Компьютерная обработка результатов измерений

Лекция 3. Систематические и случайные погрешности. Теория оценок.

Емельянов Эдуард Владимирович

Специальная астрофизическая обсерватория РАН
Лаборатория физики оптических транзиентов

22 марта 2021 года

1. Погрешность
2. Теория оценок
3. Ковариационная матрица
4. Доверительные интервалы
5. Метод наименьших квадратов
6. Ошибка МНК

Погрешность

Погрешность — отклонение измеренного значения величины от её истинного (действительного) значения.

Абсолютная погрешность, Δx (напр., RMS); относительная погрешность, $\delta x = \Delta x/\overline{x}$; приведенная погрешность $\gamma x = \Delta x/N_x$ (нормировочный коэффициент).

По причине возникновения

Инструментальные определяются погрешностями применяемых средств измерений.

Методические обусловлены несовершенством метода, а также упрощениями, положенными в основу методики.

Субъективные обусловлены качествами экспериментатора.

Погрешность

По характеру проявления

Случайные обусловлены совокупностью внешних факторов, влияющих на результат эксперимента.

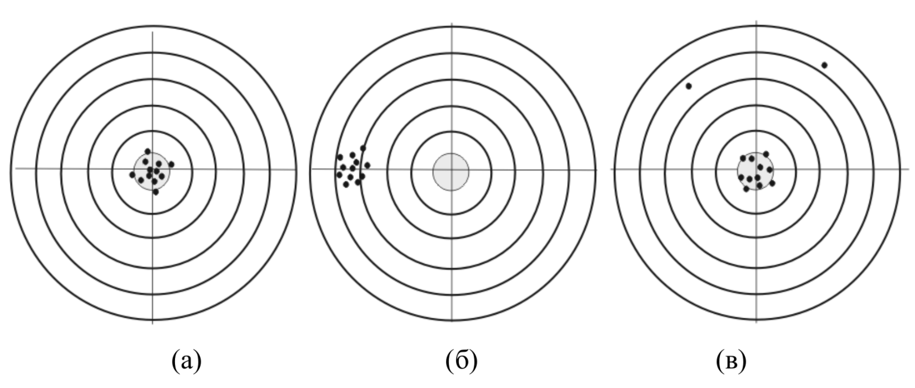
Систематические связаны с влиянием прибора на измеряемую величину или методическими ошибками, выявляются лишь сменой прибора/метода/экспериментатора.

Промахи наиболее сильно себя проявляют и связаны с неисправностью прибора или экспериментатора.

Средняя квадратическая погрешность среднего арифметического

$$\sigma_{<x>} = \frac{\sigma_x}{\sqrt{n}} = \sqrt{\frac{\sum\limits_{i=1}^{n}(\overline{x_i} - <x>)^2}{n(n-1)}}.$$

Погрешность



Примеры погрешностей: а) случайная, б) случайная и систематическая, в) случайная и промахи.

Правила вычисления погрешностей

1. $$\Delta\left(\sum a_n\right) = \sum \Delta a_n.$$

2. $$\prod(a_i \pm \Delta a_i) = \prod a_i \prod(1 \pm \delta a_i) \approx \prod a_i (1 \pm \sum \delta a_i),$$
$$\bigl(a[1 \pm \delta a]\bigr)^n \approx a^n (1 \pm n\delta a).$$

3. **В сложных функциях** вида $y = f(x_1, \ldots, x_n)$ можно оценить погрешность, воспользовавшись приближением:
$$\delta y \approx \left|\frac{dy}{y}\right| = \left|\frac{df(x_1, \ldots, x_n)}{f(x_1, \ldots, x_n)}\right|,$$
в котором следует заменить ${}^{dx_i}/_{x_i} = \delta x_i$ – относительная погрешность измерения величины x_i, $dx_i = \Delta x_i$ – абсолютная погрешность. Все слагаемые необходимо суммировать по абсолютной величине.

Правило «трех сигм»

При гауссовом распределении случайной величины вероятность

$$P(|x-\overline{x}|<3\sigma)=2\Phi(3)=0.9973.$$

(Φ – нормальное интегральное распределение).

Правило трех сигм: если случайная величина распределена нормально, то абсолютная величина ее отклонения от математического ожидания не превосходит утроенного среднего квадратичного отклонения.

Теорема Ляпунова: случайная величина, являющаяся суммой большого числа взаимно независимых случайных величин, имеет нормальное распределение.

Распределение χ^2

Распределение суммы квадратов n нормальных независимых случайных величин (x_i, $i = \overline{1, n}$, $\overline{x} = 0$, $\sigma_x = 1$): $\chi^2 = \sum_{i=1}^n x_i^2$ с $k = n$ степенями свободы. Каждое линейное соотношение уменьшает количество степеней свободы на единицу. Плотность распределения «хи квадрат»:

$$f(x) = \begin{cases} 0, & x \leqslant 0, \\ \dfrac{\mathrm{e}^{-x/2}\, x^{k/2-1}}{2^{k/2}\Gamma(k/2)}, & x > 0, \end{cases}$$

где $\Gamma(x) = \int\limits_0^\infty t^{x-1} e^{-t} dt$ – гамма-функция, в частности, $\Gamma(n+1) = n!$. $\overline{\chi^2} = k$, $\sigma^2_{\chi^2} = 2k$. Из закона больших чисел при $k \to \infty$ это распределение приближается к нормальному.

Распределение χ^2

В общем случае для любых нормальных независимых случайных величин

$$\chi^2 = \sum \left(\frac{x_i - \overline{x}}{\sigma_x} \right)^2 .$$

При $k = 2$ распределение совпадает с экспоненциальным.
Квантили распределения χ^2 вычисляются при помощи функции `chi2inv` пакета `statistics`. Например:
chi2inv([0.990:0.001:0.999], 10)
ans =
23.209 23.514 23.853 24.235 24.673 25.188 25.813 26.611 27.722 29.588
Само распределение можно отобразить при помощи `chi2pdf(x, N)`.
`chi2cdf` – интегральное распределение.

Распределение Стьюдента (t-распределение)

Строится наподобие χ^2 для $n+1$ независимой нормальной величины Y_i:

$$t = \frac{Y_0}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} Y_i^2}}.$$

Плотность распределения:

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}.$$

Распределение симметрично. $\bar{t} = 0$ при $n > 1$, $\sigma_t^2 = {}^{n}\!/_{(n-2)}$ при $n > 2$.

Распределение Стьюдента (t-распределение)

Можно представить T с k степенями свободы через нормальное и χ^2: если Z распределено нормально, а V – по закону chi^2, то

$$T = Z\sqrt{\frac{k}{V}}.$$

Распределение возникает из распределения выборочных среднего, $<X>$ и дисперсии, S:

$$\frac{<X> - \overline{X}}{S/\sqrt{n}} \propto t(n-1).$$

Аналогичные функции из пакета `statistics`: `tinv`, `tpdf`, `tcdf`:

tinv([0.990:0.001:0.999], 100)

ans =

2.3642 2.4052 2.4506 2.5012 2.5589 2.6259 2.7064 2.8077 2.9464 3.1737

Ковариационная матрица

Ковариация, $\sigma_{xx} = \sigma_x^2 = D(x) = \overline{(x-\overline{x})^2}$:

$$\operatorname{cov}(x,y) = \sigma_{xy} = \sigma_{yx} = \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}(x-\overline{x})(y-\overline{y}) = \overline{(x-\overline{x})(y-\overline{y})}.$$

Ковариационная матрица для двух и M одинаковых величин ($\operatorname{cov}(X)$):

$$C_{xy} = \begin{pmatrix}\sigma_x^2 & \sigma_{xy}\\ \sigma_{yx} & \sigma_y^2\end{pmatrix},\quad C_{x_M} = \begin{pmatrix}\sigma_1^2 & \sigma_{12} & \sigma_{13} & \dots & \sigma_{1M}\\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \dots & \sigma_{2M}\\ \dots & \dots & \dots & \dots & \dots\\ \sigma_{M1} & \sigma_{M2} & \sigma_{M3} & \dots & \sigma_M^2\end{pmatrix},\quad C = C^T.$$

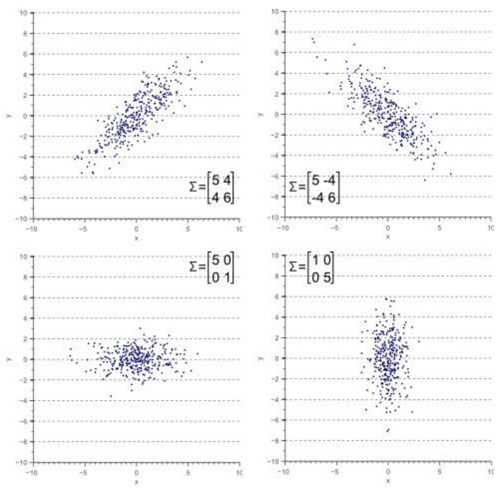
Ковариационная матрица

Пусть $\Delta x = \overline{x - \overline{x}}$, тогда C можно определить как $C = \overline{\Delta x \cdot \Delta x^T}$, где Δx – вектор-столбец. В общем случае $C_{xy} = \overline{\Delta x \Delta y^T}$ для векторов X и Y любой длины. **Свойства**:

- для независимых X и Y, $C_{XY} = C_X + C_Y$;
- $\mathrm{cov}(AX + B) = A\mathrm{cov}(X)A^T$ (A – произвольная квадратная матрица);
- $\mathrm{cov}(X, Y) = \mathrm{cov}(Y, X)^T$;
- $\mathrm{cov}(\sum X_i, Y) = \sum \mathrm{cov}(X_i, Y)$, $\mathrm{cov}(X, \sum Y_i) = \sum \mathrm{cov}(X, Y_i)$;
- для независимых X и Y $\mathrm{cov}(X, Y) = 0$.

Ковариационная матрица

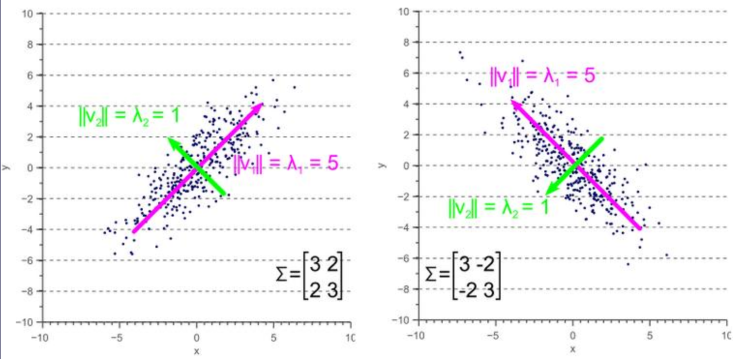
Вектор данных: $X = \binom{x_i}{y_i}$; $\Sigma = C_{xy}$. σ_x и σ_y характеризуют разброс данных по осям. σ_{xy} отражает линейную зависимость $y(x)$. В данном случае удобней было бы использовать **корреляционную матрицу**, где $\rho_{xy} = \overline{(x-\overline{x})(y-\overline{y})}\Big/\sigma_x\sigma_y$. При $\rho_{xy} = 0$ эллиптичное облако точек, при $\rho_{xy} = 1$ расположение точек вдоль отрезка.



Собственные значения ковариационной матрицы характеризуют дисперсию вдоль направления, заданного ее **собственными векторами**. $C_{xy}v = \lambda v$.

```
[v, lambda]=eig([3 2; 2 3], "vector")
```

Получили два собственных значения: 1 и 5, которым соответствуют вектора $v(1) = \begin{pmatrix}-\sqrt{2}/2\\ \sqrt{2}/2\end{pmatrix}$ и $v(5) = \begin{pmatrix}\sqrt{2}/2\\ \sqrt{2}/2\end{pmatrix}$.



Доверительный интервал

Доверительная вероятность

$p = P(X_0 \leqslant x \leqslant X_1)$

Математическое ожидание

Если известен закон распределения (мат. ожидание и дисперсия: μ и σ), то

$$P\left(<X> - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leqslant \mu \leqslant <X> + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha,$$

где z_α – α–квантиль нормального распределения
В Octave: `norminv(x)`. Например, для $1-\alpha = 0.95$, $1 - \frac{\alpha}{2} = 0.975$.

Доверительный интервал

α**–квантилем** называется число x_α: $P(X \leqslant x_\alpha) \geqslant \alpha$ и $P(X \geqslant x_\alpha) \geqslant 1 - \alpha$. Т.е. по интегральной функции распределения $F(x_\alpha) = \alpha$. А т.к. $P(a \leqslant X \leqslant b) = F(b) - F(a)$, получаем:

$$P(x_{\frac{1-\alpha}{2}} \leqslant X \leqslant x_{\frac{1+\alpha}{2}}) = \alpha.$$

Пример

В 64 наблюдениях получено: $S_1 = \sum x = 600$, $S_2 = \sum(x - \overline{x})^2 = 3800$. Вычислить 95% доверительный интервал матожидания.
Решение: $\sigma = \sqrt{S_2/(n-1)} = 7.72$; $<x> = S_1/n = 9.375$. $F(0.975) = 1.96$, отсюда найдем границы интервала $<x> \pm F(0.975)\sigma/\sqrt{n}$: $\overline{x} \in [7.484, 11.266]$ с точностью 95%.

Доверительный интервал

Математическое ожидание

Если закон распределения неизвестен, то

$$P\left(<X> - t_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}} \leqslant \mu \leqslant <X> + t_{1-\frac{\alpha}{2},n-1}\frac{S}{\sqrt{n}}\right) = 1-\alpha,$$

где S – несмещенный RMS. Величина

$$T = \frac{<X> - \mu}{S/\sqrt{n}}$$

имеет распределение Стьюдента, а $t_{\alpha,n-1}$ – его квантили.

Пример: $<X> = 10$, $S_n = 2$, $n = 11$ (10 степеней свободы), по таблице для двухстороннего распределения Стьюдента с вероятностью 95% $T_{10}^{95} = 2.228$. Тогда доверительный интервал есть $\overline{X} \pm TS_n/\sqrt{n}$, т.е. $\mu \in (8.6565, 11.3440)$. В Octave `t=tinv(0.975, 10)`, т.к. $1-\alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow 1-\frac{\alpha}{2} = 0.975$.

Доверительный интервал

Дисперсия

Если известно среднее, можно воспользоваться распределением χ^2.

$$P\left(\frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{(1+\alpha)/2,n}} \leqslant \sigma^2 \leqslant \frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{(1-\alpha)/2,n}}\right) = \alpha.$$

Если же среднее неизвестно, то

$$P\left(\frac{(n-1)S^2}{\chi^2_{(1+\alpha)/2,n-1}} \leqslant \sigma^2 \leqslant \frac{(n-1)S^2}{\chi^2_{(1-\alpha)/2,n-1}}\right) = \alpha.$$

Алгоритм обработки результатов измерений

Прямые измерения

1. Заполнить таблицу с результатами N измерений x_i.
2. Вычислить среднее арифметическое измеренной величины: $\overline{x} = \dfrac{1}{n}\sum x_i$.
3. Определить стандартный доверительный интервал:
$$S_x = \sqrt{\frac{1}{n(n-1)}\sum(x_i - \overline{x})^2}.$$
4. Задать значение коэффициента надежности α, и по нему определить значение коэффициента Стьюдента $t_{\alpha,N}$. По паспортным данным определить абсолютную погрешность измерительного прибора, $\Delta x_{\text{пр}}$. Если $\Delta x_{\text{пр}} > 4t_{\alpha,N}S_x$, представить результат в виде $x = \overline{x} \pm \Delta x_{\text{пр}}/2$, обработка окончена.
5. Если $\Delta x_{\text{пр}} < t_{\alpha,N}S_x$ считаем, что $\Delta x = t_{\alpha,N}S_x$; иначе вычисляем результирующую среднеквадратичную погрешность как $\Delta x = \sqrt{(t_{\alpha,N}S_x)^2 + (\Delta x_{\text{пр}}/2)^2}$.
6. Результат записываем как $x = \overline{x} \pm \Delta x, \quad \alpha = \ldots$

Алгоритм обработки результатов измерений

Косвенные измерения

- Вычислить для всех измеряемых величин среднее значение и погрешность прямого измерения. При этом для всех величин выбирается одно и то же значение доверительной вероятности α.
- По формуле вычислить среднее значение измеряемой величины: $\overline{w} = f(\overline{x}, \overline{y}, \cdots)$.
- Оценить погрешность косвенно измеряемой величины:

$$(\Delta w)^2 = \left(\frac{\partial w}{\partial x}\right)^2 \Bigg|_{\substack{x=\overline{x};\\ y=\overline{y};\\ \cdots}} (\Delta x)^2 + \left(\frac{\partial w}{\partial y}\right)^2 \Bigg|_{\substack{x=\overline{x};\\ y=\overline{y};\\ \cdots}} (\Delta y)^2 + \cdots .$$

- Записать результат в виде $w = \overline{w} \pm \Delta w, \quad \alpha = \ldots$.

Метод наименьших квадратов

Пусть имеется функция $f(x|a)$, зависящая от аргумента x и набора параметров a. Данной функции соответствует набор пар данных (x_n, y_n), причем $y_n = f(x_n|a) + \varepsilon_n$, где ε_n – случайная ошибка. Математическое ожидание ошибки $\overline{\varepsilon} = 0$, ее среднеквадратическое отклонение равно σ_n. Для оценки a (аппроксимации набора данных заданной функцией) необходимо минимизировать выражение

$$\Phi = \sum_{n=1}^{N} \left(\frac{y_n - f(x_n|a)}{\sigma_n^2} \right)^2.$$

При этом подразумевается, что число измерений превышает число параметров a.

Пример: линейная зависимость

Пусть $y = ax + b$, x_n известны с пренебрежимо малой погрешностью, y_n – результаты измерений с нормальным распределением, $\overline{y_i} = ax_i + b$.

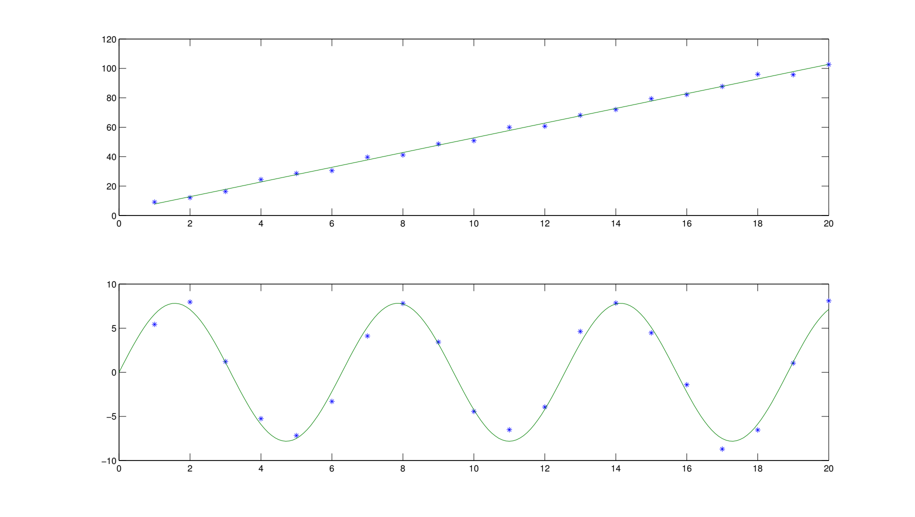
Минимизируем величину $Y = \sum(y_i - \overline{y_i})^2$, $\dfrac{\partial Y}{\partial a} = 0$, $\dfrac{\partial Y}{\partial b} = 0$:

$$a = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2} = \frac{\overline{xy} - \overline{x}\,\overline{y}}{\overline{x^2} - (\overline{x})^2},$$

$$b = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2} = \frac{\overline{x^2}\,\overline{y} - \overline{x}\,\overline{xy}}{\overline{x^2} - (\overline{x})^2}.$$

$$\sigma^2 = \frac{n}{n-2}\left(\overline{y^2} - (\overline{y})^2 - a^2\left[\overline{x^2} - (\overline{x})^2\right]\right), \qquad \sigma_a^2 = \frac{\sigma^2}{n\left(\overline{x^2} - (\overline{x})^2\right)}, \qquad \sigma_b^2 = \sigma_a^2\overline{x^2}.$$

Аппроксимация МНК



Аппроксимация МНК

Некоторые зависимости, можно свести к линейным. Например, $y = \mathrm{e}^{ax+b} \Longrightarrow \ln y = ax + b$.
Возможно также сведение зависимостей к системам линейных уравнений $A\vec{x} = \vec{b}$, ранг матрицы A должен быть больше количества искомых переменных. Минимизируем $(A\vec{x} - \vec{b})^T(A\vec{x} - \vec{b})$, что приводит к системе уравнений

$$A^T A\vec{x} = A^T\vec{b} \quad \Longrightarrow \quad \vec{x} = (A^T A)^{-1} A^T\vec{b}.$$

Или $\vec{x} = A^+\vec{b}$ (псевдообратная матрица), в Octave — «левое деление» $A\backslash b$.

Пример

Пусть некоторая величина изменяется по закону $y = a_0 + a_1\,\mathrm{e}^{-t} + a_2 t e^{-t}$. В матричном виде $Y = TA$, где T – функциональная матрица, у которой в первом столбце размещены единицы (соответствует умножению на a_0), во втором — функция e^{-t}, а в третьем — $t\,\mathrm{e}^{-t}$. Коэффициенты A найдем при помощи МНК: $A = T\backslash Y$.

```
t = [0 0.3 0.8 1.1 1.6 2.3]';
y = [0.6 0.67 1.01 1.35 1.47 1.25]';
T = [ones(size(t))  exp(-t)  t.*exp(-t)];
A = T\y
```

Ошибка МНК

МНК для линейной зависимости

Пусть наблюдаемая l имеет линейную зависимость от a, b и c:

$$l(t) = x \cdot a(t) + y \cdot b(t) + z \cdot c(t).$$

Из эксперимента получаем N наборов данных l_k, a_k, b_k и c_k:

$$l_k = xa_k + yb_k + zc_k + \Delta l_k.$$

Найдем x, y и z, минимизируя $S = \Delta l_k^2 = \sum\bigl(l_k - (xa_k + yb_k + zc_k)\bigr)^2$:

$$\frac{\partial S}{\partial \aleph} = \frac{\partial}{\partial \aleph} \sum\bigl(l_k - (xa_k + yb_k + zc_k)\bigr)^2 = 0$$

Ошибка МНК

Введем обозначения: $\sum_{k=1}^{N} \aleph_k \beth_k = [\aleph\beth]$. Тогда после дифференцирования получим систему из трех уравнений для нахождения трех неизвестных:

$$\begin{cases} x[aa] + y[ab] + z[ac] = [al], \\ x[ba] + y[bb] + z[bc] = [bl], \\ x[ca] + y[cb] + z[cc] = [cl]. \end{cases} \quad \text{или} \quad \begin{pmatrix} [aa] & [ab] & [ac] \\ [ba] & [bb] & [bc] \\ [ca] & [cb] & [cc] \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} [al] \\ [bl] \\ [cl] \end{pmatrix}.$$

Или $MK = V$, следовательно, $K = M^{-1}V$ (`K=M\V`). Аналогичную систему можно составить для погрешностей:

$$\begin{pmatrix} [aa] & [ab] & [ac] \\ [ba] & [bb] & [bc] \\ [ca] & [cb] & [cc] \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = \begin{pmatrix} [a\Delta l] \\ [b\Delta l] \\ [c\Delta l] \end{pmatrix} \quad \text{или} \quad M\Delta K = \Delta V.$$

Ошибка МНК

Итак, для погрешностей имеем: $\Delta K = M^{-1}\Delta V$. Если наблюдения — равноточные и независимые, ковариационная матрица ошибок диагональна:

$$C_L = \begin{pmatrix}\Delta l_1\\ \Delta l_2\\ \dots\\ \Delta l_N\end{pmatrix}\begin{pmatrix}\Delta l_1 & \Delta l_2 & \dots & \Delta l_N\end{pmatrix} = \sigma_0^2\begin{pmatrix}1 & 0 & 0 & \cdots & 0\\ 0 & 1 & 0 & \cdots & 0\\ \dots & \dots & \dots & \dots & \dots\\ 0 & 0 & 0 & \cdots & 1\end{pmatrix} = \sigma_0^2 E_N.$$

Аналогично построим ковариационную матрицу ошибок неизвестных:

$$C_K = \begin{pmatrix}\sigma_x^2 & \sigma_{xy} & \sigma_{xz}\\ \sigma_{yx} & \sigma_y^2 & \sigma_{yz}\\ \sigma_{zx} & \sigma_{zy} & \sigma_z^2\end{pmatrix} = \Delta K\cdot\Delta K^T = M^{-1}\Delta V\cdot\Delta V^T(M^{-1})^T$$

Т.к. $[\aleph\beth] = [\beth\aleph]$, матрица M симметрична ($M = M^T$).

Ошибка МНК

Члены ковариационной матрицы $\Delta V \cdot \Delta V^T$ на примере одного:

$$<[a\Delta l][a\Delta l]> = <\sum a_i \Delta l_i \sum a_j \Delta l_j> = <\sum\sum a_i a_j \Delta l_i \Delta l_j>,$$

т.е. $\Delta V \cdot \Delta V^T = \sum\sum a_i a_j <\Delta l_i \Delta l_j>$, а т.к. $<\Delta l_i \Delta l_j>$ равны нулю при $i \neq j$ и равны σ_0^2 при $i = j$, получим: $\sigma_0^2 \sum a_k^2 = \sigma_0^2 [aa]$. И в итоге:

$$C_K = M^{-1}\sigma_0^2 M M^{-1} = \sigma_0^2 M^{-1}!$$

Вывод: обратная матрица нормальных уравнений является матрицей весов вектора неизвестных.
Для получения несмещенной оценки σ_0^2 воспользуемся формулой:

$$\sigma_0^2 = \frac{\sum\Big(l_k - (x a_k + y b_k + z c_k)\Big)^2}{N - M},$$

где M – число неизвестных (в нашем случае — три).

По набору данных получить коэффициенты линейной зависимости и определить их погрешности.

A	1	0	-1	2	3	-2	0	4
B	0	3	2	-1	2	-1	3	1
C	2	-2	0	1	-2	3	-2	0
L	7	1	3	2	1	6	1	5

```
A=[1 0 -1 2 3 -2 0 4]'; B=[0 3 2 -1 2 -1 3 1]'; C=[2 -2 0 1 -2 3 -2 0]';
L=[7 1 3 2 1 6 1 5]'; T=[A B C]; % T - матрица данных, T*K=L
K=T \ L % искомые коэффициенты
V=T' * L; Mr=K/V; % M*K = V, Mr->M^(-1)
v=L-T*K; sigma0 = sqrt(sum(v.^2)/(8-3));
DK = sigma0 * sqrt(diag(Mr)) % искомые погрешности
for i=1:3;
printf("K%d=%.2f+-%.2f (%.1f%%)\n", i, K(i), DK(i), 100*DK(i)/K(i));
endfor
K1=0.72+-0.06 (7.7%)
K2=2.29+-0.07 (3.3%)
K3=3.06+-0.14 (4.5%)
```

Спасибо за внимание!

mailto

eddy@sao.ru
edward.emelianoff@gmail.com