24 *Практические методы вычисления МНК-оценок и их средних квадратических ошибок*

24.1 ***Метод определителей***

Этот метод применяют на практике, если число неизвестных невелико и не превосходит трех. Метод ориентирован прежде всего на “ручные вычисления” с использованием микрокалькуляторов. В основе метода лежат известные формулы Крамера.

Пусть система трех уравнений имеет вид

$$[aa]x + [ab]y + [ac]z = [al],$$

$$[ba]x + [bb]y + [bc]z = [bl],$$

$$[ca]x + [cb]y + [cc]z = [cl].$$

Вычислим определитель системы

$$D = \begin{Vmatrix} [aa] & [ab] & [ac] \\ [ba] & [bb] & [bc] \\ [ca] & [cb] & [cc] \end{Vmatrix}.$$

Заменяя последовательно первый, второй и третий столбцы матрицы нормальных уравнений на столбец правой части, вычислим определители

$$D_x = \begin{Vmatrix} [al] & [ab] & [ac] \\ [bl] & [bb] & [bc] \\ [cl] & [cb] & [cc] \end{Vmatrix}, \quad D_y = \begin{Vmatrix} [aa] & [al] & [ac] \\ [ba] & [bl] & [bc] \\ [ca] & [cl] & [cc] \end{Vmatrix}, \quad D_z = \begin{Vmatrix} [aa] & [ab] & [al] \\ [ba] & [bb] & [bl] \\ [ca] & [cb] & [cl] \end{Vmatrix}.$$

Теперь неизвестные определяются следующим образом:

$$x = \frac{D_x}{D}, \qquad y = \frac{D_y}{D}, \qquad z = \frac{D_z}{D}.$$

Вычислением неизвестных не заканчивается метод наименьших квадратов. Необходимо вычислить средние квадратические ошибки этих неизвестных.

Весовые коэффициенты суть диагональные элементы q_{11}, q_{22}, q_{33} обратной матрицы системы нормальных уравнений, т.е.

$$\begin{pmatrix} [aa] & [ab] & [ac] \\ [ba] & [bb] & [bc] \\ [ca] & [cb] & [cc] \end{pmatrix} \cdot \begin{pmatrix} \underline{q_{11}} & q_{12} & q_{13} \\ q_{21} & \underline{q_{22}} & q_{23} \\ q_{31} & q_{32} & \underline{q_{33}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Для определения q_{11} нужно заменить первый столбец матрицы нормальных уравнений на первый столбец правой части, вычислить определитель полученной матрицы и поделить на D:

$$q_{11} = \frac{1}{D} \begin{vmatrix} 1 & [ab] & [ac] \\ 0 & [bb] & [bc] \\ 0 & [cb] & [cc] \end{vmatrix} = \frac{[bb] \cdot [cc] - [cb]^2}{D}.$$

Для определения q_{22} нужно заменить второй столбец матрицы нормальных уравнений на второй столбец правой части и выполнить аналогичные вычисления:

$$q_{22} = \frac{1}{D}\begin{Vmatrix} [aa] & 0 & [ac] \\ [ba] & 1 & [bc] \\ [ca] & 0 & [cc] \end{Vmatrix} = \frac{[aa]\cdot[cc]-[ac]^2}{D} .$$

Поступая точно также, получим

$$q_{33} = \frac{1}{D}\begin{Vmatrix} [aa] & [ab] & 0 \\ [ba] & [bb] & 0 \\ [ca] & [cb] & 1 \end{Vmatrix} = \frac{[aa]\cdot[bb]-[ab]^2}{D} .$$

Теперь нужно определить квадрат средней квадратической ошибки единицы веса

$$\varepsilon_0^2 = \frac{1}{n-3}\sum_{k=1}^{n}\left[l_k - \left(a_k x + b_k y + c_k z\right)\right]^2 ,$$

где *x, y, z* - значения наших неизвестных, полученные из решения системы нормальных уравнений.

Последнюю формулу можно преобразовать так, что необходимость вычисления остаточных разностей типа $\upsilon_k = l_k - \left(a_k x + b_k y + c_k z\right)$ отпадет. Запишем $\sum \upsilon_k^2$ следующим образом

$$\sum\left(l_k - a_k x - b_k y - c_k z\right)\left(l_k - a_k x - b_k y - c_k z\right) =$$
$$= \sum l_k\left(l_k - a_k x - b_k y - c_k z\right) - x\sum\left(a_k l_k - a_k a_k x - a_k b_k y - c_k a_k z\right) -$$
$$= y\sum\left(b_k l_k - b_k a_k x - b_k l_k y - b_k c_k z\right) - z\sum\left(c_k l_k - c_k a_k x - c_k b_k y - c_k c_k z\right).$$

Используя скобки Гаусса, перепишем полученный результат

$$[\upsilon\upsilon] = [ll] - \left([la]x + [lb]y + [lc]z\right) - x\left[[al] - \left([aa]x + [ab]y + [ac]z\right)\right] -$$
$$- y\left[[bl] - \left([ba]x + [bb]y + [bc]z\right)\right] - z\left[[cl] - \left([ca]x + [cb]y + [cc]z\right)\right].$$

Легко видеть, что последние три выражения в квадратных скобках равны нулю, так как *x, y, z* удовлетворяют системе нормальных уравнений. Отсюда следует

$$[\upsilon\upsilon] = [ll] - \left([la]x + [lb]y + [lc]z\right) .$$

Поскольку $[al],[bl],[cl]$ мы уже вычислили, составляя нормальные уравнения, осталось довычислить сумму квадратов правых частей исходных уравнений (наблюдений) $[ll]$. Результатом обработки данных будут значения *x, y, z* и их ошибки:

$$x \pm \sqrt{\frac{[\upsilon\upsilon]}{n-3}}\cdot\sqrt{q_{11}} , \qquad y \pm \sqrt{\frac{[\upsilon\upsilon]}{n-3}}\cdot\sqrt{q_{22}} , \qquad z \pm \sqrt{\frac{[\upsilon\upsilon]}{n-3}}\cdot\sqrt{q_{33}} .$$

24.2 <u>*Метод последовательных исключений (схема Гаусса)*</u>

Этот метод можно применять к любому количеству неизвестных. Схема, предложенная Гауссом первоначально, была рассчитана для ручного счета с использованием специальных вычислительных бланков. Схема предусматривает:

1. составление нормальных уравнений с параллельным контролем правильности вычислений,
2. последовательное исключение неизвестных, начиная с первого. Вычисления сопровождаются контролем,
3. параллельное вычисление весовых коэффициентов,
4. определение неизвестных, начиная с последнего,
5. вычисление суммы квадратов остаточных разностей, вычисление ошибок неизвестных.

Несколько слов о терминологии.

Исходные уравнения, надлежащие математической обработке косвенных наблюдений методом наименьших квадратов, в астрономической традиции называют условными уравнениями. Этот термин заимствован из задач геодезии, где МНК применяется довольно широко. Однако, использование его в астрономической практике не имеет основания. Дело в том, что в геодезии дело имеют с линейными уравнениями, в которых на неизвестные накладываются дополнительные условия. Например, если *x, y, z* - есть три угла треугольника, то какими бы они ни были, должно выполняться равенство $x+y+z=180^0$. В противном случае отрезки прямых не образуют замкнутую геометрическую фигуру: образуется невязка - несовпадение первой и последней точки замкнутого треугольника (или полигона). Термин невязка также применяется в астрономической практике, но несколько в ином смысле: это разность (*O-C*):

$$\upsilon_k = y_k - \left(a_k x + b_k y + c_k z\right).$$

Во избежание недоразумений, величины υ_k следует называть остаточными разностями, т.е. так же, как их называют в англоязычной литературе (residual).

Для иллюстрации схемы Гаусса ограничимся всего тремя неизвестными, хотя ее можно распространить на любое их количество.

Составляется таблица исходных уравнений:

x	*y*	*z*	**l**	**s**	**Ĩ**	
a_1	b_1	c_1	l_1	s_1		v_1
a_2	b_2	c_2	l_2	s_2		v_2
a_3	b_3	c_3	l_3	s_3		v_3
a_4	b_4	c_4	l_4	s_4		v_4
...	...	...	...	...	...	...
a_n	b_n	c_n	l_n	s_n		v_n

В таблицу исходных уравнений помещают :

1) коэффициенты перед неизвестными a_k, b_k, c_k;
2) правые части l_k несовместных “уравнений” $a_k x + b_k y + c_k z = l_k$;
3) контрольные суммы $s_k = a_k + b_k + c_k + d_k$;

4) значения $\hat{l}_k = a_k\hat{x} + b_k\hat{y} + c_k\hat{z}$. Этот столбец вычисляется после того, как неизвестные будут определены;

5) остаточные разности $\upsilon_k = l_k - \hat{l}_k$ в последнем столбце таблицы.

Вычисление контрольных сумм (пятый столбец) необходимо в случае ручного счета (на микрокалькуляторе). Эти суммы предохранят от возможных промахов при составлении нормальных уравнений и их решении. Однако, числа a_k, b_k, c_k и l_k должны быть выписаны с одинаковым числом знаков после запятой. В случае, когда они определены иначе, прибегают к приему “выравнивания” порядков коэффициентов соответствующим выбором неизвестных. Поясним сказанное на примере. Пусть

$$634.8x - 1.2080y + 0.002385z = 1.709.$$

Как видно из этого уравнения, все числа четырехзначные. Следовательно, можно пользоваться четырехзначными вычислениями. Следует ожидать, что и решения будут иметь четыре значащих цифры . Однако, вычисление суммы

$$s = 634.8 - 1.2080 + 0.002385 + 1.709$$

не имеет смысла, т.к. первое число имеет только одну цифру после запятой, а третье - шесть. Дополнение “нужных мест” нулями (что возможно) приведет к резкому увеличению числа знаков для s

$$s = 634.800000 - 1.208000 + 0.002385 + 1.709000 = 635.303385.$$

Мы получим девятизначную контрольную сумму. Чтобы этого избежать, перепишем уравнение следующим образом:

$$\frac{634.8}{1000}(1000x) - 1.2080y + 100\cdot(0.002385)\frac{z}{100} = 1.709.$$

Пусть $1000x = x'$, $y = y'$, $0.01z = z'$. Теперь уравнение принимает вид

$$0.6348x' - 1.2080y' + 0.2385z' = 1.7090.$$

Вычисление контрольных сумм не встречает затруднений:

$$s = 0.6348 - 1.2080 + 0.2385 + 1.7090 = 1.3743.$$

Процедуру составления нормальных уравнений, решения и вычисления весовых коэффициентов можно изобразить следующей схемой, в которой столбцы помечены малыми латинскими буквами, а строки пронумерованы:

Схема Гаусса

	a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*
1	$[aa]$	$[ab]$	$[ac]$	$[al]$	$[as]$	Σ_1	1		
2	$[ba]$	$[bb]$	$[bc]$	$[bl]$	$[bs]$	Σ_2	0		
3	$[ca]$	$[cb]$	$[cc]$	$[cl]$	$[cs]$	Σ_3	0		
4	1	$\frac{[ab]}{[aa]}$	$\frac{[ac]}{[aa]}$	$\frac{[al]}{[aa]}$	$\frac{[as]}{[aa]}$	Σ_4	$\frac{1}{[aa]}$		
5		$[bb.1]$	$[bc.1]$	$[bl.1]$	$[bs.1]$	Σ_5	*	1	

6		$[cb.1]$	$[cc.1]$	$[cl.1]$	$[cs.1]$	$\sum_6$	*	0	
7		1	$\frac{[bc.1]}{[bb.1]}$	$\frac{[bl.1]}{[bb.1]}$	$\frac{[bs.1]}{[bb.1]}$	$\sum_7$	*	$\frac{1}{[bb.1]}$	
8			$[cc.2]$	$[cl.2]$	$[cs.2]$	$\sum_8$	*	*	1
9			1	$\frac{[cl.2]}{[cc.2]}$	$\frac{[cs.2]}{[cc.2]}$	$\sum_9$	*	*	$\frac{1}{[cc.2]}$

Комментарии к схеме Гаусса.

<u>1,2,3 строки.</u>

Столбцы “*a*”,”*b*”,”*c*”,”*d*” занимают нормальные уравнения, полученные из исходных уравнений перемножением соответствующих столбцов и суммированием произведений.

Столбец “*e*” вычислен по той же самой схеме: перемножением s_k с a_k, b_k и c_k:

$[as]=\sum a_k s_k$, $\qquad [bs]=\sum b_k s_k$, $\qquad [cs]=\sum c_k s_k$.

Контролем правильности вычислений являются суммы столбца “*f*” :

$$\sum\nolimits_1=[aa]+[ab]+[ac]+[al]=[as],$$
$$\sum\nolimits_2=[ba]+[bb]+[bc]+[bl]=[bs],$$
$$\sum\nolimits_3=[ca]+[cb]+[cc]+[cl]=[cs].$$

Столбцы ”*g*”,”*h*”,”*i*” предназначены для вычисления весовых коэффициентов q_{11}, q_{22}, q_{33} (q_x, q_y, q_z).

<u>4 строка</u>

есть уравнение исключения (элиминационное уравнение). Оно получено путем деления <u>первого нормального уравнения</u> на <u>первый коэффициент</u>. Этой процедуре подвергаются все числа, стоящие в первой строке, в том числе $[as]$ (столбец “*e*”) и 1 (столбец ”*g*”), исключая столбец “*f*”, который выполняет контрольные функции

$$\sum\nolimits_4=1+\frac{[ab]}{[aa]}+\frac{[ac]}{[aa]}+\frac{[al]}{[aa]}=\frac{[as]}{[aa]}.$$

<u>5,6 строки</u>

занимает вторая промежуточная система нормальных уравнений, полученная после исключения неизвестного *x*:

$$[bb.1]y+[bc.1]z=[bl.1],$$
$$[cb.1]y+[cc.1]z=[cl.1].$$

Исключение неизвестного *x* выполняется следующим образом. Уравнение исключения нужно умножить на коэффициент при *x* и вычесть из нормального уравнения:

$$[bb.1]=[bb]-[ba]\frac{[ab]}{[aa]}, \qquad [cb.1]=[cb]-[ac]\frac{[ab]}{[aa]},$$
$$[bc.1]=[bc]-[ba]\frac{[ac]}{[aa]}, \qquad [cc.1]=[cc]-[ca]\frac{[ac]}{[aa]}.$$

Точно также поступают со всеми столбцами схемы, за исключением столбца “f”, который служит контролем правильности выкладок:

$$\sum\nolimits_5 = [bb.1]+[bc.1]+[bl.1]=[bs.1],$$

$$\sum\nolimits_6 = [cb.1]+[cc.1]+[cl.1]=[cs.1].$$

Полученная система двух уравнений с двумя неизвестными является системой нормальных уравнений со всеми ее свойствами. Во-первых, ее матрица симметрическая $[bc.1]=[cb.1]$, поэтому не все коэффициенты этого уравнения нужно вычислять. Многие вычислители их даже не выписывают. Во-вторых, эта матрица является весовой по отношению к неизвестным y и z. Поэтому вычисления можно не доводить до конца, а на определенном этапе воспользоваться, например, методом определителей.

7 строка.

Второе элиминационное уравнение. Снова первое нормальное уравнение (промежуточной системы) разделим на первый коэффициент, стоящий перед неизвестной, подлежащей исключению на втором этапе вычислений.

8 строка.

Умножая второе элиминационное уравнение на коэффициент $[cb.1]$ (второе нормальное уравнение) и вычитая из уравнения строки (6), получим снова “систему” нормальных уравнений без неизвестных y:

$$[cc.2]=[cc.1]-[cb.1]\frac{[bc.1]}{[bb.1]}.$$

Правая часть этого уравнения равна

$$[cl.2]=[cl.1]-[cb.1]\frac{[bl.1]}{[bb.1]}.$$

Столбец “e” дает $$[cs.2]=[cs.1]-[cb.1]\frac{[bs.1]}{[bb.1]}.$$

Вычисления верны, если $\sum_8=[cs.2]$.

Заполняется последний столбец “i” схемы: ставится очередная единица.

9 строка -

новое элиминационное уравнение, которое по сути есть решение нормального уравнения (8), которое содержит лишь одно неизвестное

$$z=\frac{[cl.2]}{[cc.2]}.$$

Соблюдая правила, мы должны вычислить все числа строки 9, в том числе и для суммового контроля.

Неизвестные вычисляют из элиминационных уравнений:

$$z=\frac{[cl.2]}{[cc.2]},$$

$$y=\frac{[bl.1]}{[bb.1]}-z\frac{[bc.1]}{[bb.1]},$$

$$x = \frac{[al]}{[aa]} - z\frac{[ac]}{[aa]} - y\frac{[ab]}{[aa]}.$$

Для вычисления весовых коэффицентов используются столбцы ”*g*”,”*h*”,”*i*”. Существует следующее правило: для вычисления *q* необходимо перемножить числа, стоящие в строках первых нормальных уравнений с числами элиминационных строк и результаты сложить.

$q_x = (1)\cdot(4) + (5)\cdot(7) + (8)\cdot(9)$ - столбец ”*g*”,

$q_y = (5)\cdot(7) + (8)\cdot(9)$ - столбец ”*h*”,

$q_z = (8)\cdot(9)$ - столбец ”*i*”.

Результат вычислений записывают так:

$$x = \hat{x} \pm \varepsilon_0 \sqrt{q_x}\ ,$$

$$y = \hat{y} \pm \varepsilon_0 \sqrt{q_y}\ ,$$

$$z = \hat{z} \pm \varepsilon_0 \sqrt{q_z}\ .$$

Численный пример решения системы линейных уравнений по схеме Гаусса, который мы приводим ниже, повторяет тот же пример, которым мы иллюстрировали метод определителей.

x	y	z	l	s	$\tilde{l}$	υ
1	0	2	7	10	6.8	+0.2
0	3	-2	1	2	0.8	+0.2
-1	2	0	3	4	3.9	-0.9
2	-1	1	2	4	2.2	-0.2
3	2	-2	1	4	0.6	+0.4
-2	-1	3	6	6	5.4	+0.6
0	3	-2	1	2	0.8	+0.2
4	1	0	5	10	5.2	-0.2
					$\sum\upsilon$	+0.3
					$\sum\upsilon^2$	1.53

$$\varepsilon_0 = \sqrt{\frac{1.53}{8-3}} = 0.55$$

$\boxed{35}$	8	-8	19	54	54	1		
8	29	-20	11	28	28	0		
-8	-20	26	28	26	26	0		
1	0.228	-0.228	0.543	1.543	1.543	0.0286		
	$\boxed{27.2}$	-18.2	6.7	15.7	15.7	-0.229	1	
	-18.2	24.2	32.3	38.3	38.3	0.229	0	
	1	-0.669	0.246	0.577	0.577	-0.0084	0.0368	
		$\boxed{12.0}$	36.8	48.8	48.8	0.0761	0.670	1
		1	3.07	4.07	4.07	0.0063	0.0558	0.0833
z=	3.07							
y=	2.30				q	0.031	0.074	0.083
x=	0.72				ε	0.097	0.150	0.158

x=0.72±0.10, y=2.30±0.15, z=3.07±0.16

6.4.3. *Метод Холецкого*

Этот способ решения системы линейных уравнений может использоваться в любых случаях, не только для решения нормальных уравнений. Он легко программируется и предназначен, в основном, для машинного счета.

Основная идея метода. Допустим, что мы должны решить систему уравнений, заданную в матричной форме

$$\mathbf{A}\boldsymbol{x}=\boldsymbol{y},$$

где **A** - квадратная матрица (m х m), $\boldsymbol{x}$ - вектор неизвестных, $\boldsymbol{y}$ - известный вектор.

Метод основан на разложении квадратной матрицы (условие симметричности необязательно) на две треугольные:

$$\boldsymbol{A}=\boldsymbol{BC},$$

$$\mathbf{B}=\begin{pmatrix} b_{11} & 0 & ... & 0 \\ b_{21} & b_{22} & ... & 0 \\ ... & ... & ... & ... \\ b_{m1} & b_{m1} & ... & b_{mm} \end{pmatrix}, \qquad \mathbf{C}=\begin{pmatrix} 1 & c_{12} & ... & c_{1m} \\ 0 & 1 & ... & c_{2m} \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{pmatrix}.$$

Тогда $\boldsymbol{BCx}=\boldsymbol{y}$.

Обозначим $\boldsymbol{Cx}=\boldsymbol{z}$.

Получим $\boldsymbol{Bz}=\boldsymbol{y}$.

Теперь вместо одной мы будем иметь две системы уравнений:

$$\begin{array}{ll} b_{11}z_1 = y_1 & x_1 + c_{12}x_2 + c_{13}x_3 + \ldots + c_{1m}x_m = z_1, \\ b_{21}z_1 + b_{22}z_2 = y_2 & x_2 + c_{23}x_3 + \ldots + c_{2m}x_m = z_2, \\ b_{31}z_1 + b_{32}z_2 + b_{33}z_3 = y_3 & x_3 + \ldots + c_{3m}x_m = z_3, \\ \ldots\ldots\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\ b_{m1}z_1 + b_{m2}z_2 + \ldots + b_{nm}z_m = y_m & x_m = z_m. \end{array}$$

Из первой системы последовательно определяются вспомогательные неизвестные $z_1, z_2, \ldots, z_m$. После этого из второй системы в обратном порядке определяются искомые $x_m, x_{m-1}, \ldots, x_1$.

При решении системы нормальных уравнений нам необходимо определить обратную матрицу этой системы, так как она нужна для вычисления весовых коэффициентов неизвестных. При обращении матрицы A необходимо столбец y заменить на m столбцов единичной матрицы и решить систему m раз.

Разложение матрицы A на две треугольные производят следующим образом. Имеем тождество:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{pmatrix} = \begin{pmatrix} b_{11} & 0 & \cdots & 0 \\ b_{21} & b_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ b_{m1} & b_{m1} & \cdots & b_{mm} \end{pmatrix} \cdot \begin{pmatrix} 1 & c_{12} & \cdots & c_{1m} \\ 0 & 1 & \cdots & c_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Выполняя произведение матриц в правой части тождества и приравнивая соответствующие результаты произведений элементам матрицы A, получим последовательно элементы матриц B и C.

Вычисления облегчаются, если придерживаться следующей последовательности операций:

1. строки B x 1-й столбец C $\overset{A}{\Rightarrow}$ 1 столбец B,
2. 1 строка B x столбцы C $\Rightarrow$ 1 строка C,
3. строки B x 2-й столбец C $\Rightarrow$ 2 столбец B,
4. 2 строка B x столбцы C $\Rightarrow$ 2 строка C,
5. строки B x 3-й столбец C $\Rightarrow$ 3 столбец B,
6. 3 строка B x столбцы C $\Rightarrow$ 3 строка C,

. . . .

Для решения системы линейных уравнений метод Холецкого выгоднее метода последовательных исключений, причем, чем больше число неизвестных, тем преимущество становится больше.

<u>Численный пример.</u>

Решить систему нормальных уравнений, используя метод Холецкого:

$$\begin{pmatrix} 35 & 8 & -8 \\ 8 & 29 & -20 \\ -8 & -20 & 26 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 19 \\ 11 \\ 28 \end{pmatrix}.$$

Разложим систему уравнений на две с треугольными матрицами и решим их.

	B			C		z	x
35			1	0.2286	-0.2286	0.5429	0.720
8	27.2			1	-0.6680	0.2447	2.293
-8	-18.2	12.0			1	3.066	3.066

Вычислим матрицу весовых коэффициентов. Обозначим через Z матрицу вспомогательных неизвестных, а через Q - матрицу весовых коэффициентов. Получим

$$Z = \begin{pmatrix} 0.0286 & 0 & 0 \\ -0.0084 & 0.0368 & 0 \\ 0.0063 & 0.0558 & 0.0833 \end{pmatrix} \qquad Q = \begin{pmatrix} 0.0310 & -0.0042 & 0.0063 \\ -0.0042 & 0.0741 & 0.0556 \\ 0.0063 & 0.0556 & 0.0833 \end{pmatrix}$$

Правильность вычисления контролирует симметричность матрицы Q.

6.5 Теорема Гаусса-Маркова

Теорема Гаусса- Маркова утверждает, что МНК-оценка является наилучшей в том смысле, что её ковариационная матрица является наименьшей. Однако, как понимать, что одна матрица меньше или больше другой? Какие матрицы можно сравнивать?

Квадратная матрица $\mathbf{C_1}$ больше другой квадратной матрицы $\mathbf{C_2}$, если разность $\mathbf{C_1}$-$\mathbf{C_2}$ есть положительно определённая матрица. Квадратная матрица $\mathbf{C}$ является положительно определённой, если для произвольного вектора $\mathbf{y}$ имеет место неравенство $\mathbf{y}^\mathbf{T}\mathbf{C}\mathbf{y}>0$.

Вернёмся к нашим уравнениям. Если исходные уравнения заданы следующим образом

$$\mathbf{l} = \mathbf{A}\mathbf{x} + \Delta\mathbf{l},$$

то МНК-оценка вектора параметров **x,** как мы видели, может быть записана следующим образом

$$\hat{\mathbf{x}} = (\mathbf{A}^\mathbf{T}\mathbf{P}\mathbf{A})^{-1}\mathbf{A}^\mathbf{T}\mathbf{P}\mathbf{l},$$

где **P** – матрица весов наблюдений. Обозначим

$$\mathbf{B_0} = (\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\mathbf{A}^\mathbf{T}\mathbf{P},$$

теперь несмещённую МНК-оценку параметров получим с помощью равенства

$$\hat{\mathbf{x}} = \mathbf{B_0 l} = \mathbf{B_0}(\mathbf{Ax} + \Delta\mathbf{l})$$

Для того, чтобы эта оценка была несмещённой необходимо, чтобы выполнялись условия

$$\langle\hat{\mathbf{x}}\rangle = \mathbf{B_0 Ax} + \mathbf{B_0}\langle\Delta\mathbf{l}\rangle.\ \mathbf{B_0 A} = \mathbf{E},\ \ \langle\Delta\mathbf{l}\rangle = \mathbf{0},$$

другими словами ошибки наблюдений должны быть случайными, не содержать систематической составляющей, а матрица $\mathbf{B_0}$ должна быть псевдообратной по отношению к матрице **A,** что для уравнений МНК имеет место.

Предположим, что существует другой линейный алгоритм вычисления несмещённой оценки вектора **x.** Для этого алгоритма

$$\hat{\hat{\mathbf{x}}} = \mathbf{B_1 l} = \mathbf{B_1}(\mathbf{Ax} + \Delta\mathbf{l})$$

Снова убеждаемся, что для условия несмещённости матрица $\mathbf{B_1}$ должна быть псевдообратной по отношению к матрице **A.** Отсюда следует, что

$$(\mathbf{B}_1 - \mathbf{B}_0)\mathbf{A} = \mathbf{0}$$

Введём обозначение для ковариационной матрицы погрешностей наблюдений

$$\mathbf{R} = \langle\Delta\mathbf{l}\Delta\mathbf{l}^\mathbf{T}\rangle,$$

тогда ковариационные матрицы погрешностей оценок погрешностей оценок $\hat{\mathbf{x}}$ и $\hat{\hat{\mathbf{x}}}$ будут

$$\mathbf{C}_{\Delta\mathbf{x}} = \langle(\hat{\mathbf{x}} - \mathbf{x})(\hat{\mathbf{x}} - \mathbf{x})^\mathbf{T}\rangle = \mathbf{B_0 R B_0}^T \quad \text{и аналогично} \quad \mathbf{C}'_{\Delta\mathbf{x}} = \langle(\hat{\hat{\mathbf{x}}} - \mathbf{x})(\hat{\hat{\mathbf{x}}} - \mathbf{x})^\mathbf{T}\rangle = \mathbf{B_1 R B}_1^T.$$

Образуем разность, и выполним небольшие преобразования

$$\mathbf{C}'_{\Delta\mathbf{x}} - \mathbf{C}_{\Delta\mathbf{x}} = \mathbf{B}_1\mathbf{R}\mathbf{B}_1^T - \mathbf{B}_0\mathbf{R}\mathbf{B}_0^T = \left[\mathbf{B}_0 + (\mathbf{B}_1 - \mathbf{B}_0)\right]\mathbf{R}\left[\mathbf{B}_\mathbf{0}^\mathbf{T} + (\mathbf{B}_1^T - \mathbf{B}_\mathbf{0}^\mathbf{T})\right] - \mathbf{B}_0\mathbf{R}\mathbf{B}_\mathbf{0}^\mathbf{T} =$$

$$= (\mathbf{B}_1 - \mathbf{B}_0)\mathbf{R}(\mathbf{B}_1 - \mathbf{B}_0)^T + \mathbf{B}_0\mathbf{R}(\mathbf{B}_1^T - \mathbf{B}_0^T) + (\mathbf{B}_1 - \mathbf{B}_0)\mathbf{R}\mathbf{B}_\mathbf{0}^\mathbf{T}.$$

Однако, $(\mathbf{B}_1 - \mathbf{B}_0)\mathbf{R}\mathbf{B}_\mathbf{0}^\mathbf{T} = (\mathbf{B}_1 - \mathbf{B}_0)\mathbf{R}\left[\mathbf{PA}(\mathbf{A}^T\mathbf{PA})^{-1}\right] = 0,$ так как $\mathbf{RP} = \sigma_\mathbf{0}^2\mathbf{E_n}$, а $(\mathbf{B}_1 - \mathbf{B}_0)\mathbf{A} = \mathbf{0}.$ Матрица $\mathbf{B_0 R}(\mathbf{B}_1^T - \mathbf{B}_0^T)$ также равна нулю, так как она является транспонированной к только что рассмотренной.

Итак, $\mathbf{C}'_{\Delta\mathbf{x}} - \mathbf{C}_{\Delta\mathbf{x}} = (\mathbf{B}_1 - \mathbf{B}_0)\mathbf{R}(\mathbf{B}_1 - \mathbf{B}_0)^T.$

Полученная матрица неотрицательно определена. Из теории случайных векторов известно, что ковариационная матрица случайного вектора положительно определена, так как она является преобразованием случайного шума наблюдений линейным оператором $(\mathbf{B}_1 - \mathbf{B}_0)$. Следовательно $\mathbf{C}'_{\Delta\mathbf{x}} > \mathbf{C}_{\Delta\mathbf{x}}$. Минимальное значение ковариационная матрица погрешностей оценивания принимает только при $\mathbf{B_1} = \mathbf{B_0}$ **,** что и доказывает теорему.

6.5.1 Метод максимального правдоподобия

Является ли МНК-оценка наилучшей, наиболее достоверной? В теории вероятности разработан так называемый метод максимального правдоподобия. Как узнать, что данная оценка является более правдоподобной, чем другая? Необходим критерий. Один из критериев – дисперсия разброса экспериментальных точек на кривой, аппроксимирующей искомую зависимость. Это, кажется, в подавляющем числе случаев устраивает экспериментаторов и наблюдателей. Но не все учёные безоговорочно принимают этот критерий. Теорема Гаусса-Маркова утверждает, что метод наименьших квадратов обеспечивает лишь минимальную дисперсию погрешностей оценивания в случае оценки случайной величины и минимальную ковариационную матрицу при оценке случайного вектора. Но будет ли такая оценка более правдоподобной, чем какая-то другая оценка?

Метод максимального правдоподобия имеет дело с так называемой функцией правдоподобия, которая зависит как от компонент вектора наблюдений $\mathbf{l}(l_1, l_2, ...l_n)$, так и от компонент вектора параметров $\mathbf{x}(x_1, x_2, ...x_m)$. Для получения **наиболее правдоподобных** оценок параметров $\tilde{x}_1(l_1,l_2,..l_n)$, $\tilde{x}_2(l_1,l_2,l_3,...l_n)$, … $\tilde{x}_m(l_1,l_2,l_3,...l_n)$ решают систему уравнений $\dfrac{\partial L(l_1,l_2,...l_n;\tilde{x}_1,\tilde{x}_2...\tilde{x}_m)}{\partial \tilde{x}_k}=0$, k=1,2,3…m. , где $L(\mathbf{l},\tilde{\mathbf{x}})$ - функция правдоподобия. В качестве функции правдоподобия часто берут совместную плотность вероятности невязок. В частности, когда ошибки оценок распределены по нормальному закону, то есть

$$f(\Delta\mathbf{x})=\frac{1}{\sqrt{(2\pi)^m\det\mathbf{C}_{\Delta\mathbf{x}}}}\exp\left\{-\frac{1}{2}\Delta\mathbf{x}^{\mathbf{T}}\mathbf{C}_{\Delta\mathbf{x}}^{-1}\Delta\mathbf{x}\right\},$$

максимум этой функции достигается при соблюдении принципа Гаусса-Лежандра. Таким образом, метод наименьших квадратов обеспечивает максимальное правдоподобие оценок при условии, что погрешности наблюдений удовлетворяют нормальному закону распределения.

Уместно задать вопрос. А если ошибки наблюдений, не подчиняются нормальному закону распределения? Метод наименьших квадратов для оценки параметров не годится? Некоторые учёные именно так и думают. Однако нельзя забывать, что теорема Гаусса-Лежандра выполняется! То есть МНК обеспечивает наименьший разброс точек около эмпирической кривой, независимо от закона распределения. Оценки МНК обеспечивают наименьшую дисперсию погрешности. Если это вполне устраивает наблюдателя, то можно не заботится о законе распределения.

6.6. Определение параметров нелинейных функций посредством МНК

Рассмотрим функции с одним аргументом и *m* параметрами. Будем считать, что аргументом является время *t*, а параметрами $x_1,x_2,\ldots,x_m$. В дискретные моменты t_k имеем наблюдения этой функции l_k, содержащие погрешности Δl_k. Таким образом, имеем

$$l_1=\varphi(t_1;x_1,\ldots,x_m)+\Delta l_1,$$
$$l_2=\varphi(t_2;x_1,\ldots,x_m)+\Delta l_2,$$
$$\cdots\cdots\cdots\cdots\cdots$$
$$l_n=\varphi(t_{n1};x_1,\ldots,x_m)+\Delta l_n.$$

Используя критерий L_2 (сумму квадратов остаточных разностей), определим параметры $x_1,x_2,\ldots,x_m$:

$$L_2=\sum_{k=1}^{n}\Big[l_k-\varphi\big(t_k;x_1,\ldots,x_m\big)\Big]^2=\min.$$

В случае неравноточных независимых наблюдений критерий L_2 модифицируется

$$L_2=\sum_{k=1}^{n}p_k\Big[l_k-\varphi\big(t_k;x_1,\ldots,x_m\big)\Big]^2=\min.$$

Нормальными уравнениями будут

$$\frac{\partial L_2}{\partial x_j}=0,\qquad j=\overline{1,m}.$$

Эти уравнения нелинейны. Минимизация функционала L_2 в случае числа неизвестных более двух приводит к очень громоздким вычислениям. Обычно в этом случае прибегают к линеаризации исходных уравнений и к использованию методики МНК.

6.6.1. ***Линеаризация и решение***

Предположим, что какие-то приближенные значения параметров x_j нам известны. Обозначим их через $\bar{x}_j$. Тогда

$$x_j = \bar{x}_j + \delta x_j, \quad j = \overline{1, m},$$

где δx_j - искомая поправка. Будем также считать, что эта поправка достаточно мала, чтобы мы имели право заменить точное значение функции на приближенное, ограничиваясь разложением Тейлора до первой степени δx_j:

$$\varphi\left(t_k; \bar{x}_1 + \delta x_1, \bar{x}_2 + \delta x_2, \ldots, \bar{x}_m + \delta x_m\right) \approx$$

$$\approx \varphi\left(t_k; \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m\right) + \delta x_1 \frac{\partial \varphi(t_k)}{\partial x_1} + \delta x_2 \frac{\partial \varphi(t_k)}{\partial x_2} + \quad \ldots \quad + \delta x_m \frac{\partial \varphi(t_k)}{\partial x_m} =$$

$$= \varphi\left(t_k; \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m\right) + \sum_{j=1}^{m} \delta x_j \frac{\partial \varphi(t_k)}{\partial x_j}.$$

Обозначим

$$\delta l_k = l_k - \varphi\left(t_k; \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m\right),$$

$$a_{kj} = \frac{\partial \varphi\left(t_k; \bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m\right)}{\partial x_j}.$$

Получим систему линейных уравнений

$$\delta l_k = a_{k1}\delta x_1 + a_{k2}\delta x_2 + \ldots + a_{km}\delta x_m + \Delta l_k, \qquad k = \overline{1, n}.$$

Коэффициенты a_{kj} образуют матрицу **A**:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix}.$$

Поправки к параметрам и вариации δl_k изобразим в векторном виде:

$$\delta\mathbf{x} = \begin{pmatrix} \delta x_1 \\ \delta x_2 \\ \ldots \\ \delta x_m \end{pmatrix}, \qquad \Delta\mathbf{l} = \begin{pmatrix} \delta l_1 \\ \delta l_2 \\ \ldots \\ \delta l_n \end{pmatrix}.$$

Теперь для определения наилучшей в среднеквадратическом смысле оценки вектора параметров $\delta\mathbf{x}$ воспользуемся системой линейных нормальных уравнений, которая в матричной форме имеет вид

$$\mathbf{A}^T \mathbf{A} \delta\mathbf{x} = \mathbf{A}^T \delta\mathbf{l},$$

или $$\mathbf{A}^T \mathbf{P} \mathbf{A} \delta\mathbf{x} = \mathbf{A}^T \mathbf{P} \delta\mathbf{l}$$ в случае неравноточных наблюдений.

Процедура решения и вычисления ковариационной матрицы погрешностей оценки поправок остается такой же, что и для “классических” линейных уравнений.

После определения $\delta\tilde{x}_j$ следует уточнить параметры

$$x_j = \bar{x}_j + \delta x_j.$$

СКО для $\tilde{x}_j$ остается той же, что и для $\delta\tilde{x}_j$.

При решении нелинейных уравнений возникает следующая трудность. Представим себе, что предварительный набор параметров был выбран неудачно, так что некоторые поправки оказались слишком велики. Тогда замена нелинейной функции линейной относительно поправок может быть недостаточно строгой. В этом случае процедуру вычисления МНК-оценок поправок к параметрам следует повторить, т.е. вычислить вторую итерацию. За приближенные значения параметров берут МНК-оценки, вычисленные на первой итерации.

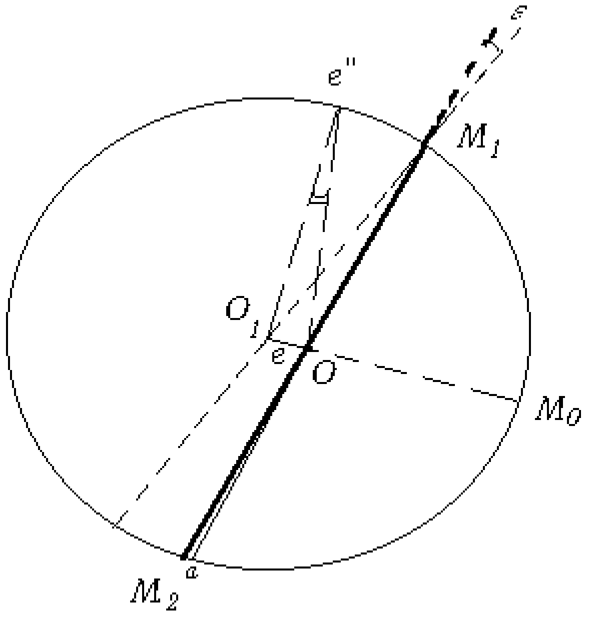
Повторяя всю процедуру составления исходных и нормальных уравнений, мы получим новые значения параметров.

Может оказаться, что и второй итерации недостаточно. Тогда за **x** принимают приближенные значения и вычисляют поправки на третьей итерации и т.д.. Процедура продолжается до тех пор, пока СКО оценок не перестанут уменьшаться, а поправки не будут превышать по модулю этих средних квадратических ошибок.

Типичный пример применения МНК к нелинейным функциям - задача улучшения орбит по заданным наблюдениям координат этих небесных тел.

Численный пример

Рассмотрим решение задачи исследования круга астрономического инструмента. Требуется определить параметры, устанавливающие связь между отсчетами двух микроскопов M_1 и M_2, отличающиеся от 180^0 из-за эксцентриситета круга и других параметров, которые определяют разность



$$M_2 - M_1 = 180^0 + a + 2e'' \sin(M_1 - M_0),$$

где a - постоянная часть, вызванная неточностью установки микроскопов, e'' - эксцентриситет круга (см. рис.), M_0 - отсчет лимба, который лежит на одной прямой, соединяющей центр окружности лимба (O_1) и ось вращения (O). Параметры a, e'', M_0 являются неизвестными, которые нужно определить методом наименьших квадратов.

Поворотом круга будем устанавливать первый микроскоп на отсчетах $0^0, 30^0, 60^0, \ldots$ через 30^0 и будем считывать со второго микроскопа M_2. Таким образом, наблюдениями в данной задаче будут разности отсчетов этих микроскопов.

В приведенной ниже таблице приведены значения M_1 (первый столбец), M_2 (второй столбец), отличие от 180^0 их разности и “центрированная” разность - приведенная к среднему значению (L'').

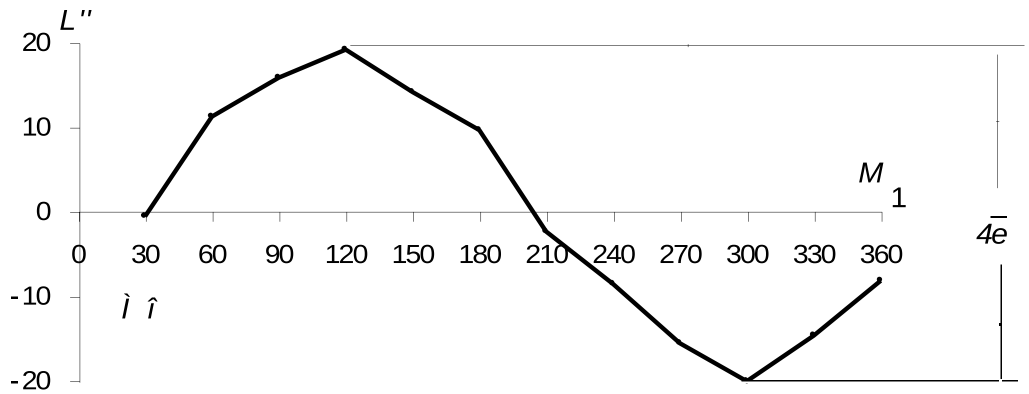
M_1	M_2	$M_2 - M_1 - 180^0$	L''
30^0	$210^{0\ 1}$’16.7’’	76.7’’	-0.6’’
60^0	$240^0$1’28.5’’	88.5’’	11.2’’
90^0	$270^0$1’33.1’’	93.1’’	15.8’’
120^0	$300^0$1’36.4’’	96.4’’	19.1’’
150^0	$330^0$1’31.5’’	91.4’’	14.1’’
180^0	$0^0$1’26.9’’	86.9’’	9.6’’
210^0	$30^0$1’14.9’’	74.9’’	-2.4’’
240^0	$60^0$1’08.7’’	68.7’’	-8.6’’
270^0	$90^0$1’01.7’’	61.7’’	-15.6’’
300^0	$120^0$0’57.2’’	57.2’’	-20.1’’
330^0	$150^0$1’02.6’’	62.6’’	-14.7’’
360^0	$180^0$1’09.1’’	69.1’’	-8.2’’

ср.зн. 77.3’’

Здесь $L'' = M_2 - M_1 - 180^0 - 77.3''$

Из основной формулы следует, что $L'' = a - 77.3'' + 2e'' \sin(M_1 - M_0)$.

Для того, чтобы линеаризовать данное выражение, нужно определить предварительные значения параметров $\bar{a}, \bar{e}, \overline{M_0}$. Это легко сделать графически. Построим график зависимости L'' от M_1.



Из графика определяем $\bar{a} = 77.3''$, $\bar{e} = 9.4''$, $\overline{M_0} = 27^0$.

Теперь запишем уравнение в вариациях. Пусть $e = \bar{e} + x$, $M_0 = \overline{M}_0 + y$, $z = a - 77.3''$.

$$L'' = z + 2(\bar{e} + x)\sin(M_1 - \overline{M}_0 - y) \approx$$
$$\approx z + 2\bar{e}\sin\left(M_1 - M_0\right) + 2x\sin\left(M_1 - \overline{M}_0\right) - 2y\bar{e}\cos\left(M_1 - \overline{M}_0\right).$$

Обозначим $l_k = L'' - 2\bar{e}\sin\left(M_1 - M_0\right)_k$,

$$a_k = 2\sin\left(M_1 - \overline{M}_0\right)_k,$$

$$b_k = -2\bar{e}\cos\left(M_1 - \overline{M}_0\right)_k,$$

$$c_k = 1.$$

Теперь имеем исходное уравнение в виде $l_k = a_k x + b_k y + c_k z + \Delta l_k$.

Заметим, что Δl_k - ошибка наблюдений. Она содержится в установке лимба и отсчете микроскопа. Строго говоря, эта ошибка входит и в M_1 и в M_2. Для того чтобы МНК был строгим, мы высказываем гипотезу о том, что ошибку содержит только M_2 (т.е. L''), а величина M_1, которая входит в формулу вычисления коэффициентом, - безупречна. Теперь имеем

M_1	a_k	b_k	c_k	l_k
30	0.1047	-18.774	1	-1.6
60	1.0593	-15.707	1	+1.0
90	1.7820	-8.535	1	-1.0
120	1.9973	+0.984	1	+0.3
150	1.6773	10.239	1	-1.7
180	0.9080	16.751	1	+1.1
210	-0.1047	18.774	1	-1.4
240	-1.0893	15.767	1	+1.6
270	-1.7820	8.535	1	+1.2
300	-1.9973	-0.984	1	-1.3
330	-1.6773	-10.239	1	+1.1
360	-0.9080	-16.751	1	+0.3

Нормальные уравнения имеют вид

$$\begin{pmatrix} 24.000 & 0.000 & 0.000 \\ 0.000 & 2120.6 & 0.000 \\ 0.000 & 0.000 & 12.000 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -5.37 \\ 18.30 \\ -0.40 \end{pmatrix}.$$

Поскольку матрица нормальных уравнений диагональная, получим

$$x = -\frac{5.37}{24.0} = -0.22, \qquad p_x = 24.0,$$

$$y = \frac{18.30}{2120.6} = 0.0086, \qquad p_y = 2120.6,$$

$$z = -\frac{0.40}{12.0} = -0.033, \qquad p_z = 12.0.$$

Сумму квадратов остаточных уклонений вычислим по формуле

$$[\upsilon\upsilon] = [ll] - ([al]x + [bl]y + [cl]z) = 16.35.$$

Поэтому

$$\varepsilon_0 = \sqrt{\frac{16.35}{12-3}} = 1.35''.$$

Таким образом,

$$x = -0.22 \pm \frac{1.35}{\sqrt{24}} = -0.22'' \pm 0.28'',$$

$$y = 0.0086 \pm 0.029(\text{р€Љ}) = 0.49^0 \pm 1.66^0,$$

$$z = -0.033 \pm \frac{1.35}{\sqrt{12}} = -0.033'' \pm 0.39''.$$

Следовательно,

$$\begin{cases} e = 9.4'' - 0.2'' \pm 0.28'' = 9.2'' \pm 0.3'' \\ M_0 = 27^0 + 0.5^0 \pm 1.7^0 = 27.5^0 \pm 1.7^0 \\ a = 77.3'' - 0.033'' \pm 0.39'' = 77.3'' \pm 0.4'' \end{cases}$$

Из-за того, что поправки получились меньше по абсолютному значению ошибок, вторая итерация в данном случае не требуется.

6.5.4. <u>*Численный пример (второй вариант)*</u>

Вернемся снова к только что рассмотренному численному примеру. Запишем основное уравнение следующим образом:

$$M_2 - M_1 = 180^0 + a + 2e''\sin(M_1 - M_0) =$$
$$= 180^0 + a + 2e''\sin M_1 \cos M_0 - 2e''\cos M_1 \sin M_0 \, .$$

Введем обозначения

$$x = 2e''\cos M_0, \qquad l_k = \left(M_2 - M_1\right)_k - 180^0 - 77.3'',$$
$$y = -2e''\sin M_0, \qquad a_k = \sin M_1,$$
$$z = a - 77.3'', \qquad b_k = \cos M_1,$$
$$c_k = 1 .$$

Теперь исходные уравнения линейны. Коэффициенты этих уравнений и правая часть приведены в таблице

M_1	a_k	b_k	c_k	l_k
30	0.500	0.866	1	-0.6
60	0.866	0.500	1	11.2
90	1.000	0.000	1	15.8
120	0.866	-0.500	1	19.1
150	0.500	-0.866	1	14.1
180	0.000	-1.000	1	9.6
210	-0.500	-0.866	1	-2.4
240	-0.866	-0.500	1	-8.6
270	-1.000	0.000	1	-15.6
300	-0.866	0.500	1	-20.6
330	-0.500	0.866	1	-14.7
360	0.000	1.000	1	-8.2

Нормальные уравнения снова имеют диагональную форму

$$\begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 12 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 97.8 \\ -50.9 \\ -0.4 \end{pmatrix},$$

$$x = \frac{97.8}{6} = 16.3'', \qquad p_x = 6,$$

$$y = -\frac{50.9}{6} = -8.5'', \qquad p_y = 6, \qquad [\upsilon\upsilon] = 15.92,$$

$$z = -\frac{0.40}{12.0} = -0.033'', \qquad p_z = 12, \qquad \varepsilon_0 = \sqrt{\frac{15.92}{12-3}} = 1.33''.$$

Следовательно,

$$x = 16.3'' \pm 0.5'',$$
$$y = -8.5'' \pm 0.5'',$$
$$z = 0.0 \pm 0.4'' .$$

Теперь нужно вычислить искомые e'', M_0 и a и их ошибки. Как следует из приведенных формул

$$e = \frac{1}{2}\sqrt{x^2 + y^2} = 9.2'',$$

$$M_0 = -arctg\frac{y}{x} = 27.5^0,$$

$$a = z + 77.3'' = 77.3''.$$

Остается вычислить ошибки

$$4e^2 = x^2 + y^2,$$

$$2 \cdot 4e\Delta e = 2x\Delta x + 2y\Delta y.$$

Поскольку матрица нормальных уравнений диагональна, ковариация между Δx и Δy равна нулю. Следовательно,

$$4e\varepsilon_e = \sqrt{x^2\varepsilon_x^2 + y^2\varepsilon_y^2},$$

$$\varepsilon_e = \frac{1}{4e}\sqrt{x^2\varepsilon_x^2 + y^2\varepsilon_y^2} = \frac{0.5''\sqrt{x^2 + y^2}}{4e''} = \frac{0.5''}{2} = 0.25'',$$

$$tgM_0 = -\frac{y}{x},$$

$$\frac{\Delta M_0}{\cos^2 M_0} = -\frac{\Delta y}{x} + \frac{y}{x}\frac{\Delta x}{x},$$

$$\frac{\varepsilon_{M_0}}{\cos^2 M_0} = \frac{1}{x}\sqrt{\varepsilon_y^2 + \left(\frac{y}{x}\right)^2 \varepsilon_x^2} = \frac{2e\varepsilon_x}{x},$$

$$\varepsilon_{M_0} = 2\frac{e''\varepsilon_x}{x^2}\cos^2 M_0 \text{р}ad = 2\frac{0.5 \cdot 9.2}{(16.3)^2}\cos^2 27.3 = 0.02734\text{р}ad = 1.6^0.$$

Итак,

$$e = 9.2'' \pm 0.3'',$$

$$M_0 = 27.5^0 \pm 1.6^0,$$

$$a = 77.3'' \pm 0.4''.$$

Полученный результат практически не отличается от того, который мы имели в предыдущем случае.

Обработка ряда наблюдений

В одном из разделов мы уже рассматривали два варианта задачи определения параметра X по заданным наблюдениям. Причём, в одном случае все наблюдения были равноточны , в другом – нет. Основными условиями определения оценки мы брали минимум дисперсии оценки и её несмещённость. Обоим этим требованиям удовлетворяет оценивание по методу наименьших квадратов. Попробуем получить основные формулы для обработки ряда наблюдений, используя метод наименьших квадратов.

Итак, пусть вектору наблюдений $\mathbf{l}(l_1, l_2, \ldots l_n)$ соответствуют веса $(p_1, p_2 \ldots p_n)$. Требуется определить параметр x, который связан с наблюдениями простым равенством $l_k = x + \Delta l_k$, причём дисперсия ошибки k- го наблюдения известна, и равна $\sigma_k^2 = \frac{\sigma_0^2}{p_k}$. Весовая матрица вектора наблюдений, в этом случае, имеет диагональных вид $\mathbf{P} = \sigma_0^2 \text{diag}\, \sigma_k^{-2}$. Запишем сходные уравнения в матричной форме

$$\mathbf{l} = x + \Delta\mathbf{l}, \text{ или}$$

$$\begin{pmatrix} 1 \\ 1 \\ \cdots \\ 1 \end{pmatrix} x + \begin{pmatrix} \Delta l_1 \\ \Delta l_2 \\ \cdots \\ \Delta l_n \end{pmatrix} = \begin{pmatrix} l_1 \\ l_2 \\ \cdots \\ l_n \end{pmatrix}$$

Нормальные уравнения с учётом весов, в этом случае, принимают вид

$$(1 \quad 1 \quad \cdots \quad 1) \begin{pmatrix} \sigma_1^{-2} & 0 & \cdots & 0 \\ 0 & \sigma_2^{-2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \sigma_n^{-2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \cdots \\ 1 \end{pmatrix} \tilde{x} = (1 \quad 1 \quad \cdots \quad 1) \begin{pmatrix} \sigma_1^{-2} & 0 & \cdots & 0 \\ 0 & \sigma_2^{-2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \sigma_n^{-2} \end{pmatrix} \begin{pmatrix} l_1 \\ l_2 \\ \cdots \\ l_n \end{pmatrix}$$

Выполнив, указанные матричные операции, получим

$$\sum_{k=1}^{n} \sigma_k^{-2} \hat{x} = \sum_{k=1}^{n} \sigma_k^{-2} l_k \ .$$

Умножим обе части равенства на σ_0^2, и воспользовавшись обозначением $p_k = \dfrac{\sigma_0^2}{\sigma_k^2}$, будем

иметь $\hat{x} = \dfrac{\sum_{k=1}^{n} p_k l_k}{\sum_{k=1}^{n} p_k}$, что полностью совпадает с полученной ранее нами формулой.

В случае равноточных наблюдений $p_k = 1$ и оценка параметра совпадает с арифметическим

средним $\hat{x} = \dfrac{\sum_{k=1}^{n} l_k}{n}$.

Наконец, нужно получить ещё выборочную (апостериорную) дисперсию единицы веса. согласно изложенному, эта дисперсия в МНК равна

$\varepsilon_0^2 = \dfrac{\sum_{k=1}^{n} p_k (l_k - \hat{x})^2}{n-1}$, что также совпадает с формулой, полученной в упомянутом разделе.

Приведённый пример показывает, формулы для обработки ряда равноточных и неравноточных наблюдений могут быть получены как частный случай метода наименьших квадратов.