Chi-Square Minimization Linear Fit
With Photon-Noise Correlated Errors

John Fowler
9 October, 2002

This discussion is excerpted from the SDS for the leariz module (Appendix C). The linear model defined as

$$y = a\,x + b$$

where y is the observed DN and x is the plane number; a and b will be computed by the chi-square-minimization procedure. Since there are N usable samples to fit an equation with two coefficients, the number of chi-square degrees of freedom is $N_f = N - 2$.

The chi-square minimization procedure will be described for the full-covariance case, i.e., with nonzero off-diagonal elements arising from the correlated photon noise that is summed up the ramp. The full covariance matrix is reconstructed as follows. The model for the photometric error at plane i is

$$\varepsilon_i = \varepsilon_{pi} + \varepsilon_{ui} = \sum_{j=1}^{i} \Delta\varepsilon_{pj} + \varepsilon_{ui}$$

where ε_{pi} is the total photon noise at plane i, which is the sum of the incremental photon noise added at each plane leading up to plane i, $\Delta\varepsilon_{pj}$ at plane j, and ε_{ui} is the rest of the photometric error at plane i, assumed to be uncorrelated with all other errors in the ramp (e.g., read noise). We are interested in the expectation value of the product of the errors at any two planes m and n:

$$\begin{aligned}
\varepsilon_m\,\varepsilon_n &= \left(\sum_{j=1}^{m} \Delta\varepsilon_{pj} + \varepsilon_{um} \right) \left(\sum_{k=1}^{n} \Delta\varepsilon_{pk} + \varepsilon_{un} \right) \\
&= \sum_{j=1}^{m} \left(\Delta\varepsilon_{pj} \sum_{k=1}^{n} \Delta\varepsilon_{pk} \right) + \varepsilon_{um}\,\varepsilon_{un} + \varepsilon_{um} \sum_{k=1}^{n} \Delta\varepsilon_{pk} + \varepsilon_{un} \sum_{j=1}^{m} \Delta\varepsilon_{pj}
\end{aligned}$$

Since ε_{um} is uncorrelated with all other errors in the ramp except for ε_{un} when $m = n$, the last two terms on the right will become zero when we take expectation values. Furthermore, each incremental photon-noise error $\Delta\varepsilon_{pj}$ is uncorrelated with each other $\Delta\varepsilon_{pk}$ except for $j = k$. For $m = n$, the expectation values are therefore

$$\left\langle \varepsilon_n^2 \right\rangle \equiv \sigma_n^2 = \left\langle \sum_{k=1}^{n} \Delta\varepsilon_{pk}^2 \right\rangle + \left\langle \varepsilon_{un}^2 \right\rangle = \sigma_{pn}^2 + \sigma_{un}^2 = \frac{y_n}{G} + \sigma_{un}^2$$

where $\sigma_n^{\,2}$ is the total photometric uncertainty at plane n, $\sigma_{pn}^{\,2}$ is the full photon-noise uncertainty at plane n, which is equal to the sample at that plane divided by the gain G (we will assume herein that the sample is in DN; if it is in electrons, then y is not divided by G), and σ_{un}^2 is the uncertainty at plane n due to all effects other than photon noise. For $m \neq n$, let $m > n$ (for the case $n > m$, the symmetry of the covariance matrix will be used; here we consider only $m > n$ to make it clear that the summation below extends only up to the lower of the two planes); then

$$\left\langle \varepsilon_m \varepsilon_n \right\rangle \equiv \sigma_{mn}^2 = \left\langle \sum_{k=1}^{n} \Delta\varepsilon_{pk}^2 \right\rangle = \sigma_{pn}^2 = \frac{y_n}{G}$$

The error covariance matrix for samples along the ramp can be constructed from these expressions, since the off-diagonal values $\sigma_{mn}^2 = \sigma_{pn}^2$ are known from the y_n and G values, and the diagonal values are known from the linearization error model (for this to work properly, input uncertainties for the DCE should be used, then the linearization uncertainties are root-sum-squared with those, and the result provides the diagonal elements of the full covariance matrix). The code checks to ensure that y_n/G is not greater than or equal to σ_n^2 by rescaling the former to 99% of the latter if it is not already at least that small; this is done to avoid absurdities due to corrupted data.

For illustration purposes, consider a ramp with $N = 5$; the error covariance matrix is then

$$\Omega = \begin{pmatrix} \sigma_1^2 & \frac{y_1}{G} & \frac{y_1}{G} & \frac{y_1}{G} & \frac{y_1}{G} \\ \frac{y_1}{G} & \sigma_2^2 & \frac{y_2}{G} & \frac{y_2}{G} & \frac{y_2}{G} \\ \frac{y_1}{G} & \frac{y_2}{G} & \sigma_3^2 & \frac{y_3}{G} & \frac{y_3}{G} \\ \frac{y_1}{G} & \frac{y_2}{G} & \frac{y_3}{G} & \sigma_4^2 & \frac{y_4}{G} \\ \frac{y_1}{G} & \frac{y_2}{G} & \frac{y_3}{G} & \frac{y_4}{G} & \sigma_5^2 \end{pmatrix}$$

The chi-square minimization employs the general definition of chi-square:

$$\chi^2 = u \; W \; u^T$$

where W is the inverse of Ω and u is a vector with components

$$u_i \equiv y_i - a\, x_i - b$$

Expanding the vector-matrix-vector multiplication yields

$$\chi^2 = \sum_i \sum_j w_{ij}\, u_i\, u_j$$

where the sums are from 1 to N and the w_{ij} are the elements of the W (weight) matrix. The derivatives of the equation above with respect to a and b are set to zero to obtain the system of equations

$$\begin{aligned} a \sum_i w_i\, x_i \quad &+ \quad b \sum_i w_i \quad = \quad \sum_i w_i\, y_i \\ a \sum_i x_i\, z_i \quad &+ \quad b \sum_i w_i\, x_i \quad = \quad \sum_i y_i\, z_i \end{aligned}$$

where

$$w_i \equiv \sum_j w_{ij}$$

$$z_i \equiv \sum_j w_{ij}\, x_j$$

This 2×2 system is easily solved for a and b, and then the value of chi-square is obtained from the equation defining it above.