For IUCr electronic archive

Optimization and Planning of Experiment in Single Crystal Diffractometry with 0-dimensional Detector

Yu. Nekrasov

Institute of Crystallography, Russian Academy of Sciences
59 Leninsky prosp. Moscow 117333, Russia

An optimizing algorithm of data collection is described. The algorithm ensures optimum quality of measurements at the set survey rate and distinguishes by the flexible distribution of time between first and second measurements of reflection, by efficiency and simplicity.

1.Introduction

A standard data collection optimization (Grant, 1973) consists in varying the second (main) measurement duration from the results of a first (estimating) measurement. This leads to approximately equal accuracies for strong and moderate reflections. Unfortunately the losses in device productivity or in quality of the experiment (because of constant estimating measurement duration unconnected with the mean intensity of reflections) and the difficulties in planning the experiment (because of unpredicted repeated measurements quantity) are inevitable.

These shortcomings are overcomed in the statistical method of optimization (Nekrasov, 1987a). This method was realized by us in 1978 in a control program for the Syntex $P2_1$ diffractometer. With its aid we investigated hundreds of protein crystals, as well as tens of organic and mineral crystals. The method ensured a given quality of the experiment (in the sense of statistical accuracy of measurements) without loss of the productivity of the diffractometer in all cases in which the mean level of scattering permits this to be done in an admissible time.

However the possibility of arbitrary distribution of time between the first measurement and the second one was absent. Moreover too many parameters were used.

This work represents a development of the statistical optimization method. An improved algorithm free from these shortcomings is described.

2. Method

Data collection is carried out under the control of optimizing program. The program operates with parameters of two categories: the control ones and the varied ones.

The control parameters are:

(1) The survey rate.

(2) The number of reflections measured with higher accuracy (HAM reflections).

The varied parameters are:

(1)The first measurement duration (or its scan rate).

(2)The second measurement duration (or its scan rate).

(3) The threshold accuracy of the first measurement for the exeqution of the second one.

The target values of control parameters are set by the user.

The varied parameters are corrected periodically after the measurement of the current group of reflections. The base for the correction is a comparison between the current and the target values of control parameters. The instrument overall efficiency is monitored at the same time. The reference reflections are not included in the analysis.

The first measurement duration is calculated from the survey rate. It depends on the number HAM reflections.

The second measurement limit duration is equal to the first one multiplied by an arbitrary coefficient. The coefficient maximum and minimum values are set by the user. If the HAM reflections quantity is smaller than the target one, the coefficient decreases, and both measurements durations become closer to each other. (A similar situation arises when the survey rate does not correspond to the mean intensity of reflections).

Such a mutual dependence of data collection basic parameters leads automatically to an optimum time distribution between the first and the second measurements at the set survey rate in a wide range of mean intensity. When the intensity is high and the HAM reflections quantity corresponds to the target one, then the main measurement is slow and the estimating one is fast. At reduced intensity, the HAM reflections quantity decreases and the estimating measurement duration increases. The main measurement duration also increases but at smaller degree (because of the opposite influence of both parameters determining its value).

As an illustration, the results of a model calculation for typical situations are given in Table.[1]

[1] Calculaton formulae: $t_0 = 60/R_0$; $t_1 = t_0 / (\bar{q}\gamma + 1)$; $t_2 = \gamma\, t_1$; $V_1 = \Delta\omega / t_1$; $V_2 = V_1/\gamma$ - are equivalent to formulae (1) and (2), see 3.2. Scan interval $\Delta\omega$ is 0.5 degree; background to scan times ratio t_B/t_P is 0.5; it is equivalent to duration of scan at $\Delta\omega$ = 0.75 degree without measurement of background.

Table. Scan rates V_1 and V_2 for 1-st and 2-nd measurements and coefficient γ as functions of the HAM reflections set quantity q_0 and their real quantity $\bar{q}$.

The set survey rate R_0= 120 reflections/hour

	q_0	0.95		0.20	
$\bar{q}/q_0$	γ	V_1	V_2	V_1	V_2
1.00	8	12.9	1.61	3.90	0.49
0.25	4	2.93	0.73	1.80	0.45

In contrast to the standard method the 2-nd measurement duration is always as long as possible (except for reference reflections). Thus the strong reflections acquire a high accuracy according to their contribution in structural information (Vainshtein and Kayushina, 1967).

If the second measurement scan rate is smaller than the apparatus minimum rate then the measurement is splitted into several ones.

The HAM reflections quantity is maintained by changing the 1-st measurement threshold accuracy.

These reflections are used also for the instrument overall efficiency monitoring. If no HAM reflections are in last groups then one matrix reflection peak intensity is checked. If that intensity is less than 0.5 the original one, an attempt to refine the orientation matrix is made. The survey is stopped in case of failure. (Monitoring may be disconnected).

3. Algorithm

3.1. Survey rate maintenance. Distribution of the time between the first and the second measurements of a HAM-reflection

First measurement duration, minutes $t_1 = \left(\overline{R}/R_0\right)\overline{t_1}$ (1)

or First scan rate, degree/min. $V_1 = \left(R_0/\overline{R}\right)\overline{V_1}$ (1')

Maximum second measurement duration $t_{2\max} = \gamma\, t_1$ (2)

or Minimum second scan rate $V_{2\min} = V_1/\gamma$ (2′)

$$V_{\lim} \le V_1 \le V_{\max}$$

$$V_{\lim} \le V_{2\min} \le V_{\max}$$

Coefficient γ (empiric formula) $\gamma = \left(\bar{q}/q_0\right)^{1/2} \gamma_{\max}$ (3)

$$\gamma_{\min} \le \gamma \le \gamma_{\max}$$

3.2. Quantity of HAM reflections maintenance

First measurement threshold accuracy $(I/\sigma)_1 = \left(\bar{q}/q_0\right)\overline{(I/\sigma)_1}$ (4)

$$(I/\sigma)_{\min} \le (I/\sigma)_1$$

3.3. Higher accuracy measurements

Second measurement duration $t_2 = \left(N_{\lim}/N_{1\max}\right) t_1$ (5)

$$t_2 \le t_{2\max}$$

or Second scan rate $V_2 = \left(N_{1\max}/N_{\lim}\right) V_1$ (5′)

$$V_{2\min} \le V_2$$

Repeated scans quantity $Z = V_{\lim}/V_2 + 0.5$ (integer) (6)

$$Z_{\min} \le Z \le Z_{\max}$$

3.4. Notations

$\overline{R} = \sum wR \big/ \sum w$: current mean survey rate, reflections/hour (7)

$R = n_0 \big/ \Delta t$: current survey rate of the m-th group (8)

Δt : measurement duration of the m-th group, hours

$\overline{q} = \sum wn \big/ n_0 \sum w$: current mean HAM reflections quantity (9)

$\overline{t_1} = \sum wt_1 \big/ \sum w$: current mean first measurement duration, minutes (10)

$\overline{V_1} = \sum wV_1 \big/ \sum w$: current mean first measurement scan rate, °/min (11)

$\overline{(I/\sigma)_1} = \sum w (I/\sigma)_1 \big/ \sum w$: current mean threshold accuracy (12)

The sums are taken over m groups.

R_0 : survey rate target value, reflections/hour,

q_0 : HAM reflections relative quantity target value,

n_0 : number of reflections in the group,

n : real number of the HAM reflections in m-th group,

w : individual weight of m-th group,

$N_{1\max}$: maximum number of counts registered on a scan step,

$N_{\lim}$: limit number of counts restricted by step memory.

$V_{\lim}$: apparatus minimum scan rate,

$V_{\max}, \gamma_{\min}, \gamma_{\max}, Z_{\min}, Z_{\max}, M$: arbitrary values.

4. Conclusion

The method ensures data collection with optimal quality at the set rate and allows to plan the experiment in relation of its duration and the number of reflections measured with higher accuracy. If the instrument works normally, no operator intervention in the course of the experiment is necessary.

The statistical optimization based on prediction of the nearest future from the results of recent past, will be the more effective the higher the homogeneity of the intensity. Therefore it is advisable to make the measurements in spherical layers in reciprocal space, by inverted passage of the rows.

Appendix

HAM reflections mean intensity calculation

$$\overline{I} = \overline{P} - \overline{B} \tag{13}$$

$$\overline{P} = t_0 \sum P \Big/ \sum t_P \tag{14}$$

$$\overline{B} = t_0 \sum B \Big/ \sum t_B \tag{15}$$

$$\sigma_I^2 = \sigma_P^2 + \sigma_B^2 \tag{16}$$

$$\sigma_P^2 = t_0^2 \Big/ \sum \left(t_P^2 / P \right) \tag{17}$$

$$\sigma_B^2 = t_0^2 \Big/ \sum \left(t_B^2 / B \right) \tag{18}$$

Or, in ‘scan rate’ terms:

$$\overline{P} = V_0^{-1} \sum P \Big/ \sum V^{-1} \tag{14’}$$

$$\overline{B} = V_0^{-1} \sum B \Big/ \sum \left(t_B / t_P \right) V^{-1} \tag{15’}$$

$$\sigma_P^2 = V_0^{-2} \Big/ \sum P^{-1} V^{-2}$$

(17′)

$$\sigma_B^2 = V_0^{-2} \Big/ \sum \left(t_B / t_P \right)^2 B^{-1} V^{-2}$$

(18′)

$\bar{I}$: the mean intensity,

$\bar{P}$: the mean quantity of counts registered on peak,

$\bar{B}$: the mean quantity of counts registered on background.

t_0 : the measurement standard duration defining the scale of intensity,

V_0 : the standard scan rate,

σ_I^2 : the mean intensity dispersion.

The sums are taken over all repeated measurements of one reflection.

The t_P / t_0 and t_B / t_0 values are used as weigths here. It avoids a small uncorrectness of traditional weigths σ^{-2} (Nekrasov, 1987b).

References

D.F.Grant (1973). *Acta Cryst.* **A29**, 217.

Yu.Nekrasov (1987a). *Sov. Phis. Crystallogr.* **32**(5), 744-745.

Yu.Nekrasov (1987b). *Sov. Phis. Crystallogr.* **32**(3), 459-460.

B.K.Vainshtein & R.L.Kayushina (1967). *Sov. Phis. Crystallogr.* **11**(4), 468-474.