  $R_i$ , и  $i = 0, 1, \dots, 2n+1$ . Вне этого интервала индексов обе гистограммы равны нулю. Функцию  $R_i$  будем считать свертывающей, например, это может быть весовая функция или функция (плотность) распределения с максимумом в точке  $i = n+1$ .

Тогда свертка этих функций (обозначим ее  $y_j$ ) будет выглядеть так

$$y_j = \sum_{i=-n}^n s_{j+i} R_{n+i}, \quad j = 0, 1, 2, \dots, 2n+1$$

Ее ДПФ будет очевидно равно

$$y_v = \frac{1}{2n+2} \sum_{k=0}^{2n+2} \left( \sum_{i=-n}^n s_{k+i} R_{n+i} \right) e^{-qkv} \quad (12)$$

Каждый член (12) можно записать как

$$s_{k+i} e^{-q(k+i)v} R_{n+i} e^{-q(n+i)v} e^{q(n)v}$$

Величину  $e^{q(n)v}$  можно вынести за знак (12), и тогда каждый член двойной суммы в (12) будет попарным произведением членов ДПФ от  $s_{k+i}$  и  $R_{n+i}$ . Другими словами, применение ДПФ к свертке отличается от непрерывного случая, а именно:

$$\hat{D}F y = \frac{1}{2n+2} \hat{D}F R \cdot \hat{D}F s \cdot e^{q(n)v}$$

Обобщением и модификацией Фурье преобразования является wavelet анализ. Преобразование Фурье имеет ряд ограничений:

- частотный состав функции (гистограммы) предполагается стационарным; иначе, сглаживание, например, даст подавление неких "средних" верхних частот, что может вести к сильным локальным искажениям результирующей функции.
- сами частоты и их состав фиксированы, что не всегда желательно с точки зрения анализа;
- форма частот берется простейшей - синусоидальной, что сильно уменьшает аналитические возможности формализма Фурье преобразований.

Вейвлиты и призваны преодолеть эти недостатки.

Одномерное вейвлет-преобразование функции  $f(x)$  имеет вид:

$$y_\psi(a, b) = C \int \psi((b-x)/a) f(x) dx \quad (13)$$

где  $C$  - нормирующая константа, функции  $\psi(x, a, b)$  - семейство "вейвлитов-обобщенных аналогов синусоидальных частот  $\exp(ixw)$  с подгоняемыми параметрами  $a, b$ , что позволяет сдвигать "частоты" уже неравномерно и с переменным масштабом.

Диапазон используемых вейвлитов исключительно велик - от, простейших ступенчатых функций до, например, производных от функции Гаусса

$$\psi_n(x) = (-1)^n \frac{d^n}{dx^n} \exp(-x^2/2), \quad n = 0, 1, 2, \dots \quad (14)$$

Вейвлит-преобразование (13) дает нам разложение функции  $f(x)$  по вейвлитам  $\psi_n(x)$ , и дальнейшие действия с этим разложением могут быть такими же, как и классическом Фурье-анализе. Но здесь при правильном выборе базиса вейвлитных функций качество результата может быть значительно лучше, чем в Фурье-анализе. К гистограммам следует применять дискретный аналог (13). Здесь математический формализм сложнее, поскольку добиться симметрической обратимости (13) такой, как в классическом Фурье-анализе, удастся не для всякого базиса вейвлитов. Но для ограниченных целей (задач сглаживания, например) особых требований к этому базису нет. Гистограмма  $h(x_k)$ ,  $k = 0, 1, 2, \dots, m$  будет иметь не более, чем  $m/2$  независимых "частот"  $\psi_n(x)$ , но выбираемых из более широкого спектра всевозможных.

Вейвлиты весьма подробно изложены в обширной литературе, и на сайтах Интернета, где доступ к информации о вейвлитах можно получить, набрав в [www.google.ru](http://www.google.ru) "вейвлет анализ".

**Задачи фильтрации данных.** Задача сглаживания гистограмм - это частный случай более общей проблемы фильтрации данных. Фильтром называется преобразование  $\hat{F}f(k) \rightarrow s(l)$ ,  $k \in K$ ,  $l \in L$ , переводящее гистограмму  $f(k)$  в гистограмму  $s(l)$ . Цели ее: усиление, подавление или выравнивание компонент.

Кроме рассмотренного метода сглаживания с помощью преобразования Фурье мы кратко рассмотрим здесь наиболее эффективные методы - это аппроксимационные фильтры и фильтры на основе сплайнов.

Аппроксимационные фильтры:  $n$  чисел  $f_k$  подгоняются ортогональными полиномами

$P_l(k)$ .

Пример: каждая тройка  $f(k-1), f(k), f(k+1)$  или  $f_{-1}, f_0, f_1$  подгоняется  $P_1(k) = ak + b$ ,  $k = -1, 0, 1$ . Так как полиномиальные члены ортогональны, мы имеем (все веса равны 1)

$$\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} \sum f_k \\ \sum f_k k \end{pmatrix}$$

Сглаживание дается  $f_0 = b$ . Поэтому,  $b = \frac{1}{3}(f_{-1} + f_0 + f_1)$ .

Другой пример: каждая пятерка  $f(k-2), f(k-1), f(k), f(k+1), f(k+2)$  или  $f_{-2}, f_{-1}, f_0, f_1, f_2$  подгоняется  $P_2(k) = a(k^2 - 2) + bk + c$ ,  $k = -2, -1, 0, 1, 2$ . Так как полиномиальные члены ортогональны, мы имеем

$$\begin{pmatrix} 5 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 14 \end{pmatrix} \begin{pmatrix} c \\ b \\ a \end{pmatrix} = \begin{pmatrix} \sum f_k \\ \sum f_k k \\ \sum f_k (k^2 - 2) \end{pmatrix}$$

Сглаживание дается  $f_0 = -2a + c$ . Поэтому,  $c = \frac{1}{5} \sum_{k=-2}^2 f_k$  и  $a = \frac{1}{14} \sum_{k=-2}^2 f_k (k^2 - 2)$   
и

$$f_0 = \frac{1}{35}(-3f_{-2} + 12f_{-1} + 17f_0 + 12f_1 - 3f_2)$$

**СПЛАЙНОВЫЕ ФИЛЬТРЫ.** Непрерывный фильтр сплайнового типа строится как задача минимизации функционала

$$\int_{t_1, t_2} F(f(t), f'(t), f''(t)) dt + \beta r(f(t), y(t)) \quad (15)$$

при заданном  $\beta$  и определенных граничных условиях на искомую функцию  $f(t)$ . Здесь  $F$  - критерий оценки требуемого качества функции  $f(t)$ , а  $r$  произвольная метрика. Смысл этой процедуры - найти функцию  $f(t)$ , близкую к исходной  $y(t)$  в смысле метрики  $r$ , и в то же время оптимальную в смысле критерия  $F$  при данном  $\beta$ .

Для численного решения задачи требуется дискретизация (15).

**ПРИМЕРЫ КРИТЕРИЕВ ОЦЕНКИ КАЧЕСТВА ФУНКЦИЙ.** Рассмотрим конкретные примеры, а именно

- меры осцилляций функции;
- меры изменчивости функций.

Тогда для сглаживания данных надо построить функцию, близкую к данным, но имеющую минимальную меру осцилляций; а для построения функции, огибающей данные снизу, взять такую, которая близка к ним снизу и имеет минимальную меру изменений. Фильтры такого типа используют не только частотные свойства данных, но также и амплитудные, что делает эти фильтры более гибкими и уменьшает искажение информационных компонент по сравнению с чисто частотными фильтрами.

Рассмотрим непрерывный аналог введенных мер. Визуальному впечатлению быстрых изменений направлений осцилляций кривой соответствует быстрое изменение знака и нормы ее 1-ой производной. С другой стороны, сравнивая две части кривой, где 1-ая производная изменяется одинаково, мы склонны считать, что

более короткая часть колеблется быстрее.

Можно изложить это иначе. Мы можем рассматривать длинную часть кривой как резонанс, и поэтому, считать, что она колеблется медленнее. В обоих случаях совершенно естественно считать отношение 1-ой производной к длине кривой в интервале  $(x, x + dx)$  как меру осцилляций кривой в этом интервале

$$c'(x) = \frac{f'(x + dx) - f'(x)}{\sqrt{(1 + f'(x)^2)dx}}.$$

Когда  $dx \rightarrow 0$  мы получим плотность этой меры

$$c(x) = \lim_{dx \rightarrow 0} c'(x) = \lim \frac{f'(x + dx) - f'(x)}{dx} \frac{1}{\sqrt{1 + f'(x)^2}} = \frac{f''(x)}{\sqrt{1 + f'(x)^2}}.$$

Интегральная мера осцилляций функции в интервале  $(x_1, x_2)$  равна

$$\mu(x_1, x_2) = \int_{x_1}^{x_2} c(x)^2 dx = \int_{x_1}^{x_2} \frac{f''(x)^2}{1 + f'(x)^2} dx. \quad (16)$$

Аналогично, мера изменчивости функций может быть определена как отношение изменений функции и аргумента

$$df(x) = (f(x + dx) - f(x))/dx.$$

Точечная мера совпадает с 1-ой производной, когда  $dx \rightarrow 0$

$$d(x) = f'(x).$$

Интегральная мера изменчивости определяется следующим образом

$$\delta(x_1, x_2) = \int_{x_1}^{x_2} d(x)^2 dx.$$

Выражение (16) допускает следующее обобщение:

- для комплексно-значных функций

$$\mu(a, b) = \int_a^b \frac{f''(x)^* f''(x)}{1 + f'(x)^* f'(x)} dx. \quad (17)$$

где  $(f'')^*$  и  $(f')^*$  являются сопряженными функциями к  $f''$  и  $f'$ .

- Произвольная степень знаменателя

$$\mu(x) = \frac{f''(x)}{\sqrt[1]{1 + f'(x)^2}} \cdot \mu(a, b) = \int_a^b \mu(x)^2 dx \quad (18)$$

Рассмотрим множество гармонических функций ( $\exp(-ikx)$ ),  $k = 0, 1, \dots, \infty$ , и определим их интегральные меры осцилляций в интервале  $(0, 2\pi)$  с использованием (18)

$$\mu(0, 2\pi) = \int k^4/(1+k^2)^b dx = 2\pi k^4/(1+k^2)^b.$$

Они приблизительно равны

$$c \cdot k^{4-2b}, \quad (19)$$

т.е. мера основных гармонических функций  $\exp(-ikx)$  при  $b = 1$  растет пропорционально квадрату частот  $k$ . Этот закон может быть изменен, если использовать выражение (18) с  $b$  не равным единице как меру осцилляций.

В частности, кривизна  $f''/(1+f'^2)^{1.5}$  является линейной функцией гармонических частот.

Рассмотрим далее множество функций, зависящих от параметров следующим образом

$$f(x, a, p, w) = a \cdot m((x-p)/w).$$

Оно включает, например, функции экспоненциального типа. Их интегральная мера осцилляций будет равна

$$((a^2 m''^2/w^4)/(1+(a^2 m'^2/w^2)^b) \simeq a^{2b-2}/w^{2b-4}$$

Отсюда можно видеть, что при  $b = 1$ , функция  $f(x)$  мера осцилляций инвариантна по отношению к "высоте" фигуры, контур которой описывается функцией  $f(x)$ , а при  $b = 2$ , по отношению к "ширине".

**АЛГОРИТМ ФИЛЬТРАЦИИ СПЛАЙНОВОГО ТИПА.** Начнем с задачи построения играющего очень важную роль в практических приложениях фильтра для низкочастотной огибающей снизу  $f(t)$  для функции  $y(t)$  на интервале  $(t_1, t_2)$ .

Для дискретизации (15) рассмотрим случай взвешенной квадратичной метрики и критерия  $F(f(t)) = f'(t)^2$ . Имеем

$$\sum_{i=2}^{m-1} f'(i)^2 + \beta q \sum_{i=1}^m w(i)(f(i) - y(i))^2 \quad (20)$$

при граничных условиях  $f(1) = y(1)$ ,  $f(m) = y(m)$  и  $q = \frac{1}{\sum_i w(i)}$ .

Веса  $w(i)$  должны быть обратно пропорциональны

1. некоторой степени  $y(i)$  или  $y''(i)$  на 1-ом шаге;
2. некоторой степени от разности  $y(i)$  и текущего  $f(i)$  на следующих шагах; эти веса  $w(i)$  можно строить и робастным способом, например, так

$$w(i) = \begin{cases} 1 & \text{если } |f_0(i) - y(i)| \leq c\sigma(i); \\ 1/(1+|f_0(i) - y(i)|) & \text{иначе} \end{cases}$$

где  $c$  - константа, например  $c = 3$ .

Для простоты записи нормируем веса:  $w(i) \rightarrow w(i)q$ . Для нахождения минимума функционала (20) можно применять т.н. метод 'прогонки'. Его процедура является итеративной. Уравнение Эйлера для минимизации (20) выглядит так

$$f(i+1) - 2f(i) + f(i-1) - \beta w(i)f(i) = -\beta w(i)y(i), \quad i = 2, \dots, m-1,$$

при  $g(1) = f(1)$ ,  $g(m) = f(m)$ .

Эти уравнения могут быть решены устойчивым способом следующим образом. Представим

$$f(i) = c(i)f(i+1) + h(i), \quad c(1) = 0, h(1) = y(0), i = 1, \dots, m-1,$$

где коэффициенты  $c(i)$ ,  $h(i)$ ,  $i > 1$  являются пока неизвестными. Подставив  $f(i)$  в уравнение Эйлера, мы получим

$$f(i+1) - 2g(i) + c(i-1)f(i) + h(i-1) - \beta w(i)f(i) = -\beta w(i)y(i).$$

Отсюда мы получаем формулу для вычисления коэффициентов  $c(i)$ ,  $h(i)$ :

$$c(i) = 1/(2 + \beta w(i) - c(i-1)), \quad h(i) = c(i)(\beta w(i)f(i) + h(i-1)),$$

при начальных  $c(1) = 0, h(1) = y(1)$ , используя которые, мы находим последовательно нужное решение  $f(i)$ .

Следующая важная задача - сглаживание функций. Для заданной функции  $y(i)$  ее сглаженный аналог  $f(i)$  находится из условия минимума выражения

$$\sum_{i=2}^{m-1} f''(i)^2 e(i) + \alpha \sum_{i=1}^m w(i)(y(i) - f(i))^2 \quad (21)$$

при граничных условиях  $f(1) = y(1), f(m) = y(m), f''(1) = 0, f''(m) = 0$ .

Т.е. в данном случае критерий  $F(f''(t)) = f''(t)^2 e(t)$ ;  $e(t)$  и  $w(t)$  - тем или иным способом выбранные веса. Например,  $e(i) = 1/(1 + y'(i)^2)$ , а  $w(i) = 1/\sigma(i)^2$ , или  $w(i) = 1$ .

Уравнение Эйлера для минимизации (21) выглядит сложнее, так как приводит к 5-точечным разностным схемам. Для его решения приходится прибегать уже к матричной 'прогонке'. Уравнения для минимизации (21) выглядят так

$$e(i+1)d''(i+1) - 2e(i)d''(i) + e(i-1)d''(i-1) + \alpha \cdot w(i)f(i) = \alpha \cdot w(i)y(i). \quad (22)$$

где  $d(i) = f''(i)$ . Граничные условия имеют вид

$$f(1) = y(1); f(m) = y(m); f''(1) = 0; f''(m) = 0;$$

Построим формулы для матричного аналога метода 'прогонки' для решения этих уравнений. Мы имеем

$$P(i) = A_i \cdot P(i+1) + B(i) \quad (23)$$

где  $P(i)$  является вектором  $(f(i), f''(i))$ , а  $A_i$  - матрицей размера  $2 \cdot 2$ , и  $B(i)$  - вектором размера 2. Мы имеем

$$A_i(1,1) = u_i(1,1); A_i(1,2) = u_i(1,2);$$

$$A_i(2,1) = u_i(2,1); A_i(2,2) = u_i(2,2);$$

$$A_1(1,1) = A_1(1,2) = A_1(2,1) = A_1(2,2) = 0;$$

$$B_1(i) = u_i(1,1)B_1(i-1) - u_i(1,2)(\alpha \cdot w(i)f(i) - e(i-1)B_2(i-1))/e(i+1);$$

$$B_2(i) = u_i(2, 1)B_1(i - 1) - u_i(2, 2)(\alpha \cdot w(i)f(i) - e(i - 1)B_2(i - 1))/e(i + 1);$$

$$B_1(1) = f(1); B_2(1) = 0;$$

где  $u_i(k, j)$  - элементы матрицы, обратной к

$$\begin{pmatrix} 2 - A_{i-1}(1, 1) & 1 - A_{i-1}(1, 2) \\ -(e(i - 1)A_{i-1}(2, 1) + \alpha \cdot w(i))/e(i + 1) & (2e(i) - e(i - 1)A_{i-1}(2, 2))/e(i + 1) \end{pmatrix}.$$

Используя эти формулы, сначала вычисляем коэффициенты прямой процедуры, после чего решение  $(f(i), f''(i))$  получается из (23).

В заключение можно кратко упомянуть пико- и периодо-усиливающие фильтры, например: разностный оператор как средство усиления периодичностей, или фильтр типа

$$f(x) = \max\left(0, -\frac{r''(x)}{\sqrt{1 + r'(x)^2}}\right)$$

как средство усиления пико-образных компонент.

## ЛЕКЦИЯ 11. ОСНОВНЫЕ ЗАДАЧИ АНАЛИЗА ЭКСПЕРИМЕНТАЛЬНЫХ СПЕКТРОВ.

Важнейшими при анализе экспериментальных данных в физике твердого тела являются следующие задачи

- Формонезависимый многоспектрный и многофазный Ритвельд-анализ (в первую очередь, для RTOF-спектров).
- Определение симметричной структуры порошка (Powder Match).
- Автоматический 3-мерный синтез Фурье.
- Индексация многофазных порошков.
- Формонезависимая подгонка одно- и многомерных пиков.
- Различные типы фильтрации данных, и т.д.

Весьма универсальной и удобной как для целей использования, так и для демонстрации является

### ПРОГРАММА VMRIA.

Программа VMRIA (Visual Multiphase Rietveld Analysis) ориентирована на самый широкий класс нейтронно-дифракционных данных от многофазных поликристаллов, как времяпролетных, так и угловых; как прямых, так и Фурье (Reversed Time of Flight) и открыта для всевозможных модификаций и расширений; способна анализировать несколько спектров многофазных структур одновременно (например, измеренных при различных углах); независима от формы пика и может работать как с регулярной формой (Гауссиан, Лоренциан, и т.д.), так и с наиболее экзотичной, заданной, если нужно, графически или таблицей значений.

Для обработки многомерных спектров используется программа VDOMUS - аналог программы VMRIA, схожий с ней по интерфейсу пользователя. Этими программами можно воспользоваться для практического усвоения материала, изложенного в пособии, т.к. при решении вышеназванных задач использовались методы, описанные в пособии. Обе программы написаны на языке программирования DELPHI и работают в операционной среде Windows XX.

**Задача анализа пиков.** Задача имеет 2 фазы:

1. качественная, - определение числа пиков в спектре и их приближенных центров (положений);
2. количественная, - построение оценок параметров этих пиков.

Формально метод решения задачи (алгоритм UPEAK) заключается в следующем.

Одномерный спектр записывается так

$$s(x) = \sum_{i=1}^n f_i(x) + b(x) + e(x)$$

где  $f_i$  - полезная компонента; это может быть пик, или экспонента, или другая функция, отличная от нуля только в локальной области;  $b(x)$  - функция фона,  $e(x)$  - ошибка спектра. Переменная  $x$  обычно называется каналом, и измеряется в условных

целочисленных единицах  $x = 1, 2, \dots, m$ .

Для каждого типа компонент  $f_i$  строится модель  $m_i(x)$ , также и  $m_b(x)$ . Эти модели могут быть аналитическими функциями или экспериментальными гистограммами, которые параметризуются следующим образом

$$f_i(x, A, P, W, C) = A \cdot m_i((x - P)/(C \cdot x + W))$$

Параметры  $A, P, W, C$  имеют ясный физический смысл:

1.  $A$  - амплитуда (или площадь) компоненты;
2.  $P$  - положение (центр) компоненты;
3.  $W$  - полуширина компоненты;
4.  $C$  - коэффициент линейной зависимости полуширины от канала.

Далее начинается поиск пиков, включающий в себя следующие шаги.

1. Спектр приводится к стандартному масштабу по амплитуде умножением на соответствующий скаляр, и полуширине пиков компрессией по каналам так, чтобы амплитуда не превосходила 1000, а полуширина 6 каналов.

2. Затем осуществляются сглаживание спектра, построение низкочастотной огибающей снизу и вычитание ее из спектра.

3. Далее производится поиск вершин пиков по критерию квази-кривизны: в каждой точке вычисляются величины  $s(x)''/\sqrt{1 + s(x)'^2}$  и точки локальных максимумов этой величины  $x_j$  при условии, что  $s(x_j) > err(x_j)$ , где  $err(x)$  - максимальный уровень ошибки (с учетом фона) спектра в точке  $x$ , принимаются за приближенные положения пиков.

4. Далее производится анализ найденных пиков по критериям чувствительности (отсев слабых пиков) и разрешения (слияние слишком близких).

5. На основе найденных пиков строится расчетный спектр, подгоняемый по амплитуде к  $s(x)$ , который затем вычитается из  $s(x)$ .

6. Остаток анализируется на присутствие скрытых наиболее сильных пиков, которые, если обнаружены, добавляются к списку ранее найденных.

Затем выполняется та или иная процедура аналитической подгонки параметрической модели спектра к  $s(x)$  в метрике наименьших квадратов или робастной, в результате чего мы получаем оценки параметров  $A, P, W, C$  и их погрешности для каждого пика. Дифференцирование по параметрам, необходимое для применения методов подгонки, осуществляется следующим образом ( $z = (x - P)/(C \cdot x + W)$ ):

$$\frac{\partial f_i}{\partial A} = m_i((x - P)/(C \cdot x + W))$$

$$\frac{\partial f_i}{\partial P} = -A \cdot \frac{\partial m_i(z)}{\partial x} / (C \cdot x + W)$$

$$\frac{\partial f_i}{\partial W} = -A \cdot \frac{\partial m_i(z)}{\partial x} z / (C \cdot x + W)$$

$$\frac{\partial f_i}{\partial C} = -A \cdot \frac{\partial m_i(z)}{\partial x} z \cdot x / (C \cdot x + W)$$

**Задача анализа пиков в многомерных спектрах. Алгоритм DOMUS.**  
 Многомерный спектр  $s(X)$  записывается следующим образом

$$s(X) = \sum_{i=1}^n f_i(X) + b(X) + e(X)$$

где  $f_i$  -  $i$ -ая полезная компонента спектра (многомерный пик),  $b$  - многомерный фон,  $e$  - многомерная ошибка,  $X$  - вектор каналов.

Каждый пик  $f_i$  имеет многомерную модель  $m(X)$ ,

$$A_i m(M_i \cdot Z_i), \text{ где } Z_i = \frac{x_k - P_{ik}}{C_{ik}x_k + W_{ik}},$$

$A_i$  - амплитуда  $i$ -го пика,  $x_k$  -  $k$ -ая компонента вектора  $X$ , а  $P_{ik}, C_{ik}, W_{ik}$  - проекции параметров положения пика и коэффициентов зависимости его полуширины от канала на  $k$ -ую ось, а матрица  $M_i$  осуществляет поворот пика в пространстве значений  $X$ . В двумерном случае эта матрица имеет вид:

$$M = \begin{Bmatrix} 1 & \sigma_{12} \\ \sigma_{21} & 1 \end{Bmatrix}.$$

где  $\sigma_{12}, \sigma_{21}$  - параметры. В 3-мерном случае таких  $\sigma_{ij}$ ,  $i \neq j$  будет уже 6, и т.д. Фон описывается полиномом от компонент вектора  $X$ . В многомерном случае фон выражен гораздо слабее, чем в одномерном (преимущество многомерных измерений), и можно брать полиномы невысоких степеней, а при размерности вектора  $X$ , большей 3 (а иногда и 2), можно фоном пренебречь вообще.

В 3-мерном случае можно использовать методы прямого разделения пиков (без подгонки), поскольку в этом случае пики не перекрываются так сильно, как в 1-мерном, и, по крайней мере, по одной координате они разделены.

Поиск 3-мерных пиков осуществляется так.

1. Строится сглаженная копия спектра;
2. Находятся локальные максимумы  $P_i = x_i, y_i, z_i$ ,  $i = 1, \dots, n$ :

$$s(x_i, y_i, z_i) > s(x_i \pm 1, y_i \pm 1, z_i \pm 1),$$

такие, что удовлетворяют критерию чувствительности  $s(P_i) > c_0 \sqrt{s(P_i)}$ .

3. Находятся квази-кривизны:

$$q_x = s''_{xx}(x, y, z) / \sqrt{1 + s_x'^2(x, y, z)}$$

$$q_y = s''_{yy}(x, y, z) / \sqrt{1 + s_y'^2(x, y, z)}$$

$$q_z = s''_{zz}(x, y, z) / \sqrt{1 + s_z'^2(x, y, z)}$$

4. Условием присутствия пика является отрицательность всех  $q$  и

$$\max(|q_x|, |q_y|, |q_z|) > 1.5$$

5. Найденные пики проверяются на разрешение: если для расстояния между пиками  $i, j$  имеет место

$$\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} < res,$$

пики  $i$  и  $j$  сливаются:

$$\hat{x} = k_1 x_i + k_2 x_j, \quad \hat{y} = k_1 y_i + k_2 y_j, \quad \hat{z} = k_1 z_i + k_2 z_j,$$

где  $k_1 = s(P_i)/(s(P_i) + s(P_j))$ ,  $k_2 = s(P_j)/(s(P_i) + s(P_j))$ .

6. Затем оцениваются полуширины  $w_x, w_y, w_z$ ;

7. Затем суммируются отсчеты исходного спектра, попадающие в область

$$\left(\frac{x - x_i}{w_x}\right)^2 + \left(\frac{y - y_i}{w_y}\right)^2 + \left(\frac{z - z_i}{w_z}\right)^2 = r^2,$$

и их сумма является оценкой гиперобъема 3-мерного  $i$ -го пика.

**Кристаллографическая задача (Rietveld analysis).** Упрощенно кристалл можно трактовать как некую базисную конфигурацию нескольких атомов (элементарную ячейку кристалла), периодически повторенную по всем направлениям в пространстве. Эта ячейка образует многогранник, в котором можно определить локальную систему координат (сингонию), взяв за координатные оси ребра многогранника. В общем случае нам понадобятся 3 линейно-независимых ребра с длинами  $A, B, C$  и углами  $\alpha = \angle BC$ ,  $\beta = \angle AC$ ,  $\gamma = \angle AB$ . Для  $j$ -го атома в этом многограннике можно вычислить координаты  $x_j, y_j, z_j$ , определенные в интервале  $[-1, 1]$  - получаем т.н. представление ячейки кристалла в прямом пространстве. Дополнительным к прямому служит понятие обратного пространства кристалла. Это тоже многогранник, имеющий ребра с длинами  $a^*, b^*, c^*$  и углы  $\alpha^*, \beta^*, \gamma^*$ , и связанный с исходным так, что объем ячейки в обратном пространстве обратно пропорционален объему в прямом. Важное для дифракции понятие - система плоскостей отражения - определяется как множество векторов  $H$  в обратном пространстве, координаты которых - целочисленные величины  $h, k, l$  (индексы Миллера), а межплоскостное расстояние в системе есть величина

$$d_H = 1/\sqrt{(H, H)}, \quad (24)$$

где скалярное произведение  $(H, H)$  записывается в терминах обратного пространства так:

$$(H, H) = a^{*2}h^2 + b^{*2}k^2 + c^{*2}l^2 + 2a^*b^*\cos(\gamma^*)hk + 2a^*c^*\cos(\beta^*)hl + 2b^*c^*\cos(\alpha^*)kl;$$

Дифракционный рефлекс  $H$  характеризуется также структурным фактором в ячейке атомов  $X_j, j = 1, 2, \dots, na$  - величиной, описываемый выражением

$$F_H = \sum_j b_j \cdot N_j \cdot \exp(2\pi i(H, X_j)) \cdot \exp(-(B_j \cdot H, H)) \quad (25)$$

Здесь  $b_j, N_j, B_j$  - соответственно, длина когерентного рассеяния нейтронов, атомная плотность и тепловой тензор  $j$ -го атома.

Формально профиль экспериментального нейтронно-дифракционного спектра  $y(D)$  может быть записан следующим образом

$$y(D) = f(D) \cdot \left( \sum_{i=1}^{nph} a_i \cdot R_i(D) + back(D) \right) + E(D), \quad (26)$$

где  $D$  является каналом (функцией межплоскостного расстояния  $d$ ),  $D_1 \leq D \leq D_2$ ;  $back$  является фоновым распределением нейтронов, для которых описание может быть двояким:

1. полиномиальное  $back(D) = \sum_{i=1}^5 q_i \cdot Q_i(Z)$ ,

$$Q_0(Z) = 1; \quad Q_1(Z) = Z; \quad Q_2 = 3 \cdot Z^2 - 1; \quad Q_3(Z) = 5 \cdot Z^3 - 3 \cdot Z$$

$$Q_4(Z) = 35 \cdot Z^4 - 30 \cdot Z^2 + 3; \quad Q_5(Z) = 63 \cdot Z^5 - 70 \cdot Z^3 + 15 \cdot Z,$$

2. обобщенным описанием квази-максвелловского типа

$$back(D) = (q_1 + q_2 \cdot Z + q_3 \cdot Z^2) \cdot \exp\left(-\frac{D - q_4}{q_5 \cdot Z + q_6}\right),$$

Здесь  $Z = (2 \cdot D - D_1 - D_2)/(D_2 - D_1)$ , а  $q_i$  - искомые параметры;  $E$  является статистической ошибкой, которая имеет нулевое среднее и дисперсию  $v(D)$ , и независима в каждой точке  $D$ .

$f(D)$  является эффективным спектром, умноженным на величину  $\exp(-x_i \cdot Z)$  (для коррекции спектра на эффекты, подобные абсорбции); здесь  $x_i$  является параметром.  $R_i$  является вкладом  $i$ -ой фазы порошка в распределение  $y$  с параметром  $a_i$ :

$$R_i = \sum_{Hi}^{nr_i} \mu_{Hi} \cdot |F_{Hi}|^2 \cdot m_i \left( \frac{D - D_{Hi}}{W_i + c_i \cdot D} \right).$$

Здесь  $F_{Hi}$  - структурный фактор (25),  $Hi$  = вектор индексов Миллера  $i$ -ой фазы,  $m_i(D)$  - модель, а  $W_i$  и  $c_i$  - параметры полуширины дифракционного пика  $i$ -ой фазы. Положение времяпролетного дифракционного пика  $D_H$  вычисляется согласно

$$D_H = const \cdot L \cdot \sin(\theta + \theta_0) d_H - D_0$$

$const = 505.556/r$ ,  $r$  = ширина канала (в мкс),  $\theta, \theta_0$  - угол Брэгга и его поправка;  $L$  = времяпролетная база, а  $d_H$  - межплоскостное расстояние (24).

$\mu_{Hi}$  - множественности  $Hi$ -ого рефлекса, нормированные на фактор Лоренца ( $1/d_H$  в 4-й степени), а также умноженные на первичную экстинкцию и поправленные на фактор преимущественной ориентации, геометрию счетчика, поглощение воздуха и образца и т.д.

Выражения, входящие в (26), могут быть упрощены. Например, можно ввести понятия симметричных атомов, т.е. выделить в ячейке некоторые базисные атомы, а остальные считать полученными из базисных (независимых) с помощью определенных преобразований симметрии. Тогда, если независимый атом имеет вектор координат  $X = (x, y, z)$  и тепловой тензор  $B$ , то симметричный ему атом будет иметь вектор координат  $R = U \cdot X + T$ , а также, для случая центральной симметрии,  $R = -U \cdot X + T$ , и тепловой тензор  $B_1 = U' \cdot B \cdot U$ , где  $U$  и  $T$  являются

матрицей преобразований симметрии и вектором трансляций, соответственно. Для каждой ячейки кристалла существует свой, зависящий от сингонии, набор таких матриц и векторов (пространственная группа). Математически преобразования симметрии соответствуют унитарным (сохраняющим скалярное произведение) преобразованиям. Использование преобразований симметрии упрощает запись структурного фактора в (26).

Можно также сократить число рефлексов  $Hi$  для вычисления выражения (26), воспользовавшись тем, что для унитарных матриц  $(H, UX) = (U^*H, X)$  и  $(UH, UH) = (U^*H, U^*H) = (H, H)$  (здесь  $U^*$  - транспонированная матрица), и тем, что трансляционная часть преобразования симметрии, хотя и меняет фазу структурного вектора, но не меняет квадрат его модуля (а в (26) фигурирует именно он), т.е.

$$\begin{aligned} \exp(2\pi i(H, UX_j)) &= \exp(2\pi i(U^*H, X_j)), \\ \exp(-(B_j \cdot UH, UH)) &= \exp(-(B_j \cdot U^*H, U^*H)), \end{aligned}$$

и, следовательно, рефлексы  $U^*H$  имеют одинаковые  $d_H$  и квадраты модулей структурных факторов. Мы можем исключить эти рефлексы из списка, и ограничиться только рефлексом  $H$ , умножив квадрат его структурного фактора на коэффициент повторения (множественность рефлекса  $\mu_{Hi}$ ).

Кроме того, частично слагаемые в (25) могут взаимно погасить друг друга за счет разных знаков, что тоже упрощает (25).

Задачей профильного анализа является уточнение характеристик элементарной ячейки кристалла, т.е. параметров  $A, B, C, \alpha, \beta, \gamma$ , и характеристик ее атомов, т.е., параметров  $X_j, B_j$  и  $N_j$ .

Метод Rietvelda может также использоваться и для учета дифракционного рассеяния нейтронов за счет взаимодействия нейтронных спинов с магнитными моментами атомов. Предполагается, что магнитные атомы ячейки кристалла образуют отдельную фазу, которая дает свой вклад в общую интенсивность (26).

**Метод решения задачи** состоит в следующем. Модель пика задается функцией  $m(x)$  (в виде таблицы или формулы) и параметризуется как и в алгоритме UP-EAK. В VMRIA пики  $i$ -ой фазы описываются независимыми моделями  $m_i$ ; не более чем 4 вместе. По умолчанию могут быть взяты некоторые широко применяемые функции, такие, как Гауссиан, Лоренциан, Фойгтман или свертка гауссовой функции и экспоненциальных, но во многих случаях (в частности, у RTOF-спектров) пики имеют очень неправильную, с резкой асимметрией или с отрицательными хвостами, форму, так что практически графический способ является единственным способом для преодоления проблемы формы пика.

Это относится и к фоновой модели.

Параметры получаются как описанные в курсе нелинейные мнк-оценки, т.е., с помощью итерационного процесса, который, начиная с начального вектора параметров  $P_0$ , получает уточненные вектора  $P_k$  минимизацией выражения

$$F(P) = \sum w(D) \cdot (y(D) - m(D, P))^2 \quad (27)$$

в пространстве значений параметров при граничных условиях

$$p_{il} \leq p_i \leq p_{iu} \quad (28)$$

$$p_j = a_{ji} \cdot p_i + b_{ji}, \quad j > i. \quad (29)$$

Чтобы найти минимум (27), используется демпфированный полно-матричный процесс Гаусса-Ньютона, включающий обобщенные веса как средство сделать процесс робастным, и проектирующий параметры, получаемых в итерациях, на ограничения (28,29). Равенство  $p_{il} = p_i = p_{ii}$  означает фиксацию  $i$ -ого параметра. В VMRIA используется компрессия мнк-матрицы за счет удаления строк и столбцов, соответствующим фиксированным и зависимым параметрам и это ведет к существенной экономии памяти ЭВМ.

**Задача Powder Match.** Алгоритм Rietveld метода работает надежно, если начальные значения параметров элементарной ячейки и атомов кристалла заданы достаточно близко к истинным. Если такой информации у физика на начальном этапе анализа данных нет, помочь может разбивка задачи на отдельные части и решение задачи по частям. Powder Match есть процедура уточнения одних параметров элементарной ячейки, а функции, описывающие структурный фактор заменяются обычными пиковыми функциями. Положения у них будут вычисляться по кристаллографическим формулам, но амплитуды (площади) и полуширины описываться параметрами  $A, W$ , как в алгоритме UPEAK.

Эти параметры могут не представлять интереса не только на начальном этапе анализа, но и в дальнейшем, если, например, ординаты спектра искажены неучтенными в модели эффектами (такими, как текстурные или др.) Но это не имеет значения, если интерес представляют только параметры ячейки.

Впрочем результат решения задачи Powder Match: уточненные параметры ячейки и площади пиков, - может оказаться полезным для полного анализа кристалла. Мы можем преобразовать площади и положения пиков в структурные факторы; применить к последним процедуру Fourier синтеза, и получить, по крайней мере, приближенную качественную информацию об атомах ячейки, и далее уточнить ее методом Rietvelde.

**Задача Fourier синтез.** Пусть  $r(X)$  - функция, определенная на  $-1 \leq x_i \leq 1$ ] (атомная плотность идеального кристалла), и вектор  $X = (x, y, z)$  (координаты атома в элементарной ячейке); и пусть далее задан целочисленный вектор  $H = (h, k, l)$ ,  $h = 0, 1, \dots, h_{max}, k = 0, 1, \dots, k_{max}, l = 0, 1, \dots, l_{max}$  (индексы Миллера).

Структурный фактор  $f$  упрощается

$$f(H_k) = \frac{1}{n} \sum_{j=1}^m r(X_j) \exp(i2\pi(H_k, X_j)), \quad k = 1, 2, \dots, n. \quad (30)$$

Здесь  $n$  - число всех комбинаций  $h, k, l$ . Выражение (30) похоже на 3-хмерное обобщение стандартного ДПФ (10), но в отличие от него здесь  $n$  может быть больше, чем  $m + 1$ . Отметим некоторые свойства (30):

1. Если мы заменим некоторые  $x_j$  на  $x_j \pm 1$ , это не повлияет на (30); поэтому, мы можем свести интервал  $[-1, 1]$  к  $[0, 1]$ .
2.  $f(-H_k) = f^*(H_k)$ .

Возникает задача: имея множество  $f(H_k), k = 1, \dots, n$ , для которого справедливо (30), восстановить функцию  $r(X)$  в точках  $X_1, \dots, X_m$ . Естественным методом для решения этой задачи является обобщение обратного ДПФ:

$$g(X_j) = \frac{1}{n} \sum_{k=1}^n f(H_k) \exp(-i2\pi(H_k, X_j)); \quad (31)$$

Условие для того, чтобы (31) было обратным оператором к (30), т.е., чтобы  $g(X_j)$  было подобно  $r(X_j)$ , выглядит так

$$(h_{max} + 1)(x_l - x_j) = I_{1lj}, \quad (k_{max} + 1)(y_l - y_j) = I_{2lj}, \quad (l_{max} + 1)(z_l - z_j) = I_{3lj}, \quad (32)$$

где  $I_{1lj}, I_{2lj}, I_{3lj}$  - целые числа.

Другим условием является то, что последовательность индексов не должна содержать пропусков.

(32) может быть применено для определения разрешения в дискретном синтезе Фурье:

$$\delta x = \frac{1}{\sqrt{2}(h_{max} - h_{min} + 1)} \quad \delta y = \frac{1}{\sqrt{2}(k_{max} - k_{min} + 1)} \quad \delta z = \frac{1}{\sqrt{2}(l_{max} - l_{min} + 1)} \quad (33)$$

Два или более атомов, которые расположены ближе друг к другу, чем (33), не могут быть различены на карте Фурье. В этом отличие дискретного от непрерывного синтеза Фурье. В последнем такое важное в практике понятие как разрешение отсутствует вообще.

Другое отличие состоит в том, что если (32) не выполняется, восстановленная функция  $g(X_l)$  содержит ошибку (помеху), и если некоторое  $r(x_j)$  имеет слишком маленькую амплитуду, оно может быть подавлено этой помехой, и также не быть замечено на карте Фурье. Это - чувствительность - величина  $s$ , для которой

$$\|r(x_j)\| \geq s > 0 \quad (34)$$

должно выполняться для каждого  $j$ .

С помощью (31) мы можем решить синтеза Фурье. Если мы имеем набор структурных факторов  $f(H_k)$ ,  $k = 1, \dots$ , и выполнены условия (33), мы можем вычислить функцию  $G(x, y, z)$ , и если она подобна  $r(x, y, z)$ , она будет иметь пики там же, где и последняя. Имеем задачу поиска пиков функции  $G(x, y, z)$ , и тем самым, определения (хотя бы качественного) картины распределения атомной плотности. Для поиска пиков можем применить описанный выше метод (алгоритм DOMUS).

**Задача Powder indexing. Алгоритм AUTOX.** Задача возникает при необходимости определить параметры элементарной ячейки ab initio, т.е. не располагая никакой априорной информацией и формально ставится следующим образом: пусть задано множество межплоскостных расстояний ( $d_j$ ),  $j = 1, 2, \dots, m$  (как значения  $d$ , углы или времяпролетные каналы).

Каналы  $c$  (или углы  $a$ ) могут быть переведены в  $d$  с помощью формулы Брэгга:  $d = c / (\text{const} \cdot \sin(\theta)L)$  ( $d = \lambda / (2\sin(a))$ ).

Каждое  $d_j$  ( $d$ -spacing) является функцией параметров ячейки и трех индексов Миллера  $i$ -ой фазы ( $i = 1, \dots, n$ ) порошка

$$d_j = d(P_i, H_{ij}), \quad \text{где} \quad (35)$$

$P_i$  - вектор параметров прямой ячейки;  $A, B, C, \cos(\alpha), \cos(\beta), \cos(\gamma)$   $i$ -ой фазы;  $H_{ij}$  - вектор  $h, k, l$ ,  $i$ -ой фазы, отнесенный к  $j$ -ому рефлексу.

Требуется на основании заданных множества ( $d_j$ ) и  $n$  определить фазу, параметры  $P_i$  и  $H_{ij}$  для каждого  $d_j$ .

Функционал качества индексации выглядит так:

$$F = \sum_{i=1}^n \sum_{j=1}^m \rho[d_j, d(P_i, H_{ij})]. \quad (36)$$

Здесь  $\rho$  - метрика. Обычно используются квадратичная и модульная метрики. Решение задачи получается при минимуме (36) по  $P_i$  и  $H_{ij}$ , соответственно. Это решение является неединственным. Для обеспечения единственности, а также чтобы сузить область поиска, следующие условия являются строго необходимыми:

1. Ограничения на максимальное значение объема прямой элементарной ячейки и размеров ее ребер, и на область допустимых индексов.

$$v(P_i) \leq V_{imax} \quad (37)$$

$$\begin{array}{lll} A_{li} \leq A_i \leq A_{ui} & B_{li} \leq B_i \leq B_{ui} & C_{li} \leq C_i \leq C_{ui} \\ \alpha_{li} \leq \alpha_i \leq \alpha_{ui} & \beta_{li} \leq \beta_i \leq \beta_{ui} & \gamma_{li} \leq \gamma_i \leq \gamma_{ui} \end{array} \quad (38)$$

$$\|h_i\| \leq h_{imax} \quad \|k_i\| \leq k_{imax} \quad l_i \leq l_{imax} \quad (39)$$

где  $i$  - номер фазы, а  $v$  является функцией объема; некоторые параметры фиксированы и/или зависимы, если сингония не является триклинной;

2. Для триклинной и моноклинной сингоний углы прямой решетки должны быть максимально близкими к 90 градусам.
3. Для многофазного случая недопустима полная пересекаемость областей определения  $P_i$  и/или  $H_j$ .

Дополнительно можно использовать следующий прием. Решение ищется как такое множество параметров и индексов Миллера, которое минимизирует выражение

$$F = F_1(P_i, H_j^*) + \sum_{i=1}^n \kappa_i q_i (A_i^2 + B_i^2 + C_i^2). \quad (40)$$

где  $F_1$  есть выражение типа (36). Здесь  $\kappa_i$  являются заданными множителями, делающими влияние  $F_2$  на решение большим или меньшим, и они играют роль, схожую со штрафами в методе штрафных функций;  $q_i$  являются нормирующими множителями, а  $A_i, B_i, C_i$  - параметрами  $i$ -ой фазы прямой ячейки.

Для триклинной симметрии угловые ограничения должны позволять лишь одному из них варьироваться в диапазоне  $(0, \pi)$ , тогда как другие должны принадлежать или интервалу  $(a_{min}, \frac{\pi}{2})$ , или  $(\frac{\pi}{2}, a_{max})$ , в зависимости от того, какой угол (т.е. из какого интервала) ближе к  $\frac{\pi}{2}$ . Это следует из того, что  $\cos(a) = -\cos(\pi - a)$ .

То же самое применимо к случаю примитивной моноклинной (или ромбоэдрической) симметрии. Ограничение имеет вид:

$$\beta_{min} \leq \beta \leq 90, \text{ или } 90 \leq \beta \leq \beta_{max}.$$

Итак, задача ставится следующим образом: минимизировать (36) (или (40)) при условиях (37).

Минимизация  $F$  по  $P_i$  и  $H_j$  осуществляется с помощью серии процессов, каждый из которых выглядит так:

- Подбирается методом Монте-Карло или детерминированным образом начальное значение вектора  $P_c$  (в общем случае, многофазного), удовлетворяющее ограничениям (37);

- Подбираются индексы Миллера  $H_{ij}$ , которые при  $P_c$  дают минимальное расстояние от  $d(P_c, H_{ij})$  до  $d_j$ ; фаза последнего полагается равной тому  $i$ , при котором это расстояние минимально.
- Вычисляется начальное значение критерия  $F(P_c, H)$ .
- Осуществляется итеративное уточнение параметров на основе робастного метода Гаусс - Ньютона, т.е. на каждой итерации, начиная с  $P_c, H_c$ , мы получаем новые параметры  $\hat{P}_c$  и новые  $\hat{H}_j$ . Новое значение  $F$  сравнивается со старым: если оно меньше, уточнение продолжается; если нет, текущий процесс подгонки останавливается. В случае сходимости мы имеем текущие оптимальные  $P$  и  $H$  и текущую точность подгонки  $\epsilon = \text{MAX} | d_j - d(P_i, H_{ij}) |$ .

Решение с минимальным  $\epsilon$  по всем процессам считается решением задачи индексации порошка, если значение  $\epsilon$  приемлемо.

Уникальными чертами метода, реализованного в программе, являются возможность анализировать многофазные поликристаллы и то, что он не привязан к определенным ( в частности, первым) индексам Миллера.

## ПРИЛОЖЕНИЕ. МАТЕМАТИЧЕСКИЙ МИНИМУМ.

Целью данного раздела является напоминание некоторых математических фактов из анализа, вычислительной математики, алгебры, теории вероятностей и статистики, необходимых для понимания методов, используемых для обработки экспериментальных данных вообще, и, в частности, в предлагаемой работе.

**Матричная алгебра.** Вектором  $P_n$  размерности  $n$  называется набор чисел  $(p_1, p_2, \dots, p_n)$ . Евклидова норма вектора  $P_n$  это:

$$\|P\| = \sqrt{\sum_{i=1}^n p_i^2}. \quad \text{Взвешенная: } \|P\| = \sqrt{\sum_{i=1}^n w_i p_i^2}, \quad w_i > 0$$

Скалярное произведение:

$$(X, Y) = \sum_{i=1}^n x_i y_i \quad \text{или} \quad (X, Y) = \sum_{i=1}^n w_i x_i y_i$$

Иначе:

$$(X, Y) = \|X\| \|Y\| \cos(\phi); \quad \phi = \angle XY$$

В комплексном пространстве  $(X, Y) = \sum_{i=1}^n x_i y_i^*$ .

Ортогональность векторов  $X$  и  $Y$ :  $(X, Y) = 0$ ,  $X \neq 0$ ,  $Y \neq 0$ .

Вектора  $Q_1, Q_2, \dots, Q_k$  называются линейно-независимыми, если не существуют одновременно не равные нулю числа  $c_j$ ,  $j = 1, n$  такие, что имеет место

$$\sum_{j=1}^k c_j Q_j \equiv 0, \quad c_j \neq 0$$

Ортогональность есть частный случай линейной независимости.

Матрица  $A_{mn}$  с элементами  $a_{ij}$ ,  $i = 1, m$ ;  $j = 1, n$ :

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Определены следующие операции над матрицами:

Транспонирование матрицы - переход к матрице  $B$  такой, что  $b_{ij} = a_{ji}$ .

Обозначение:  $A'$ .

Сумма матриц  $A_{mn}$  и  $B_{mn}$  есть матрица  $C_{mn}$  такая, что

$$c_{ij} = a_{ij} + b_{ij}, \quad \equiv C = A + B$$

Произведение матриц  $A_{mn}$  и  $B_{nl}$  есть матрица  $C_{ml}$  такая, что

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}, \quad \equiv C = AB$$

Произведение матрицы  $A_{mn}$  на скаляр  $\alpha$  есть матрица  $C_{ml}$  такая, что

$$c_{ij} = \alpha \cdot a_{ij}, \quad \equiv C = \alpha A$$

Частный случай произведения матриц есть произведение матрицы  $A_{mn}$  на вектор  $X_m$ : это вектор  $Y_n$  такой, что  $Y = AX$

Наиболее интересным является случай квадратных матриц, когда  $n = m$

Среди них различаются:

1. Симметричные матрицы.  $a_{ij} = a_{ji}$ .
2. Диагональные матрицы.  $a_{ij} = 0$ , если  $i \neq j$ .
3. Ортогональные матрицы - такие, что любые 2 столбца (строки) этой матрицы являются ортогональными векторами.

Обобщением ортогональной матрицы на случай комплексных чисел является унитарная матрица. Для обеих справедливо  $A^* = A^{-1}$ .

Единичная (диагональная) матрица  $I$ ,  $i_{ii} = 1$ . Обратная матрица  $A^{-1}$ ;  $AA^{-1} = I$ .

Ранг матрицы = максимальное число линейно-независимых столбцов (или строк).

Вырожденность матрицы  $A_{nn}$  имеет место, если ее ранг меньше  $n$ . Вырожденная матрица не имеет обратной.

В компьютерных операциях с матрицами помимо вырожденности может иметь место плохая обусловленность: это когда теоретически невырожденная матрица может быть сколь угодно близка к вырожденной в силу приближенного характера представления чисел на ЭВМ и действий с ними. Количественной характеристикой такой близости может служить детерминант матрицы  $Det(A)$ ; если он близок к нулю или даже является машинным нулем (т.е. числом, не представимым на ЭВМ), то компьютерный представитель матрицы может оказаться вырожденным. Исходную матрицу в этом случае называют плохо обусловленной.

След матрицы  $Sp(A)$ ,  $Sp = \sum_{i=1}^n a_{ii}$ .

Квадратный корень из матрицы  $A$ . Это матрица  $B$  такая, что  $A = BB$ . Квадратный корень интересен тем, что его обусловленность лучше, чем у исходной матрицы, а так как  $A^{-1} = B^{-1}B^{-1}$ , то плохо обусловленную матрицу лучше всего обращать с помощью обращения ее квадратного корня. Однако не каждая матрица имеет квадратный корень, а если он существует, то в разных классах матриц он имеет разный вид.

Геометрические иллюстрации. Эффект воздействия матрицы на вектор это вращение вектора и/или изменение его длины.

Положительно-определенная матрица  $A$  определяется из условия

$$(AX, X) > 0, \quad \text{для любого вектора } X.$$

Наглядный смысл такой матрицы следует из:  $\|AX\| \|X\| \cos(\phi) > 0 \rightarrow |\phi| < 90^\circ$

т.е. положительно-определенная матрица поворачивает вектор  $X$  менее чем на  $90$  градусов. Для диагональной матрицы положительная определенность означает положительность ее диагональных элементов: все  $a_{ii} > 0$ . Это следует из

$$(AX, X) = \sum_{i=1}^n a_{ii} x_i^2$$

Собственным значением и собственным вектором матрицы являются вектор  $E$  и число  $\lambda$  такие, что:

$$AE = \lambda E$$

Собственные значения обратной матрицы  $1/\lambda$ ; следовательно, если среди собственных чисел прямой матрицы есть числа, близкие к нулю, то у обратной

матрицы будут собственные числа, близкие к бесконечности. Вырожденная матрица имеет по крайней мере одно  $\lambda = 0$ .

Собственные значения диагональной матрицы:  $a_{ii}$ , а ее собственные вектора это вектора, у которых только  $ii$ -ая координата отлична от нуля.

У положительно-определенной матрицы собственные значения положительны.

Матрица - отличная модель линейного оператора, идеально подходящая для работы с такими на ЭВМ. В евклидовом базисе норма матрицы  $A$  равна корню квадратному из максимального собственного значения матрицы  $AA'$ .

УПРАЖНЕНИЕ 1. Доказать (или хотя бы показать на примере), что матрица, у которой в каждой строке и столбце есть только один элемент, равный  $\pm 1$ , а остальные равны нулю, является ортогональной.

Эти матрицы играют большую роль в кристаллографии, где они являются представлениями пространственных групп симметрии.

УПРАЖНЕНИЕ 2. Показать, что произведение матриц вышеописанного типа есть матрица этого же типа. а умножение такой матрицы на вектор эквивалентно перестановке компонент этого вектора.

**Минимум функций нескольких переменных.** Функционалом называется отображение абстрактного пространства на числовую ось. Пример.

$$F(f(x)) = \int_X (y(x) - f(x))^2 dx + \alpha \int_X f''^2(x) dx.$$

Здесь функционал  $F$  для заданных  $X$ ,  $y(x)$  и  $\alpha$  определен на пространстве дважды дифференцируемых функций  $f(x)$ , интегрируемых в квадрате вместе с производными  $f''(x)$  и  $y(x)$  на  $X$ .

Он становится обычной функцией  $n$  переменных  $F(x_1, x_2, \dots, x_n)$ , если  $f(x)$  фиксирована, но зависит от переменного вектора  $P$ :

$$F(P) = \int_X (y(x) - f(x, P))^2 dx + \alpha \int_X f''^2(x, P) dx$$

**ЭКСТРЕМУМЫ.** Сложная процедура нахождения экстремумов функционалов существенно упрощается в случае обычных гладких функций  $n$  переменных.

Необходимое условие минимума:

$$\frac{\partial f}{\partial x_i} = 0 \quad \equiv \quad grad f = 0$$

Достаточные условия.

- матрица 2-х производных положительно - определена = минимум;
- матрица 2-х производных отрицательно - определена = максимум;
- матрица 2-х производных ни одно из двух = седло;
- матрица 2-х производных имеет собственное значение 0 = экстремум не единственен;

Доказательство следует из формулы  $f(x+h) \simeq f(x) + (h, f') + \frac{1}{2}(f''h, h)$ , где  $x$  - точка экстремума. Так как  $f' = 0$ , то  $f(x+h)$  будет больше или меньше, чем  $f(x)$  в

зависимости от  $(f''h, h)$ .

ПРИМЕРЫ. Данные условия очень наглядно иллюстрируются примерами экстремумов функций двух переменных  $x, y$ :  $x^2 + y^2$ ,  $-x^2 - y^2$ ,  $x^2 - y^2$ ,  $x^2$ .

Функционалы общего типа тоже играют большую роль в экспериментальной физике, особенно типа

$$F_1(y, y') = \int_a^b f(x, y(x), y'(x)) dx$$

$$F_2(y, y', y'') = \int_a^b f(x, y(x), y'(x), y''(x)) dx$$

в непрерывной форме интегралов по  $x$  или дискретной - сумм по  $x_i$ . Это связано с тем, что в роли  $y$  часто выступают экспериментальные данные, так что эффекты от них и их производных целесообразно оценивать просуммированными и усредненными величинами.

Экстремумами таких функционалов занимается математическая дисциплина - вариационное исчисление. Центральным его понятием является вариация - если  $F(y(x))$  есть функционал на пространстве функций  $y(x)$ , то элемент  $h(x)$ , на который изменился  $y(x) \rightarrow y(x) + h(x)$ , называется вариацией переменной  $\delta y$ , а изменение  $F(y(x) + h(x)) - F(y(x))$  - вариацией функционала  $\delta F$ .

Условие экстремума  $F$  похоже на аналогичные условия для функций:  $\delta F = 0$ .

Для функционалов типа  $F_1$  или  $F_2$  это условие имеет специфическую форму, т.н. уравнения Эйлера. Для  $F_1$  и  $F_2$  это будут

$$\frac{\partial F_1}{\partial y} - \frac{d}{dx} \frac{\partial F_1}{\partial y'} = 0; \quad (41)$$

$$\frac{\partial F_2}{\partial y} - \frac{d}{dx} \frac{\partial F_2}{\partial y'} + \frac{d^2}{dx^2} \frac{\partial F_2}{\partial y''} = 0; \quad (42)$$

УПРАЖНЕНИЕ. Попробуем вывести (41) для  $F_1$  в простейшем случае  $y(a) = a_0$ ;  $y(b) = b_0$ . Если  $h(x)$  - вариация  $y(x)$ , то при постоянных  $a_0, b_0$   $h(a) = 0$ ,  $h(b) = 0$ , а вариация  $y'(x)$  очевидно равна  $h'(x)$ . Следовательно, вариация  $F_1$  будет иметь вид

$$\delta F_1(y, y') = \int_a^b \left( \frac{\partial F_1}{\partial y} h(x) + \frac{\partial F_1}{\partial y'} h'(x) \right) dx$$

Этот интеграл есть сумма двух интегралов, и интегрируя второй по частям, и учитывая, что  $h(a) = 0$ ,  $h(b) = 0$ , мы получим

$$\delta F_1(y, y') = \int_a^b \left( \frac{\partial F_1}{\partial y} h(x) - \frac{d}{dx} \frac{\partial F_1}{\partial y'} h(x) \right) dx$$

Чтобы  $\delta F_1(y, y') = 0$ , необходимо выполнение (41).

**Разностные схемы.** Теоретическим моделям, описываемым как правило, непрерывными (и обычно достаточно гладкими) функциями, в эксперименте соответствуют дискретные функции. Т.е. множеству  $X : x_0 \leq x \leq x_1$  и теоретической функции  $f(x)$  соответствуют дискретное множество  $X_m : x = x_1, x_2, \dots, x_m$  и функция

$f(i)$ , где  $i$  - номер точки  $x_i$ . Значения функции  $f$  в точках, отличных от  $x_i$ ,  $i = 1, \dots, m$ , могут быть получены с помощью интерполяции, в результате чего мы получаем новую функцию, совпадающую со старой в точках  $x_i$  и отличающуюся от нее в остальных точках. Эти отличия можно минимизировать различными способами, в зависимости от различных критериев оценки качества приближения, так что семейство всевозможных интерполяторов очень велико.

Естественно, такие фундаментальные понятия математического анализа как производная и интеграл заменяются их дискретными аналогами - разностями и суммами.

**ПРИМЕРЫ** дискретных конструкций.

1. Простейший (линейный) интерполятор: если  $x_1 \leq x \leq x_2$ , то

$$f(x) = f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1} (x - x_1).$$

2. Простейшая 1-я разность  $\Delta_1 f(i) = \frac{f(i+1) - f(i)}{(x_{i+1} - x_i)}$ .

3. Простейшая 2-я разность  $\Delta_2 f(i) = \frac{f(i+1) - f(i-1)}{(x_{i+1} - x_{i-1})}$ .

4. Простейшая интегральная сумма  $S = \sum_{i=1}^{m-1} f(x_i)(x_{i+1} - x_i)$ .

Наиболее просто результат дискретной операции строится как линейная комбинация величин  $f(i)$ :  $r(i) = \sum_{i=1}^m a_i f(i)$ . Коэффициенты  $a_i$  выбираются из предположений о возможном характере непрерывной функции  $f(x)$  и требований к качеству приближения ее дискретным аналогом. Легко видеть, что приведенные выше примеры исходят из предположения о кусочно-линейном характере функции  $f(x)$  в промежутках между узловыми точками  $x_i$ ; иначе, из требования, чтобы кусочно-линейные функции приближались в них точно во всех точках, и если такие функции действительно приближают функцию с приемлемой точностью, то все в порядке.

Наиболее совершенные методы построения коэффициентов  $a_i$  - вариационные, в частности, сплайновые. Они будут рассмотрены в главе о фильтрации данных. Другой подход состоит в построении оптимальных узлов  $x_i$  - если позволяют условия эксперимента.

**УПРАЖНЕНИЕ.** Для простоты предположим, что все  $x_i = hi$ .

Возьмем  $\Delta_{11} f(i) = (f(i) - f(i-1))/h$  и  $\Delta_{12} f(i) = (f(i+1) - f(i-1))/(2h)$ . Оценим точность приближения 1-й производной этими величинами.

Имеем

$$f(i-1) \sim f(i) - hf'(i) + 0.5h^2 f''(i) - (1/6)h^3 f'''(i)$$

$$f(i+1) \sim f(i) + hf'(i) + 0.5h^2 f''(i) + (1/6)h^3 f'''(i)$$

Отсюда можно показать, что:

1. асимптотическая точность у  $\Delta_{12}$  выше, чем у  $\Delta_{11}$ ;
2. на немонотонных участках  $\Delta_{12}$  может давать очень большую погрешность, больше чем  $\Delta_{11}$ .

Решения разностных уравнений часто имеют вид:  $f(x_{k+1}) = \sum_{i=1}^k a_i f(x_i)$ , т.е.  $f(k+1)$  вычисляется через предыдущие значения  $f$ . Такие схемы могут накапливать погрешность так, что решение может стать неустойчивым при росте  $k$ . Более устойчивы схемы с опережением, т.е. использующие все значения  $f$ . С одной такой схемой - 'прогонкой' - мы познакомимся в главе о фильтрации данных.

**Теория вероятностей.** Случайная величина  $x$  как и обычная переменная  $x$  принимает значения из некоторой области  $X$ , но случайная величина принимает эти значения с определенными вероятностями  $P$ .

Таким образом, основной объект, с которым имеет дело теория вероятностей это тройка понятий: случайная величина, событие, функция распределения.

$$\xi, \quad \xi < t, \quad F(t) = P(\xi < t), \quad x \in X, \quad t \in [-\infty, +\infty].$$

В данном примере имеется случайная величина  $\xi$  со значениями из области  $X$ , событие, состоящее в том, что  $\xi < t$ , и соответствующая функция распределения  $P$ .

Дискретная и непрерывная случайные величины.

Если  $X$  - дискретное множество, целесообразно считать событиями равенство какому-либо числу и задавать функцию распределения так:  $F(t) = P(\xi = t)$ . В непрерывном случае такое задание невозможно, и там событие и функция распределения задаются так, как указано выше. Это самый общий способ задания этих понятий. Из него можно получить, например, вероятность такого события как

$$P(t_1, t_2) = P(t_1 \geq \xi \leq t_2) = F(t_2) - F(t_1)$$

Дискретный случай легко описывается общей формой:

$$P(\xi < t) = \sum_{t_i < t} F(\xi = t_i)$$

Гладкая функция распределения в непрерывном случае имеет производную, называемую плотностью функции распределения  $f(x)$ , так что имеет место

$$F(t) = \int_{-\infty}^t f(x)dx, \quad \text{где } f(x) = \frac{dF}{dx}.$$

Сумма событий - это  $n$  событий  $A_i$ , связанных связкой "или". Если  $A_i$  не пересекаются, то имеет место:  $P(\sum_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$

Произведение событий - это  $n$  событий  $A_i$ , связанных связкой "и". Если  $A_i$  независимы друг от друга, то  $P(\prod_{i=1}^n A_i) = \prod_{i=1}^n P(A_i)$

Важную роль играют в теории вероятностей и статистике центральные моменты 1-го и 2-го порядка случайной величины  $\xi$ : математическое ожидание (или среднее)  $\hat{E}\xi$ , дисперсия  $\hat{V}\xi$ , и корень квадратный из нее - среднеквадратичное отклонение или сигма  $\sigma$  (или ошибка как ее называют в практике анализа экспериментальных данных в физике), и ковариации двух случайных величин  $\xi_1$  и  $\xi_2$ .

Эти понятия определяются следующим образом.

Непрерывный случай.

$$\hat{E}\xi = \int_{-\infty}^{+\infty} xp(x)dx = m \tag{43}$$

$$\hat{V}\xi = \int_{-\infty}^{+\infty} (x - m)^2 p(x)dx = v^2 \tag{44}$$

Дискретный случай

$$\hat{E}\xi = \sum_{i=-\infty}^{\infty} x_i P(\xi = x_i) = m \tag{45}$$

$$\hat{V}\xi = \sum_{i=-\infty}^{\infty} (x_i - m)^2 P(\xi = x_i) = v^2 \tag{46}$$

$$\hat{V}\xi = \hat{E}\xi^2 - m^2 \quad (47)$$

Ковариации:  $cov(\xi_1\xi_2) = \hat{E}((\xi_1 - m_1)(\xi_2 - m_2))$ .

Корреляции.

$$\hat{C}(\xi_1\xi_2) = \frac{cov(\xi_1\xi_2)}{v_1v_2}. \quad (48)$$

Корреляции двух случайных величин это ковариации, нормированные на сигмы этих величин. Они служат мерой линейной зависимости между двумя случайными величинами, и благодаря нормировке и вытекающего отсюда стандартного интервала значений от -1 до +1 являются очень удобной мерой этой зависимости. Следует подчеркнуть, что в отличие от физики, где термин "корреляции" обозначает меру произвольной зависимости между произвольными величинами, в математике он служит для обозначения меры лишь линейной зависимости между случайными величинами. Далее, в математике (опять же в отличие от физики) корреляционная зависимость не используется для обозначения зависимости средних двух случайных величин (для последней используется термин "регрессионная зависимость").

## ЛИТЕРАТУРА.

1. З.Брандт. Анализ данных. Мир. 2003, Москва. (S.Brandt. Data analysis. Springer.)
2. Ватутин В.А., Ивченко Г.И., Медведев Ю.И., Чистяков В.П. Теория вероятностей и математическая статистика. Агар, М.2003.
3. Eady W.T., Dryard D., James F.E. et al. Statistical methods in experimental physics, NHRС, Amsterdam - London. 1971.
4. Lanczos C. Applied analysis. Prentice hall,inc. 1956.
5. Гришин В.К., Живописцев Ф.А., Иванов В.А. Математический анализ и интерпретация физического эксперимента. М., МГУ, 1988.
6. Злоказов В.Б. Математические методы для анализа экспериментальных спектров и спектро-подобных распределений. ФЭЧАЯ, 1985, т.16-5, с.1126-1163.
7. Ю.П.Пытьев. Методы анализа и интерпретации эксперимента. М., МГУ, 1990.

## ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА.

1. В.П.Чистяков. Курс теории вероятностей. 2003, С.Петербург, Москва.
2. Jenkins G., Watts D. Spectral analysis and its application. 1971.
3. Бахвалов Н.С. Численные методы. т.1, М., 1975.
4. Wilks S.S. Mathematical Statistics, N.Y.-Lond; J.Wiley, 1962.

Полезный материал также можно найти на сайте [www.ccp14.ac.uk](http://www.ccp14.ac.uk)

Учебное пособие

**Виктор Борисович Злоказов**

**МАТЕМАТИЧЕСКАЯ ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ  
ДАНЫХ НЕЙТРОННОГО РАССЕЙНИЯ В ФИЗИКЕ НИЗКИХ  
ЭНЕРГИЙ.**

Работа поступила в ОНТИ 15.11.2007

Тираж 50 экз.

Отпечатано в типографии «КДУ»  
Тел./Факс: (495) 939-40-36 E-mail: [pess@kdu.ru](mailto:pess@kdu.ru)