UCLA

Department of Physics & Astronomy

Definition of χ^2 for Parabolic Surface

Author:
Nicolas Frega

Supervisor:
Dr. Vladimir Vassiliev

April 22, 2014

1 Intro

The goal is find the set of parameters that will minimize the discrepancy between a given set of measured data points:

$$\hat{\mathbf{r}}_n = \begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix}_n \tag{1.1}$$

and a model that generates the surface given by:

$$\mathbf{r}\left(\tau_n, \phi_n\right) = \begin{pmatrix} \sqrt{\tau_n}\cos\phi_n \\ \sqrt{\tau_n}\sin\phi_n \\ f\left(\tau_n\right) \end{pmatrix} \tag{1.2}$$

These parameters will be part of a *transformation* that applied to (1.1) satisfies:

$$transformation(data) - model = \vec{\zeta}_n \tag{1.3}$$

where $\vec{\zeta}_n$ is the noise due to uncertainties in the measurements.

2 Parameters

The set of parameters necessary are found in the usual linear transformation of any given point $\hat{\mathbf{r}}_n$:

$$\vec{B}_n = R_z\left(\theta\right) R_y\left(\beta\right) R_x(\alpha)\hat{\mathbf{r}}_n + \vec{T} \tag{2.1}$$

where the matrices R are the rotation matrices given by:

$$R_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.2}$$

$$R_y(\beta) = \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} \tag{2.3}$$

$$R_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} \tag{2.4}$$

and $\vec{T}$ is the translation vector given by:

$$\vec{T} = \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \tag{2.5}$$

thus, (1.3) can be expressed as:

$$R_z\left(\theta\right) R_y\left(\beta\right) R_x(\alpha)\hat{\mathbf{r}}_n + \vec{T} - \mathbf{r}\left(\tau_n, \phi_n\right) = \vec{\zeta}_n \tag{2.6}$$

3 Surface

Taking advantage of the symmetry of the surface, it is possible to rewrite (2.6) as:

$$R_y(\beta)\, R_x(\alpha)\hat{\mathbf{r}}_n + R_z(-\theta)\, \vec{T} - \mathbf{r}(\tau_n, \phi_n - \theta) = R_z(-\theta)\, \vec{\zeta}_n \tag{3.1}$$

Letting $R_z(-\theta)\, \vec{T} = \vec{t}$, and $\vec{S}_n = R_y(\beta)\, R_x(\alpha)\hat{\mathbf{r}}_n + \vec{t}$, (3.1) becomes:

$$\vec{S}_n - \mathbf{r}(\tau_n, \phi_n - \theta) = R_z(-\theta)\, \vec{\zeta}_n \tag{3.2}$$

It follows from (3.2) that the parameters are: $\vec{p}_\mu = (\beta, \alpha, t_x, t_y, t_z)$, in which the greek indices enumerates the given parameters: $\vec{p}_2 = \alpha$. The data set will fit the model by the usual χ^2 distribution:

$$\chi^2 = \sum_{n=1}^{N} \left[\vec{S}_n - \mathbf{r}(\tau_n, \phi_n - \theta)\right]^T \left[\vec{S}_n - \mathbf{r}(\tau_n, \phi_n - \theta)\right] \tag{3.3}$$

It is possible to rewrite (3.3) in terms of the components of the vector $\vec{S}_n$:

$$\chi^2 = \sum_{n=1}^{N} \left[\left((S_n)_x - (r(\tau_n, \phi_n - \theta))_x\right)^2 + \left((S_n)_y - (r(\tau_n, \phi_n - \theta))_y\right)^2 + \left((S_n)_z - (r(\tau_n, \phi_n - \theta))_z\right)^2\right] \tag{3.4}$$

Due to the choice of surface the first two terms of (3.4) can be set to zero. By letting $\lambda_n = (S_n)_z$ and $(r(\tau_n, \phi_n - \theta))_z = f(\tau_n)$, it is possible to rewrite (3.4) as:

$$\chi^2 = \sum_{n=1}^{N} (\lambda_n - f(\tau_n))^2 \tag{3.5}$$

For a parabolic surface we let $f(\tau_n) = \frac{1}{4}\tau_n$, and from (1.2) and (2.1) it follows that:

$$\tau_n(\vec{p}) = (S_n)_x^2 + (S_n)_y^2 \tag{3.6}$$

4 Expanding χ^2

It is now possible to expand (3.5) with respect to $\vec{p}_\mu$ in order to find the next iteration of parameters:

$$\vec{p}^{\ (i+1)} = \vec{p}^{\ (i)} + \delta\vec{p} \tag{4.1}$$

where i is the iteration number.

4.1 First Derivative of χ^2

Taking the first derivative of (3.5) with respect to $\vec{p}_\mu$:

$$\frac{\partial\chi^2}{\partial\vec{p}_\mu} = \sum_{n=1}^{N} 2\left(\lambda_n - f(\tau_n)\right)\left(\frac{\partial\lambda_n}{\partial\vec{p}_\mu} - \left.\frac{df(\tau_n)}{d\tau}\right|_{\tau=\tau_n(\vec{p})}\frac{d\tau}{d\vec{p}_\mu}\right)\Bigg|_{\vec{p}=\vec{p}^{\ (i)}} \tag{4.2}$$

If we let the right hand side of (4.2):

$$V_n\left(\vec{p}_\mu\right) = 2\left(\lambda_n - f(\tau_n)\right)\left(\frac{\partial\lambda_n}{\partial\vec{p}_\mu} - \left.\frac{df(\tau_n)}{d\tau}\right|_{\tau=\tau_n(\vec{p})}\frac{d\tau}{d\vec{p}_\mu}\right)\Bigg|_{\vec{p}=\vec{p}^{\ (i)}} \tag{4.3}$$

It then follows that:

$$V\left(\vec{p}_\mu\right) = \sum_{n=1}^{N} V_n\left(\vec{p}_\mu\right) \tag{4.4}$$

From which (4.2) can be expressed as:

$$\frac{\partial\chi^2}{\partial\vec{p}_\mu} = V\left(\vec{p}_\mu\right) \tag{4.5}$$

4.2 Second Derivative of χ^2

The second derivative of (3.5) with respect to each parameter is given by:

$$\begin{aligned}\frac{\partial^2\chi^2}{\partial\vec{p}_\mu\partial\vec{p}_\nu} = \sum_{n=1}^{N} 2\Bigg[&\left(\left(\frac{\partial\lambda_n}{\partial\vec{p}_\mu} - \left.\frac{df(\tau_n)}{d\tau}\right|_{\tau=\tau_n(\vec{p})}\frac{d\tau}{d\vec{p}_\mu}\right)\left(\frac{\partial\lambda_n}{\partial\vec{p}_\nu} - \left.\frac{df(\tau_n)}{d\tau}\right|_{\tau=\tau_n(\vec{p})}\frac{d\tau}{d\vec{p}_\nu}\right)\right)\Bigg|_{\vec{p}=\vec{p}^{\ (i)}} + \\ &\left(\lambda_n - f(\tau_n)\right)\left(\frac{\partial^2\lambda_n}{\partial\vec{p}_\mu\partial\vec{p}_\nu} - \left.\frac{df(\tau_n)}{d\tau}\right|_{\tau=\tau_n(\vec{p})}\frac{\partial^2\tau}{\partial\vec{p}_\mu\partial\vec{p}_\nu} - \left.\frac{d^2f(\tau_n)}{d\tau^2}\right|_{\tau=\tau_n(\vec{p})}\frac{d\tau}{d\vec{p}_\mu}\frac{d\tau}{d\vec{p}_\nu}\right)\Bigg]\end{aligned} \tag{4.6}$$

Setting the right hand side of (4.6) equal to $M_n\left(\vec{p}_{\mu\nu}\right)$:

$$M\left(\vec{p}_{\mu\nu}\right) = \sum_{n=1}^{N} M_n\left(\vec{p}_{\mu\nu}\right) \tag{4.7}$$

Consequently (4.6) can be expressed as:

$$\frac{\partial^2\chi^2}{\partial\vec{p}_\mu\partial\vec{p}_\nu} = M\left(\vec{p}_{\mu\nu}\right) \tag{4.8}$$

5 Minimizing χ^2

Using (4.1) to write the expansion of χ^2:

$$\chi^2\left(\vec{p}^{\,(i+1)}\right) = \chi^2\left(\vec{p}^{\,(i)}\right) + V\left(\vec{p}^{\,(i)}\right)_\mu \delta\vec{p}_\mu + \frac{1}{2!}\left(M\left(\vec{p}^{\,(i)}\right)\right)_{\mu\nu}\delta\vec{p}_\mu\delta\vec{p}_\nu \tag{5.1}$$

Taking the derivative with respect to $\delta\vec{p}_\eta$:

$$\frac{\partial\chi^2}{\partial\delta\vec{p}_\eta} = V_\eta^{(i)} + \frac{1}{2!}\left(M_{\mu\nu}^{(i)}\delta\vec{p}_\mu + \delta\vec{p}_\nu M_{\mu\nu}^{(i)}\right) \tag{5.2}$$

Setting $\frac{\partial\chi^2}{\partial\delta\vec{p}_\eta} = 0$ and simplifying:

$$0 = V_\eta^{(i)} + M_{\eta\mu}^{(i)}\delta\vec{p}_\mu \tag{5.3}$$

Solving the above equation for $\delta\vec{p}_\mu$:

$$\delta\vec{p}_\mu = -\left(M_{\eta\mu}^{(i)}\right)^{-1} V_\eta^{(i)} \tag{5.4}$$

6 Set Point Option

If it is assumed that the error in (3.2) is negligible, it follows that:

$$\mathbf{r}\left(\tau_n, \phi_n - \theta\right) = R_y\left(\beta\right) R_x(\alpha)\hat{\mathbf{r}}_n + \vec{t} \tag{6.1}$$

At any given point the translation necessary to satisfy (6.1) is given by:

$$\vec{t} = -R_y\left(\beta\right) R_x(\alpha)\hat{\mathbf{r}}_0 + \mathbf{r}\left(\tau_0, \phi_0 - \theta\right) \tag{6.2}$$

replacing in (6.1):

$$\mathbf{r}\left(\tau_n, \phi_n - \theta\right) = R_y\left(\beta\right) R_x\left(\alpha\right)\left(\hat{\mathbf{r}}_n - \hat{\mathbf{r}}_0\right) + \mathbf{r}\left(\tau_0, \phi_0 - \theta\right) \tag{6.3}$$

Equation (6.3) shows that the problem can be reduced to two dimensions by approximating a selected point in the model frame of reference to its corresponding point in the transformed frame of reference. Then χ^2 can be minimized by simply shifting the data set by the selected point $\hat{\mathbf{r}}_0$ and iterating over i. Finally, the translations can be calculated, after convergence, by equation (6.2).