Notes on modelling gyrochronology model for Ruth's Kepler asteroseismic sample

Suzanne Aigrain

February 4, 2014

1 Preliminaries

1.1 The data

The dataset consists of N stars, labelled by the index $n = 1, \ldots, N$. For each star, we have the following main observables:

- the effective temperature T, derived from the star's spectrum, in degrees Kelvin;
- the rotation period P, derived from the star's light curve, in days;
- the age A, derived from asteroseismology, in Gyr.
- the surface gravity $\log g$ (not sure if it's derived from the star's spectrum, asteroseismology, or both), in $\mathrm{g\,cm}^3$. For convenience of notation, we define $G = \log g$;

All the above have measurement uncertainties which are assumed to be known, Gaussian and independent, i.e.

$$p(\hat{X}_n|X_n) = \mathcal{N}(\hat{X}_n|X_n, \sigma^2_{X,n}) \tag{1}$$

where X represents any of the above variables, X_n is the true value of this variable in observation n, $\hat{X}_n$ is the noisy observed value and $\sigma_{X,n}$ the associated measurement uncertainty.

1.2 The astrophysics

The range of the key observables covered by our sample is as follows: $5500 < T < 6500$, $1 < P < 50$, $1 < A < 10$, and $3.5 < G < 4.5$. Main sequence stars cooler (or bluer) than the so-called Kraft break (which is somewhere in the range $6000 < T_K < 6500$), are assumed to follow a unique gyrochronology relation relating T (or mass, or $B-V$ or $J-K$ colour), P (or $v \sin i$) and A. The goal of this study is to use the dataset to improve the calibration of this gyrochronology relation. A commonly adopted form for the gyrochronology relation in the recent literature (see e.g. **??**) is:

$$P = a\,[(B-V) - c]^b \times A^d, \tag{2}$$

with typical values for the parameters $a \sim 0.4$, $b \sim 0.3$–0.6, $c \sim 0.4$–0.5 and $d \sim 0.5$–0.6. In this expression, c corresponds to the Kraft break, and the gyronochronology relation is not defined for stars bluer than this. Since we are working with T rather than colour, we might adopt a similar relation:

$$P = a(T - T_K)^b \times t^d. \tag{3}$$

Our sample also includes main sequence stars hotter than the Kraft break, which show very little rotational evolution: they tend to have relatively short rotation periods (1–10 d) but aside from that their periods display little dependence on either effective temperature or age. Although the Kraft break is related to the size of the outer convection zone, the effective temperature (or colour) at which it occurs is not well established, and it is even possible that stars close to the Kraft break could display either behaviour (rotational evolution following the gyrochronology relation, or no evolution), seemingly at random. Finally, our sample also contains stars which have recently left the main sequence, i.e. sub-giants, which on the contrary experience very rapid spin-down. Contrarily to the Kraft break, we can assume that we know (from stellar evolution models) the location of the main sequence turnoff as a function of effective temperature and surface gravity.

1.3 The statistical approach

We want model the data in such a way as to

- enable us to compute a predictive distibution for the age of a star given estimates of its effective temperature and rotation period.
- take into account the uncertainties on all three observables in our sample;
- allow for the fact that there are three distinct populations in our sample, and that it may not be possible to uniquely assign a given star to a specific population;
- readily enable us to incorporate other datasets (e.g. from clusters, including cooler stars) in the future.

The first point above suggests that we should treat A as the dependent variable and T and P as the independent variables. This implies re-arranging Eq. 3:

$$\log A = \alpha + \beta \log (T - T_K) + \delta \log P, \tag{4}$$

where $\alpha = -\log(a)/d, \beta = \log(b)/d$, and $\delta = -1/d$.

The second point can be addressed by using a sampling approach to marginalise over the uncertainties on the independent variables, as described in Section **??**. The third point suggests either a composite model function, or a mixture model approach, as described in Section **??**. The fourth point is trivial in a hierarchical Bayesian framework, but it does mean that we want our model to be defined over a range of temperatures wider than the one in our sample.

2 Constructing the model

We start by writing down by writing the likelihood marginalised over the observational errors:

$$p(\{\hat{P}_n, \hat{A}_n, \hat{T}_n, \hat{G}_n\}|\theta, T_K) = \prod_{n=1}^{N} \int p(\hat{A}_n, \hat{T}_n, \hat{P}_n, \log g_n, A_n, T_n, P_n, G_n|\theta, T_K) \mathrm{d}A_n \mathrm{d}T_n \mathrm{d}P_n \mathrm{d}G_n \tag{5}$$

where $\theta = \{\alpha, \beta\, \delta\}$ represents the parameters of the gyrochronology relation. The joint probability distribution is then given by:

$$p(\hat{A}_n, \hat{T}_n, \hat{P}_n, \log g_n, A_n, T_n, P_n, G_n|\theta, T_K) = p(A_n, T_n, P_n, G_n|\theta, T_K) p(\hat{A}_n|A_n) p(\hat{T}_n|T_n) p(\hat{P}_n|P_n) p(\hat{G}_n|G_n), \tag{6}$$

If we chose to treat the transition between the different populations of stars in our sample as a step function, we can write the marginalised likelihood for a single star as the sum of three integrals corresponding to the three different regimes:

$$p(\hat{P}_n, \hat{A}_n, \hat{T}_n, \hat{G}_n|\theta, T_K) = \sum_{k=1}^{3} \int_k p_k(A_n, T_n, P_n, G_n|\theta, T_K) p(\hat{A}_n|A_n) p(\hat{T}_n|T_n) p(\hat{P}_n|P_n) p(\hat{G}_n|G_n), \mathrm{d}A_n \mathrm{d}T_n \mathrm{d}P_n \mathrm{d}G_n \tag{7}$$

where $k = 1$ corresponds to main sequence stars cooler than the Kraft break, $k = 2$ to main sequence stars hotter than the Kraft break, and $k = 3$ to stars which have left the main sequence, and $\int_k f(x)\mathrm{d}x$ is the integral of $f(x)$ with respect to x, evaluated over regime k. Note that each star may still have a finite probability of belonging to any one of the individual regimes, because of the uncertainties on the observables.

The joint probability distribution for the true observables takes a different from in each of the three regimes:

$$\begin{aligned} p_1(A_n, T_n, P_n, G_n|\theta, T_K) &= p_1(T_n|T_K) p_1(G_n) p_1(P_n) p_1(A_n|T_n, P_n, \theta, T_K), && (8) \\ p_2(A_n, T_n, P_n, G_n|\phi_2, T_K) &= p_2(T_n|T_K) p_2(G_n) p_2(P_n|phi_2) p_2(A_n), \text{ and} && (9) \\ p_3(A_n, T_n, P_n, G_n|\phi_3) &= p_3(T_n) p_3(G_n) p_3(P_n|\phi_3) p_3(A_n), && (10) \end{aligned}$$

where:

$$p_1(A_n|T_n, P_n, \theta, T_K) = \delta \left\{ \log A_n - \left[\alpha + \beta \log (T - T_K) + \delta \log P_n \right] \right\}. \tag{11}$$

These equations merit a few comments:

- The form of Eqs. (9) and (10) implies that there is no correlation between P and T or A in regimes 2 and 3. This is reasonable for regime 2, where we do not expect any rotational evolution, as discussed in Section 1.2. It is less appropriate, at first sight, for regime 3, where on the contrary we expect very rapid rotational evolution, but this regime is not the focus of the present study, we are less interested in modelling it in detail.
- The priors are defined separately over each regime, for two reasons. First, the priors over T and G need to reflect the different ranges for these variables in the three regimes. Additionally, we can choose different period priors in the different regimes. For example we might adopt a log-normal prior with a different mean and/or variance in each regime, or a log-uniform prior defined over a different range, to reflect our expectation that stars tend to have longer or shorter rotation periods in a given regime. In Eqs. (9) and (10), we have specified a prior with hyper-parameters ϕ, which could be inferred or marginalised over at the same time as θ and T_K.
- The use of mutually independent priors for T, G and A in Eqs. (8–10) also implies that there is no correlation between these quantities. Of course, this isn't true: there are built-in correlations because of stellar evolution. In principle, it would be more adequate to construct a joint, age dependent prior on T and G, using theoretical evolutionary models, assuming an initial mass distribution and a star formation history (i.e. a prior on A). However, that would make the results dependent on the very models we are trying to test, and in any case the quality of our date probably doesn't warrant such a level of sophistication. So for now I would say use a uniform prior on T and G and a log-uniform prior on A.

The three terms in Eq. (7) become:

$$
\begin{aligned}
p_1(\hat{P}_n,\hat{A}_n,\hat{T}_n,\hat{G}_n|\theta,T_K) &= \int_1 p_1(T_n|T_K)p_1(G_n)p_1(P_n)p_1(A_n|T_n,P_n,\theta,T_K)p(\hat{A}_n|A)p(\hat{T}_n|T_n)p(\hat{P}_n|P_n)p(\hat{G}_n|G_n)\mathrm{d}A_n\mathrm{d}T_n\mathrm{d}P_n\mathrm{d}G_n \\
p_2(\hat{P}_n,\hat{A}_n,\hat{T}_n,\hat{G}_n|\phi,T_K) &= \int_2 p_2(T_n|T_K)p_2(G_n)p_2(P_n|\phi)p_2(A_n)p(\hat{A}_n|A)p(\hat{T}_n|T_n)p(\hat{P}_n|P_n)p(\hat{G}_n|G_n)\mathrm{d}A_n\mathrm{d}T_n\mathrm{d}P_n\mathrm{d}G_n \\
p_3(\hat{P}_n,\hat{A}_n,\hat{T}_n,\hat{G}_n) &= \int_3 p_3(T_n)p_3(G_n)p_3(P_n)p_3(A_n)p(\hat{A}_n|A)p(\hat{T}_n|T_n)p(\hat{P}_n|P_n)p(\hat{G}_n|G_n)\mathrm{d}A_n\mathrm{d}T_n\mathrm{d}P_n\mathrm{d}G_n
\end{aligned}
$$

In the single regime case, if using fixed priors, these act as constant normalisation constants, and can be ignored. However, if we have multiple regimes with different priors, and / or if the priors depend on the (hyper-)parameter we wish to infer or marginalise over, as is the case here, we must keep track of them in the calculation.

To evaluate the marginalised likelihood, we are going to use importance sampling. Say we have J_n samples:

$$
\begin{aligned}
T_n^{(j)} &\sim p(T_n|\hat{T}_n) \\
P_n^{(j)} &\sim p(P_n|\hat{P}_n) \\
A_n^{(j)} &\sim p(T_n|\hat{A}_n) \\
G_n^{(j)} &\sim p(G_n|\hat{G}_n)
\end{aligned}
\tag{15}
$$

Eq. (12) can be approximated as

$$
p_1(\hat{P}_n,\hat{A}_n,\hat{T}_n,\hat{G}_n|\theta,T_K) \approx \frac{1}{J_n}\sum_{j=1}^{N} f_1(T_n^{(j)},G_n^{(j)}|T_K)p_1(T_n^{(j)}|T_K)p_1(G_n^{(j)})p_1(P_n^{(j)})p(\hat{A}_n|\tilde{A}_n^{(j)}) \tag{16}
$$

where $f_1(T_n^{(j)},G_n^{(j)}|T_K)=1$ if the sample falls in regime 1 and 0 otherwise, and $\tilde{A}$ is the value of A predicted by the gyrochronology relation given $T^{(j)}$ and $P^{(j)}$:

$$
\log\tilde{A}_n^{(j)} = \alpha + \beta\log\left(T_n^{(j)} - T_K\right) + \delta\log P_n^{(j)}. \tag{17}
$$

Similarly, Eq. (13) can be approximated as

$$
p_2(\hat{P}_n,\hat{A}_n,\hat{T}_n,\hat{G}_n|\phi,T_K) \approx \frac{1}{J_n}\sum_{j=1}^{N} f_2(T_n^{(j)},G_n^{(j)}|T_K)p_2(T_n^{(j)}|T_K)p_2(G_n^{(j)})p_2(P_n^{(j)}|\phi)p_2(A_n^{(j)}). \tag{18}
$$

where $f_2(T_n^{(j)}, G_n^{(j)}|T_K) = 1$ if the sample falls in regime 2 and 0 otherwise. Finally, Eq. (14) can be approximated as

$$p_3(\hat{P}_n, \hat{A}_n, \hat{T}_n, \hat{G}_n) \approx \frac{1}{J_n} \sum_{j=1}^{N} f_3(T_n^{(j)}, G_n^{(j)}) p_3(T_n^{(j)}) p_3(G_n^{(j)}) p_3(P_n^{(j)}) p_3(A_n^{(j)}). \tag{19}$$

where $f_3(T_n^{(j)}, G_n^{(j)}) = 1$ if the sample falls in regime 3 and 0 otherwise.

The full marginalised likelihood is then

$$\log p(\{\hat{A}_n, \hat{T}_n, \hat{P}_n, \hat{G}_n\}|\theta, \phi, m) \approx sum_{i=1}^{n} \log \left(\frac{1}{J_n} \sum_{j=1}^{N} \mathcal{P}^{(j)} \right) \tag{20}$$

where

$$\begin{aligned} \mathcal{P}^{(j)} \quad = \quad & f_1(T_n^{(j)}, G_n^{(j)}|T_K) p_1(T_n^{(j)}|T_K) p_1(G_n^{(j)}) p_1(P_n^{(j)}) p(\hat{A}_n|\tilde{A}_n^{(j)}) + \\ & f_2(T_n^{(j)}, G_n^{(j)}|T_K) p_2(T_n^{(j)}|T_K) p_2(G_n^{(j)}) p_2(P_n^{(j)}|\phi) p_2(A_n^{(j)}) + \\ & f_3(T_n^{(j)}, G_n^{(j)}) p_3(T_n^{(j)}) p_3(G_n^{(j)}) p_3(P_n^{(j)}) p_3(A_n^{(j)}) \end{aligned} \tag{21}$$