PTA analysis

Ignas Anikevičius

July 16, 2013

Contents

1 Introduction

2 Background Theory

In this chapter I will go through the context of this research as well as derivations and definitions of some formulae and mathematical notation used through this report.

2.1 Pulsar Timing and Gravitational Wave Search

Pulsars are highly predictable cosmic clocks, some of which rival even atomic clocks <citation needed>. At the moment there is an ongoing effort to establish collaborations between many radio telescopes in order to make observations of pulsars across the whole sky <citation needed>. The so called International Pulsar Timing Array (IPTA) is precisely that.

The idea, that pulsars can be used for investigation of gravitational waves by investigating correlated fluctuations of time of arrival (TOA) measurements is not new <citation needed>. However, only recently data analysis has been started as sufficient amount of data has been gathered <citation needed>. The main interest is in seeing whether the residuals of the timing model carry any information of gravitational wave source and how to extract it quickly. The residual is defined as:

$$\delta\mathbf{t} = t_{TOA} - t_{expectedTOA} \tag{2.1}$$

The procedure is usually to observe the pulsars for a long time and choose only the most precise, so called millisecond pulsars, which have a very low and steady period and fit the TOAs to obtain the residuals. Then the residuals need to be marginalised over the timing model parameters, so that the actual data is independent of the fitting model. However, the scope of this project is to develop a new way of processing the marginalised residuals and hence we shall not concern ourselves with the marginalisation of the fitting parameters at this stage and we will assume, that the residuals are known to be exact within some error.

2.2 The theoretical model for the signal

2.3 Search of single sources

When the data is being processed, usually Markov Chain Monte Carlo (MCMC) methods are chosen to maximise the Bayesian likelihood function. The probability of noise defined as a multivariate Gaussian <citation needed>:

$$p\left(\mathbf{n}\right) = \frac{1}{\sqrt{\det 2\pi \mathrm{C}}} \exp\left(-\frac{1}{2}\mathbf{n}^T \mathrm{C}^{-1}\mathbf{n}\right) \tag{2.2}$$

If we assume that our residuals consist of a signal and a noise, we can rearrange the expression for noise:

$$\mathbf{n} = \delta\mathbf{t} - \mathbf{h} \tag{2.3}$$

where $\mathbf{h}$ is the signal template which is being fitted.

The matrix C in the inner product in the exponent is called covariance matrix, which contains the information about the correlations between the noise series $\mathbf{n}_i$ for each ith pulsar. It is symmetric and the form of it can be rationalized as follows:

$$C_{ij} = C_{ij}^{GWB} + C_{ij}^{WN} + C_{ij}^{RN} + C_{ij}^{PLN} = \langle \mathbf{n}_i \mathbf{n}_j \rangle \tag{2.4}$$

where the angular brackets denote the cross-correlation of the noise time series from each pulsar <citation needed>. As you can see from the same expression, the matrix has various contributions from stochastic processes — Gravitational Wave Background (GWB), the white (WN), red (RN) and power-law noises (PLN) <citation needed>.

The matrix present in the inner product in the exponent makes them particularly expensive to evaluate as the operation scales as $\mathcal{O}\left(n^3\right)$, where n is the dimensionality of the vectors. Also, because of the fact, that the matrix needs to be inverted before carrying out evaluations doesn't help either.

From now on, we shall use a shorthand notation for such inner products, by employing the Dirac's bra-ket notation:

$$\langle \mathbf{a} | \mathbf{b} \rangle = \mathbf{a}^T \mathrm{C}^{-1} \mathbf{b} \tag{2.5}$$

Given this notation, we can express the logarithm of posterior probability function ($\ln p$) as:

$$\ln p = -\frac{1}{2} \langle \delta\mathbf{t} | \delta\mathbf{t} \rangle + \langle \delta\mathbf{t} | \mathbf{h} \rangle - \frac{1}{2} \langle \mathbf{h} | \mathbf{h} \rangle \tag{2.6}$$

During the MCMC algorithm run, we would try to maximize $\ln p$ by varying the signal template and since the dimensions of the vectors used in this calculation tend to be large, these are computationally very intensive. It is only the first term, which can be precomputed before doing the calculations, however, it does not improve the computation time drastically.

2.4 Reduced Order Modelling

Reduced order modelling is not new and such methods rely on a single assumption, that the dimensionality of the problem can be reduced by trying to assess the information content. Quite recently there was a paper published by <canizares citation needed> which investigates such models in the context of LIGO <citation needed> detector. It relies on three crucial steps before the MCMC method application in the maximization of the likelihood.

3 The Reduced Basis Construction

It is important to note, that during this stage we can precompute the norm of the signal (i.e. $\langle \mathbf{h}|\mathbf{h}\rangle$), which is very useful for constructing the reduced basis *and* posterior evaluation. This means, that during the MCMC cycles we need to compute only one product, which we shall speed up by using the Reduced Order Quadrature (ROQ) rule, which is described in the next section.

4 Reduced Order Quadrature (ROQ) rule construction

This is a completely general method, which is described in the <Canizares citation needed>, but I will include the discussion of the posterior calculation here as well.

We know, that our signal template can be expressed as a linear combination of the precomputed Reduced Basis functions:

$$\mathbf{h}\left(x\right) \approx \sum_i a_i \mathbf{e}_i\left(x\right) \tag{4.1}$$

We also make it exactly match at the empirical interpolation points:

$$\mathbf{h} = \sum_i a_i \mathbf{e}_i\left(F_k\right) \tag{4.2}$$

Hence, we can reexpress the above equation as a system of equations in matrix notation (using summation convention):

$$\mathbf{h}_i = A_{ij} a_j \implies a_i = A^{-1}_{ij} h_j \tag{4.3}$$

By expanding our residuals in a similar fashion, we can easily compute the inner product between the residuals and the signal template (using the summation convention):

$$\delta\mathbf{t} \approx b_i \mathbf{e}_i \implies b_i = \langle \delta\mathbf{t} | \mathbf{e}_i \rangle \tag{4.4}$$

$$\langle \delta\mathbf{t} | \mathbf{h} \rangle \approx b_i \langle \mathbf{e}_i | \mathbf{e}_j \rangle a_i = b_i G_{ij} a_i \tag{4.5}$$

$$= b_i G_{ij} A^{-1}_{jk} h_k \equiv \widetilde{\delta t}_i h_i \tag{4.6}$$

Since the $\widetilde{\delta\mathbf{t}}$ can be precomputed before using the MCMC method, this greatly speeds up the posterior evaluations during the MCMC cycles.

Given all the derivations we have done, after constructing the ROQ rule, our evaluation of the posterior becomes:

$$\ln p = -\frac{1}{2} \langle \delta\mathbf{t} | \delta\mathbf{t} \rangle + \widetilde{\delta\mathbf{t}}^T \mathbf{h} - \frac{1}{2} \langle \mathbf{h} | \mathbf{h} \rangle \tag{4.7}$$

of which only the middle term needs to be recalculated and the dot product for 2 vectors scales as $\mathcal{O}\left(n^2\right)$. Thus, by constructing the basis, we can bring the computational cost a lot.