Bayesian Estimation of Point Source Probability

David Makovoz, January 11, 2001

1. General Form

The value of each input image pixel s_i is a sum of the background and point source contributions:

$$s_i \;=\; b_i \;+\; \sum_j H_{i\text{-}j}\, x_j \quad ,$$

Equation 1

where b_i is the sky and instrument background and the measurement noise, x_i is the intensity of the point sources in the sky multiplied by the calibration. It is convolved with the overall point spread function (PSF). We assume that the PSF is independent of the position in the image and is given by the normalized function H_j. The summation is performed over the whole image, but effectively only extends so far as the PSF remains non-negligible.

The general problem is to estimate the probability of the point source being at a particular pixel i given a measurement vector **s** in a certain window W surrounding this pixel. We assume that the point source and background noise are characterized by the distribution functions for point sources $f_x(\mathbf{x})$ and $f_b(\mathbf{b})$ for the background. We consider two hypotheses for the pixel i . The first hypothesis h_1 is that there is a point source at the pixel, and the second h_2 (null hypothesis) is that there is not a point source at the pixel. Hipothesis h_2 includes the possibility of having a bad pixel (a radhit).

The probability of k^{th} hypothesis conditioned on the measurement **s** is given by Bayesian theorem:

$$P(h_k \mid \mathbf{s}) = \frac{f(\mathbf{s} \mid h_k) P(h_k)}{f(\mathbf{s})} \quad ,$$

Equation 2

where $P(h_k)$ is the a priori probability of the k^{th} hypothesis, $f(\mathbf{s})$ is the probability density of observing the set of pixel values **s**. Assuming completeness of the hypothesis' set, i.e. $P(h_2)$+ $P(h_2)$=1, we obtain for $f(\mathbf{s})$

$$f(\mathbf{s}) = \sum_{k=1,2} f(\mathbf{s} \mid h_k) P(h_k) \quad .$$

Equation 3

The probability density of measurement **s** under the null hypothesis is given simply by the background distribution function $f_b(\mathbf{s})$. The probability density of measurement **s** under the point source hypothesis is the result of integration over all possible point source inputs **x**:

$$f(\mathbf{s}\,|\,h_1) = \int d\mathbf{x}\, f(\mathbf{s}\,|\,\mathbf{x}) f_x(\mathbf{x}) = \int d\mathbf{x}\, f_b(\mathbf{s}-\mathbf{Hx}) f_x(\mathbf{x})\,.$$

Equation 4

Here the distribution function $f(\mathbf{s}|\mathbf{x})$ of measured values **s** conditioned on the point source contribution **x** is reduced to the background distribution function $f_b(\mathbf{s}\text{-}\mathbf{Hx})$ because of the additive character of the point source and background contributions (**Equation 1**). Matrix **H** is constructed from the point spread function: $\mathbf{H}_{ij} = H_{i-j}$. Combining everything we obtain the final expression for the quantity in question:

$$P(h_1\,|\,\mathbf{s}) = \frac{P(h_1)\int d\mathbf{x}\, f_b(\mathbf{s}-\mathbf{Hx}) f_x(\mathbf{x})}{P(h_1)\int d\mathbf{x}\, f_b(\mathbf{s}-\mathbf{x}) f_x(\mathbf{x}) + P(h_2) f_b(\mathbf{s})} =$$

$$\left(1 + \frac{(1-P(h_1)) f_b(\mathbf{s})}{P(h_1)\int d\mathbf{x}\, f_b(\mathbf{s}-\mathbf{Hx}) f_x(\mathbf{x})}\right)^{-1}$$

Equation 5

2. Simplifications

To evaluate **Equation 5** we need to make some assumptions about the point source and background distribution functions. The Gaussian distribution is a realistic approximation for the background distribution function.

$$f_b(\mathbf{b}) \;=\; N_b \exp(-\frac{1}{2}\delta\mathbf{b}^{\mathrm{T}}\mathbf{C}_b^{-1}\delta\mathbf{b})\ ,$$
$$N_b = (2\pi)^{-W/2}\ \ (Det(\mathbf{C}_b^{-1}))^{1/2}\ \ , \delta\mathbf{b} = \mathbf{b} - \overline{b}\,.$$

Equation 6

Here $\mathbf{C}_b$ is the background covariance matrix, *Det* denotes determinant.

The problem is that now even if we assume the Gaussian distribution for the point sources Equation 5 is computationally impossible to evaluate, even though it can be found in the closed form.

The hypothesis we entertain is the presence of a point source at a given pixel *i*. We make two assumptions. The first assumption is that this is the only point source present in the window *W*, i.e. to say that we working in the limit of very low density of point sources. The second assumption is that we can estimate the strength x_0 of the point source at the pixel *i* given the data, thus reducing the distribution function to a delta-function $f_x(x_j) = \delta_{ji}(\delta(x_i\text{-}x_0))$. Using the above approximations we obtain an expression for the probability $P(h_1|\mathbf{s})$:

$$P(h_1 \mid \mathbf{s}) = \left(1 + \frac{(1-P(h_1))\exp(-\frac{1}{2}(\mathbf{s}-\bar{b})^{\mathrm{T}}\mathbf{C}_b(\mathbf{s}-\bar{b})}{P(h_1)\exp(-\frac{1}{2}(\mathbf{s}-\bar{b}-\mathbf{H}x_0)^{\mathrm{T}}\mathbf{C}_b(\mathbf{s}-\bar{b}-\mathbf{H}x_0)} \right)^{-1}$$

Equation 7

Further simplification is achieved if the background is assumed to be uncorrelated for different pixels, i.e.

$$[C_{ij}]_b = \delta_{ij}\sigma_b^2$$

Equation 8

$$P(h_1 \mid \mathbf{s}) = \left(1 + \frac{1-P(h_1)}{P(h_1)} \exp\left(\frac{1}{2\sigma_b^2} (\sum_i (s_i - \bar{b} - H_i x_0)^2 - \sum_i (s_i - \bar{b})^2) \right) \right)^{-1}$$

Equation 9

The value of x_0 for the pixel flux density itself can be estimated by filtering the input image. In general then we will need two images input for probability estimator: the original input image before and after filtering.

The alternative is to derive a simple filter that can be applied on the fly and incorporated into **Equation 9**. A point source of flux density x_0 at pixel i has the following response p_j in the window W: $p_j = x_0 H_{i-j}$. We minimize the mean-squared-error (*MSE*) between the data s_j and the point source contribution p_j with respect to the flux density of the point source.

$$\partial MSE / \partial x_0 = \partial \left(\sum_i (s_i - \bar{b} - x_0 H_i)^2 \right) / \partial x_0 = -2\sum_i (s_i - \bar{b}) H_i + x_0 \sum_i H_i^2 = 0.$$

$$x_0 = \frac{\sum_i (s_i - \bar{b}) H_i}{\sum_i H_i^2}$$

Equation 10

After substituting this expression into **Equation 9** we get for the probability of point source presence at a given pixel:

$$P(h_1\,|\,\mathbf{s}) = \left(1 + \frac{1-P(h_1)}{P(h_1)}\exp\left(-\frac{1}{2\sigma_b^2}\frac{\left(\sum_i (s_i - \bar{b})H_i\right)^2}{\sum_i H_i^2}\right)\right)^{-1}$$

Equation 11

Another approach is to assume uncorrelated Gaussian distribution for the point sources and background:

$$f_b(b) \quad = \quad \frac{1}{\sqrt{2\pi}\sigma_b}\exp\left(-\frac{\delta b^2}{2\sigma_b^2}\right) \ ,$$

$$f_x(x) \quad = \quad \frac{1}{\sqrt{2\pi}\sigma_x}\exp\left(-\frac{\delta x^2}{2\sigma_x^2}\right)$$

Equation 12

We also assume that there is only one point source, i.e. $f_x(x_j) = f_x(x_0)\,\delta(x_i - x_0)$. Then integration in **Equation 4** can be performed.

$$f(\mathbf{s}\,|\,h_1) = \frac{1}{(\sqrt{2\pi}\sigma_b)^W\sqrt{2\pi}\sigma_x}\int dx\,\exp\left(-\frac{\sum_i (s_i - \bar{b} - H_i\delta x)^2}{2\sigma_b^2} - \frac{\delta x^2}{2\sigma_x^2}\right) =$$

$$\frac{1}{(\sqrt{2\pi}\sigma_b)^W\sqrt{2\pi}\sigma_x}\exp\left(-\frac{\sum_i (s_i - \bar{b})^2}{2\sigma_b^2} + \frac{\left(\sum_i (s_i - \bar{b})H_i\right)^2}{2\sigma_b^4/\sigma_T^2}\right)\int dy\,\exp\left(-\frac{y^2}{2\sigma_T^2}\right) =$$

$$\frac{\sigma_T}{(\sqrt{2\pi}\sigma_b)^W\,\sigma_x}\exp\left(-\frac{\sum_i (s_i - \bar{b})^2}{2\sigma_b^2} + \frac{\left(\sum_i (s_i - \bar{b})H_i\right)^2}{2\sigma_b^4/\sigma_T^2}\right)$$

Equation 13

Here

$$\frac{1}{\sigma_T^2} = \frac{1}{\sigma_x^2} + \frac{\sum_i H_i^2}{\sigma_b^2}$$

Equation 14

Plug it into **Equation 5** to obtain

$$P(h_1 \mid \mathbf{s}) = \left(1 + \frac{1-P(h_1)}{P(h_1)} \frac{\sigma_x}{\sigma_T} \exp\left(-\frac{\left(\sum_{\mathrm{i}} (\mathrm{s_i} - \bar{b}) H_i \right)^2}{2\sigma_b^4 / \sigma_T^2} \right) \right)^{-1}$$

Equation 15

Appendix

$$\frac{\sum_{\mathrm{i}} (\mathrm{s_i} - \bar{b} - H_i \delta x)^2}{2\sigma_b^2} + \frac{\delta x^2}{2\sigma_x^2} = \frac{\sum_{\mathrm{i}} (t_i - H_i y)^2}{2\sigma_b^2} + \frac{y^2}{2\sigma_x^2}$$

$t_i = s_i - \bar{b}, y = \delta x.$

$$\frac{\sum_{\mathrm{i}} (\mathrm{t_i} - H_i y)^2}{2\sigma_b^2} + \frac{y^2}{2\sigma_x^2} = \frac{t^2 - 2(t \cdot H) y + H^2 y^2}{2\sigma_b^2} + \frac{y^2}{2\sigma_x^2} = \frac{t^2}{2\sigma_b^2} + y^2 (\frac{1}{2\sigma_x^2} + \frac{H^2}{2\sigma_b^2}) - \frac{2(t \cdot H) y}{2\sigma_b^2},$$

$$t^2 = \sum_{\mathrm{i}} \mathrm{t_i}^2, H^2 = \sum_{\mathrm{i}} \mathrm{H_i}^2 . \frac{1}{\sigma_T^2} = \frac{1}{2\sigma_x^2} + \frac{H^2}{2\sigma_b^2}$$

$$\frac{y^2}{2\sigma_T^2} - \frac{2(t \cdot H) y}{2\sigma_b^2} = \frac{1}{2} (\frac{y^2}{\sigma_T^2} - \frac{2(t \cdot H) y}{\sigma_b^2} + A - A) = \frac{1}{2} (\frac{y}{\sigma_T} - \frac{\sigma_T (t \cdot H)}{\sigma_b^2})^2 - \frac{1}{2} A.$$

$$A = \frac{\sigma_T^2 (t \cdot H)^2}{\sigma_b^4}$$