УДК 528.1:519.92

Л. К. ВОЙСЛАВСКИЙ

О КОЛИЧЕСТВЕ ИНФОРМАЦИИ, ПОЛУЧАЕМОЙ ПРИ ОБРАБОТКЕ ГЕОДЕЗИЧЕСКИХ ИЗМЕРЕНИЙ

Ошибки, возникающие при геодезических измерениях, порождают некоторую неопределенность результатов. Мерой этой неопределенности будет остаточная (апостериорная) энтропия [3]. Поскольку наряду с необходимыми выполняют, как правило, также избыточные измерения, в процессе обработки часть остаточной энтропии будет устранена, и мы получим дополнительную информацию об измеренной величине.

Пусть l_i ($i=1, 2, \ldots, n$) — результаты непосредственных равноточных измерений некоторой величины L. Если предположить, что систематические ошибки отсутствуют, а случайные ошибки $\Delta_i (i=1, 2, \ldots, n)$ имеют нормальное распределение с параметрами 0, σ, остаточная энтропия может быть представлена в виде [2]

$$H(l) = \log_a \left(\frac{\sqrt{2\pi e}\,\sigma}{\delta}\right), \tag{1}$$

где σ — стандарт, зависящий от условий, в которых выполнены измерения; δ — интервал разрешения шкалы или отсчетного устройства измерительного прибора.

Основание логарифма Q в (1) определим при выборе единицы измерения дополнительной информации. Величины σ и δ должны быть согласованы, так как, следуя гипотезе аддитивной структуры [5], случайная ошибка Δ_i порождается множеством элементарных ошибок, в том числе и ошибкой, обусловленной интервалом разрешения (округления). Поэтому между стандартом σ, характеризующим точность измерений в целом, и стандартом, характеризующим только ошибки округления, наблюдается зависимость

$$\sigma_{ок} = \alpha\sigma, \quad (\alpha < 1). \tag{2}$$

Если отсчет по шкале дает значение величины l_i (или ее части), как например, при измерении линий, то ошибки округления можно считать распределенными равномерно в интервале $-\frac{\delta}{2}$, $+\frac{\delta}{2}$. Тогда

$$\sigma_{ок} = \frac{\delta}{\sqrt{12}}. \tag{2а}$$

Когда результат l_i получен как разность отсчетов, например при измерении горизонтального угла или превышения нивелиром, для ошибок округления более подходящим будет треугольное распределение Симпсона с параметрами 0, δ [1]. При этом

$$\sigma_{ок} = \frac{\delta}{\sqrt{6}}. \tag{2б}$$

Подставляя (2а) и (2б) в (2), получим необходимые условия согласования

$$\delta = \alpha\sigma\sqrt{12}, \ \delta = \alpha\sigma\sqrt{6}, \quad (\alpha<1). \tag{3}$$

Остаточная энтропия вероятнейшего значения $l_0 = \frac{[l_i]}{n}$ записывается в виде

$$H(l_0) = \log_a \left(\frac{\sqrt{2\pi e}\, \sigma_0}{\delta_0} \right), \tag{4}$$

где $\sigma_0 = \frac{\sigma}{\sqrt{n}}$ — стандарт арифметической середины; δ_0 — интервал округления арифметической середины.

Разность энтропий $H(l)$ и $H(l_0)$ дает количество информации, получаемое в результате вычисления арифметической середины из ряда равноточных измерений величины L. Допустим, что $\delta_0 = \delta$. Тогда

$$J = H(l) - H(l_0) = \frac{1}{2} \log_a n. \tag{5}$$

Следовательно, в данном случае количество информации — функция числа выполненных измерений.

Чтобы найти вероятнейшее значение l_0, необходимо выполнить, как минимум, два измерения. Проффесор С. Хаусбрандт называет систему, состоящую только из двух измерений, простейшей [9]. Поэтому целесообразно, на наш взгляд, принять в данных условиях за единицу количество информации, получаемое при обработке простейшей системы из двух измерений, $\frac{1}{2} \log_a 2 = 1$ е. д. и. (единица дополнительной информации). Откуда находим

$$a = \sqrt{2},$$

$$\frac{1}{2} \log_{\sqrt{2}} n = \log_2 n. \tag{6}$$

1 е. д. и. $= \frac{1}{2}$ дв. ед. (бит).

С учетом (6) выражение (5) принимает вид

$$J = H(l) - H(l_0) = \log n \text{ е. д. и.} \tag{7}$$

Здесь и в дальнейшем символ log означает двоичный логарифм.

Предположим теперь, что для определения величины L выполнен ряд неравноточных измерений в условиях, характеризуемых стандартами σ_i $(i = 1, 2, \ldots, n)$. Если подобрать произвольную постоянную $c > 0$ так, чтобы для весов измерений выполнялось условие

$$p_1 + p_2 + \cdots + p_n = 1, \tag{8}$$

остаточная энтропия, обусловленная случайными ошибками, принимая во внимание (6), будет равна

$$H(l) = p_1 \log 2\pi e \sigma_1^2 + p_2 \log 2\pi e \sigma_2^2 + \ldots + p_n \log 2\pi e \sigma_n^2 - \log \delta^2.$$

Из (8) следует, что стандарт единицы веса в данном случае будет равен стандарту арифметической середины σ_0. Будет очевидным также равенство

$$\sigma_i^2 = \frac{\sigma_0^2}{p_i}.$$

Поэтому, учитывая (8), можем записать

$$H(l) = \log 2\pi e\sigma_0^2 - (p_1 \log p_1 + p_2 \log p_2 + \cdots + p_n \log p_n) - \log \delta^2. \quad (9)$$

Принимая $\delta_0 = \delta$, на основании формул (4), (6), (9) получим

$$J = -p_1 \log p_1 - p_2 \log p_2 - \cdots - p_n \log p_n. \quad (10)$$

Чтобы определить величины $-p_i \log p_i$, можно использовать таблицы, приведенные в [2] или [8], где обозначим

$$-p_i \log p_i = \eta(p_i).$$

После этого (10) можно представить в виде

$$J = \sum_{i=1}^{n} \eta(p_i).$$

Из неравенства

$$\sum_{i=1}^{n} \eta(p_i) \leqslant \log n, \quad (10а)$$

следует, что при фиксированном значении σ_0 дополнительная информация из общей арифметической середины всегда меньше или, в крайнем случае, равна дополнительной информации, полученной из простой арифметической середины.

При выводе формул (7) и (10) было принято, что $\delta_0 = \delta$. Если число измерений n достаточно велико, можно допустить, чтобы $\delta_0 < \delta$. Однако это приведет к потере дополнительной информации, так как чем точнее мы хотим определить вероятнейшее значение l_0, тем большую степень неопределенности нужно будет устранить. Поэтому, допуская $\delta_0 < \delta$, необходимо поставить условие, чтобы количество дополнительной информации было по крайней мере не меньше, чем в случае простейшей системы из двух измерений, то есть

$$H(l) - H(l_0) \geqslant \log 2.$$

Решив это неравенство относительно δ_0 с учетом формул (1), (4) и (9), получим

$$\log \frac{\delta_0}{\delta} \geqslant \frac{1}{2} \log 2 - \frac{1}{2} \sum_{i=1}^{n} \eta(p_i) \geqslant \frac{1}{2} \log \frac{2}{n}.$$

В результате найдем

$$\delta_0 \geqslant \delta \sqrt{\frac{2}{n}}. \quad (11)$$

Формула (11) может служить критерием при выборе интервала округления арифметической середины.

Выражения (7), (10) будут справедливы и тогда, когда результаты измерений наряду со случайной Δ_i содержат также постоянную систематическую ошибку θ, так как энтропия случайной величины

$$l_i = L + \theta + \Delta_i$$

зависит только от стандарта σ и не зависит от ее математического ожидания $L + \theta$ [2]. Данный вывод согласуется с результатами исследований В. Н. Зимовнова [4].

Для определения остаточной энтропии результатов измерений, содержащих переменные систематические ошибки θ_i ($i = 1, 2, \ldots, n$), необходимо знать распределение вероятностей этих ошибок. Как известно, при фиксированном стандартном отклонении σ максимум энтропии соответствует нормальному распределению [7]. Поэтому остаточная энтропия, порождаемая переменными систематическими ошибками, в силу самого характера этих ошибок всегда будет меньше энтропии, обусловленной случайными ошибками, имеющими нормальное распределение. При этом имеется в виду, что и случайные, и систематические ошибки имеют одинаковую меру рассеивания.

Другим источником получения дополнительной информации об измеренных величинах является уравновешивание по способу наименьших квадратов. Для весов непосредственных измерений до и после уравновешивания в среднем будет справедливо отношение [6]

$$\frac{P}{p}=\frac{n}{k}=\frac{n}{n-r}, \tag{12}$$

где k — число определяемых неизвестных; r — число избыточных измерений.

Остаточные энтропии непосредственно измеренных величин, выраженные в е. д. и., будут соответственно равны:

а) до уравновешивания

$$H(l)=\log\left(\frac{2\pi e\sigma^2_{\text{изм}}}{\delta^2}\right); \tag{13}$$

б) после уравновешивания

$$H(l_{\text{ур}})=\log\left(\frac{2\pi e\sigma^2_{\text{ур}}}{\delta^2}\right). \tag{14}$$

Так как веса обратно пропорциональны квадратам стандартов, исходя из выражений (12), (13), (14), среднее количество дополнитель-

Порядок величин J и J′ для различных геодезических построений

Виды геодезических построений	Среднее количество информации: на непосредственное измерение J	Среднее количество информации: на определяемую величину J'
Разомкнутый полигонометрический ход с числом сторон:		
а) 3	0,81	1,42
б) 5	0,46	0,63
в) 10	0,22	0,26
Триангуляция:		
цепь из 20 треугольников между двумя жесткими сторонами, уравненная за все условия;	0,74	1,23
то же из 10 треугольников;	0,91	1,71
центральная система из 6 треугольников;	0,85	1,53
геодезический четырехугольник;	1,00	2,00
сплошная сеть треугольников;	1,56	4,6
Трилатерация:		
центральная система из шести треугольников;	1,12	0,13
геодезический четырехугольник;	0,26	0,31
сплошная сеть треугольников;	0,7	1,1

ной информации, полученное в результате уравновешивания, в расчете на одно непосредственное измерение будет равно

$$J = H(l) - H(l_{\text{ур}}) = \log \frac{n}{k} = \log \frac{n}{n-r}. \quad (15)$$

Отношения (12) выражают число непосредственных измерений, которое приходится в среднем на одну определяемую величину. Соответственно среднее количество информации в расчете на определяемую величину составит

$$J' = \frac{n}{k} \log \frac{n}{k} = \frac{n}{n-r} \log \frac{n}{n-r}. \quad (16)$$

В таблице показано, что количество информации, полученное в процессе уравновешивания, оказывается очень малым даже по сравнению с простейшей системой, состоящей только из двух равноточных измерений. Тем не менее выражения (15) и (16), характеризующие жесткость геодезических систем, могут служить мерой при сравнении различных конструктивных схем геодезических сетей.

Для несвободных сетей значения J и J' будут меньше, чем в формулах (15) и (16), за счет увеличения энтропии $H(l_{\text{ур}})$, обусловленной ошибками исходных данных. Увеличение энтропии $H(l_{\text{ур}})$ встречаем и при замене строгих способов уравновешивания упрощенными.

ЛИТЕРАТУРА

1. Бородачев Н. А. [и др.]. Точность производства в машиностроении и приборостроении. М., 1973.
2. Вентцель Е. С. Теория вероятности. М., 1962.
3. Елисеев С. В. Потери информации в процессе измерения углов и расстояний. — «Изв. вузов. Геодезия и аэрофотосъемка», 1966, вып. 3.
4. Зимовнов В. Н. Способ наименьших квадратов в приложении к измерениям, сопровождающимся постоянными погрешностями. М., 1960.
5. Кемниц Ю. В. Математическая обработка зависимых результатов измерений. М., 1970.
6. Чеботарев А. С. Способ наименьших квадратов с основами теории вероятностей. М., 1958.
7. Шеннон К. Э. Работы по теории информации и кибернетике. М., 1963.
8. Яглом А. М., Яглом И. М. Вероятность и информация. М., 1973.
9. Hausbrandt S. Rachunek wyrównawczy i obliczenia geodezyjne, t. 2. Warszawa, 1971.

Работа поступила 5 мая 1974 года. Рекомендована кафедрой геодезии и графики Херсонского сельскохозяйственного института.