# Fitting a plane with arbitrary observational uncertainties

Daniel Foreman-Mackey[1,2]

In this note, I'll describe a method for inferring a planar relationship between three variables given a set of noisy (and possibly non-Gaussian) observations of a set of systems. For specificity, consider the problem of gyrochronology: there is postulated to be a linear relationship between the (true) age $A_n$ of a star $n$ and its (true) effective temperature $T_n$ and rotational period $P_n$

$$\log A_n \quad = \quad m_0 + m_1\,\log T_n + m_2\,\log P_n \quad . \tag{1}$$

We would like to determine the best-fit value (or sample the posterior probability) of the parameter vector $m = \{m_0, m_1, m_2\}$ conditioned on a set of noisy observations $\hat{A}_n$, $\hat{T}_n$, and $\hat{P}_n$.

For this purpose, we need to compute the marginalized likelihood

$$p(\{\hat{A}_n, \hat{T}_n, \hat{P}_n\}\,|\,m) \quad = \quad \prod_{n=1}^{N} \int p(\hat{A}_n, \hat{T}_n, \hat{P}_n, A_n, T_n, P_n\,|\,m)\,\mathrm{d}A_n\,\mathrm{d}T_n\,\mathrm{d}P_n \quad . \tag{2}$$

This joint probability function is shown in figure 1 and this factorization can be written as

$$p(\hat{A}_n, \hat{T}_n, \hat{P}_n, A_n, T_n, P_n\,|\,m) \quad = \quad p(T_n)\,p(P_n)\,p(A_n\,|\,T_n, P_n, m)\,p(\hat{A}_n\,|\,A_n)\,p(\hat{T}_n\,|\,T_n)\,p(\hat{P}_n\,|\,P_n)$$

where (by equation 1)

$$p(A_n\,|\,T_n, P_n, m) \quad = \quad \delta\left[\log A_n - (m_0 + m_1\,\log T_n + m_2\,\log P_n)\right] \quad . \tag{3}$$

Now, let's say that we have $J_n$ posterior samples

$$\begin{aligned} T_n^{(j)} &\sim p(T_n\,|\,\hat{T}_n) \\ P_n^{(j)} &\sim p(P_n\,|\,\hat{P}_n) \end{aligned} \tag{4}$$

and that we can evaluate $p(\hat{A}_n\,|\,A_n)$ up to a normalization constant. If you don't have the posterior samples, generate them from the uncertainties provided in the catalog and if you

[1]To whom correspondence should be addressed: danfm@nyu.edu

[2]Center for Cosmology and Particle Physics, Department of Physics, New York University, 4 Washington Place, New York, NY, 10003, USA

don't have functional form for the age likelihood, make one up using a KDE or similar. Now, we can evaluate the marginalized likelihood for a single star very efficiently as follows

$$
\begin{aligned}
p(\hat{A}_n, \hat{T}_n, \hat{P}_n \,|\, m) &= \int p(T_n)\, p(P_n)\, p(A_n \,|\, T_n, P_n, m)\, p(\hat{A}_n \,|\, A_n)\, p(\hat{T}_n \,|\, T_n)\, p(\hat{P}_n \,|\, P_n)\, \mathrm{d}A_n\, \mathrm{d}T_n\, \mathrm{d}P_n \\
&\propto \int p(A_n \,|\, T_n, P_n, m)\, p(\hat{A}_n \,|\, A_n)\, p(T_n \,|\, \hat{T}_n)\, p(P_n \,|\, \hat{P}_n)\, \mathrm{d}A_n\, \mathrm{d}T_n\, \mathrm{d}P_n \\
&\approx \frac{1}{J_n} \sum_{j=1}^{J_n} p(\hat{A}_n \,|\, A_n^{(j)}) \qquad (5)
\end{aligned}
$$

where $A_n^{(j)}$ is computed from the posterior samples using equation 1.

Finally, the full marginalized log-likelihood is

$$
\log p(\{\hat{A}_n, \hat{T}_n, \hat{P}_n\} \,|\, m) \approx \log \mathcal{Z} + \sum_{n=1}^{N} \log \left[ \sum_{j=1}^{J_n} p(\hat{A}_n \,|\, A_n^{(j)}) \right] \qquad (6)
$$

where $\mathcal{Z}$ is an irrelevant normalization constant. You can then toss this into an optimizer and find the maximum likelihood plane or assert a prior $p(m)$ and draw posterior samples.

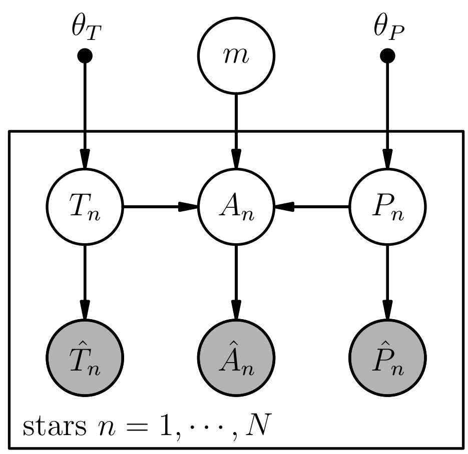


Fig. 1.—