Universidad Nacional
Autónoma de México
UNAM

Note for the fitting lecture

Sebastien Fromenteau (ICF-UNAM) and Mariana Vargas Magaña (IF-UNAM)

June 27, 2021

1 χ^2 estimator and distribution

1.1 χ^2 estimator

The χ^2 estimator correspond to the sum of the distances between a model ($y_i^{mod}(\vec{\theta})$) and data (y_i^D) normalized to the variance of the data (σ_i^2). Here the model depend on parmeters $\vec{\theta}$. If the data points are independent, the estimator is simply:

$$\chi^2(\vec{\theta}) = \sum_{i=1}^{N} \left(\frac{y_i^{mod}(\vec{\theta}) - y_i^D}{\sigma_i} \right)^2 \tag{1}$$

Where N is the number of data points. The goal is find the parameters $\vec{\theta}^b$ which minimize this estimator. However, the best set of parameter can be bad especially if the model is not well adapt to fit the data. So we want to estimate the goodness of the fit which correspond to estimate the probability to produce the data-points assuming a gaussian distribution for the errors. This is very important, all the calculations we will derive below assume the data probability distribution is a gaussian centered in y_i^D with a standard deviation σ_i.

In the case where the data points are not independent, we have to take into account the covariance between all of them. So we have to measure the covariance matrix where each element is defined as:

$$Cov_{ij} = \left\langle (y_i^D - E(y_i^D)).(y_j^D - E(y_j^D)) \right\rangle \tag{2}$$

If the data points i and j are independent, then the corresponding covariance matrix element is null $Cov_{ij} = 0$. If all the data points are independent, then all the off-diagonal terms are null. In this case we just have the diagonal terms wich correspond to the variance of each data points :

$$Cov_{ii} = \left\langle (y_i^D - E(y_i^D)).(y_i^D - E(y_i^D)) \right\rangle = \sigma_i^2 \tag{3}$$

In any case, the general way to write the χ^2 estimator is using the matrix formalism:

$$\chi^2(\vec{\theta}|\vec{y}^D) = \left(\vec{y}^{mod}(\vec{\theta}) - \vec{y}^D \right)^t Cov^{-1} \left(\vec{y}^{mod}(\vec{\theta}) - \vec{y}^D \right), \tag{4}$$

where $\vec{y}^D$ is the vector of data, $\vec{y}^{mod}(\vec{\theta})$ is the vector of points from the model for the parameters $\vec{\theta}$ and Cov^{-1} is invert covariance matrix.

In the case of independent data points, the covariance matrix is the diagonal one with $Cov_{ii} = \sigma_i^2$ and so the invert matrix is simply $Cov_{ii}^{-1} = 1/\sigma_i^2$. You can do the calculation and retrieve the equation (1).

We have defined the estimator, now we want to define the goodness of the fit considering the value obtain with the estimator. We can calculate the probability to obtain a given value.

1.1.1 Number of Degree of freedom N_{dof}

One important thing to define is the number of degree of freedom (N_{dof}). While the χ^2 estimator seems to depend on the number of data points, it's not as direct as it appears. Basically, if you have 1data point and 1 parameter to fit, you will directly found the value where you model pass by the data point, so with a null value. If you have 2 data points and 1 parameter, that's different. If you fix your parameter to pass by the first data point, then you have another point which is not perfectly dtermined, that's a degree of freedom.

We can think in the same way with 2 points and 2 parameters. With 2 parameters you can define a linear function $y = ax + b$ and so you can find the value for a and b for which the model pass perfectly by the 2 points. So there is no degree of freedom.

The general definition for the number of degree of freedom is the number of data points minus the number of parameters of the model. We want to calculate the probability to obtain the sum of the distances for a given number of freedom $N_{dof} = N_{points} - N_{params}$.

1.2 χ^2 distribution

We will calculate the χ^2 probability distribution function (pdf) by induction over the number of dof. We will start to estimate the explicitly the pdf for $N_{dof} = 1$

1.2.1 $N_{dof} = 1$

For the case of $N_{dof} = 1$, we have $Y = x^2$ with $x \sim \mathcal{N}(0,1)$. We are interested in the probability distribution of the estimator Y.

$$P_1(Y \leq s) = \int_{-\sqrt{s}}^{\sqrt{s}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \tag{5}$$

$$P_1(Y \leq s) = \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx - \int_{-\infty}^{-\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx \right) \tag{6}$$

$$\int_{-\infty}^{-\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx = 1 - \int_{-\infty}^{\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx \tag{7}$$

$$P_1(Y \leq s) = \frac{1}{\sqrt{2\pi}} \left(2 \times \int_{-\infty}^{\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx - 1 \right) \tag{8}$$

$$\int_{-\infty}^{\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx = \frac{1}{2} + \int_{0}^{\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx \tag{9}$$

$$P_1(Y \leq s) = \frac{1}{\sqrt{2\pi}} \left(2 \times \frac{1}{2} + 2 \times \int_{0}^{\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx - 1 \right) \tag{10}$$

$$P_1(Y \leq s) = \sqrt{\frac{2}{\pi}} \int_{0}^{\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx \tag{11}$$

We want to calculate the probability the probability distribution function (pdf), so $p_1(Y = s)$ which is the derivative of the cumulative function $P_1(Y \leq s)$. So we have

$$p_1(Y = s) = \frac{dP_1(Y \leq s)}{ds} \tag{12}$$

We first have to do a substitution $z = x^2$ in order to have the expression in term of x^2 (which is the estimator : $Y = x^2$) and not in function of x. So we have

$$z = x^2 \quad \Rightarrow \quad \frac{dz}{dx} = 2x \quad \Rightarrow \quad dx = \frac{dz}{2x} = \frac{dz}{2\sqrt{z}} \tag{13}$$

The upper limit become $(\sqrt{s})^2 = s$

$$P_1(Y \leq s) = \sqrt{\frac{2}{\pi}} \int_{0}^{\sqrt{s}} \exp\left(-\frac{x^2}{2}\right) dx = \sqrt{\frac{2}{\pi}} \int_{0}^{s} \exp\left(-\frac{z}{2}\right) \frac{dz}{2\sqrt{z}} \tag{14}$$

$$p_1(Y = s) = \frac{dP_1(Y \leq s)}{ds} = \sqrt{\frac{2}{\pi}} \exp\left(-\frac{s}{2}\right) \frac{1}{2\sqrt{s}} \tag{15}$$

$$\boxed{p_1(Y = s) = \frac{1}{\sqrt{2\pi}\sqrt{s}} \exp\left(-\frac{s}{2}\right)} \tag{16}$$

This result will be fundamental for the rest of the demonstration because we will use it for each step of the induction.

1.2.2 $N_{dof} = 2$

$$p_2(Y=s) = \int_0^s p_1(s-\alpha)p_1(\alpha)d\alpha \tag{17}$$

$$p_2(Y=s) = \int_0^s \frac{1}{\sqrt{2\pi}\sqrt{s-\alpha}}\exp\left(-\frac{s-\alpha}{2}\right)\frac{1}{\sqrt{2\pi}\sqrt{\alpha}}\exp\left(-\frac{\alpha}{2}\right)d\alpha \tag{18}$$

$$p_2(Y=s) = \frac{1}{2\pi}\int_0^s \frac{1}{\sqrt{\alpha}\sqrt{s-\alpha}}\exp\left(-\frac{s-\alpha+\alpha}{2}\right)d\alpha \tag{19}$$

$$p_2(Y=s) = \frac{1}{2\pi}\exp\left(-\frac{s}{2}\right)\underbrace{\int_0^s \frac{1}{\sqrt{\alpha}\sqrt{s-\alpha}}d\alpha}_{=\pi} \tag{20}$$

$$\boxed{p_2(Y=s) = \frac{1}{2}\exp\left(-\frac{s}{2}\right)} \tag{21}$$

1.2.3 $N_{dof} = 3$

$$p_3(Y=s) = \int_0^s p_2(s-\alpha)p_1(\alpha)d\alpha \tag{22}$$

$$p_3(Y=s) = \int_0^s \frac{1}{2}\exp\left(-\frac{s-\alpha}{2}\right)\frac{1}{\sqrt{2\pi}\sqrt{\alpha}}\exp\left(-\frac{\alpha}{2}\right)d\alpha \tag{23}$$

$$p_3(Y=s) = \frac{1}{2\sqrt{2\pi}}\int_0^s \frac{1}{\sqrt{\alpha}}\exp\left(-\frac{s-\alpha+\alpha}{2}\right)d\alpha \tag{24}$$

$$p_3(Y=s) = \frac{1}{2\sqrt{2\pi}}\exp\left(-\frac{s}{2}\right)\underbrace{\int_0^s \frac{1}{\sqrt{\alpha}}d\alpha}_{=2\sqrt{s}} \tag{25}$$

$$\boxed{p_3(Y=s) = \frac{\sqrt{s}}{\sqrt{2\pi}}\exp\left(-\frac{s}{2}\right)} \tag{26}$$

1.2.4 $N_{dof} = 4$

$$p_4(Y=s) = \int_0^s p_2(s-\alpha)p_2(\alpha)d\alpha \tag{27}$$

$$p_4(Y=s) = \int_0^s \frac{1}{2}\exp\left(-\frac{s-\alpha}{2}\right)\frac{1}{2}\exp\left(-\frac{\alpha}{2}\right)d\alpha \tag{28}$$

$$p_4(Y=s) = \frac{1}{4}\exp\left(-\frac{s}{2}\right)\underbrace{\int_0^s d\alpha}_{=s} \tag{29}$$

$$\boxed{p_4(Y=s) = \frac{s}{4}\exp\left(-\frac{s}{2}\right)} \tag{30}$$

1.2.5 $N_{dof} = 5$

$$p_5(Y=s) = \int_0^s p_2(s-\alpha)p_3(\alpha)d\alpha \tag{31}$$

$$p_5(Y=s) = \int_0^s \frac{1}{2}\exp\left(-\frac{s-\alpha}{2}\right)\frac{\sqrt{\alpha}}{\sqrt{2\pi}}\exp\left(-\frac{\alpha}{2}\right)d\alpha \tag{32}$$

$$p_5(Y=s) = \frac{1}{2\sqrt{2\pi}}\exp\left(-\frac{s}{2}\right)\underbrace{\int_0^s \sqrt{\alpha}d\alpha}_{=\frac{2}{3}s^{3/2}} \tag{33}$$

$$\boxed{p_5(Y=s) = \frac{s^{3/2}}{3\sqrt{2\pi}} \exp\left(-\frac{s}{2}\right)} \tag{34}$$

1.2.6 $N_{dof} = k$

We can see a pattern in the previous results which is:

$$p_k(Y=s) \propto s^{k/2-1} \exp\left(-\frac{s}{2}\right) \tag{35}$$

We have to use the induction to prove it. So we can start assuming that this relation is true for $N_{dof} = k$ and show that it's true for $N_{dof} = k+1$. I propose to use an even number for k. We can write the probability ti have $Y = s$ considering all the possibilities to have $Y_k = s - \alpha$ and $Y_1 = \alpha$.

$$p_{k+1}(Y=s) = \int_0^s p_k(s-\alpha)p_1(\alpha)d\alpha \tag{36}$$

$$p_{k+1}(Y=s) \propto \int_0^s (s-\alpha)^{k/2-1} \exp\left(-\frac{s-\alpha}{2}\right)\alpha^{-1/2} \exp\left(-\frac{\alpha}{2}\right)d\alpha \tag{37}$$

$$p_{k+1}(Y=s) \propto \exp\left(-\frac{s}{2}\right) \int_0^s (s-\alpha)^{k/2-1}\alpha^{-1/2}d\alpha \tag{38}$$

We can develop $(s-\alpha)^{k/2-1}$ with the Binomial theorem as:

$$\binom{n}{p} = \frac{n!}{p!(n-p)!}. \tag{39}$$

In order to simplify the notations we will use $\beta = k/2 - 1$ for the next development.

$$(s-\alpha)^\beta = \sum_{i=0}^{\beta} \binom{\beta}{i} s^i(-\alpha)^{\beta-i} \tag{40}$$

$$p_{k+1}(Y=s) \propto \exp\left(-\frac{s}{2}\right) \int_0^s \sum_{i=0}^{\beta} \binom{\beta}{i} s^i(-\alpha)^{\beta-i}\alpha^{-1/2}d\alpha \tag{41}$$

Since the sum converge (because the result is a binomial development for s and α which are finite numbers) we can reverse the sum and th integral. Moreover, the integral is over α so s^i can be remove from the integral too:

$$p_{k+1}(Y=s) \propto \exp\left(-\frac{s}{2}\right) \sum_{i=0}^{\beta} \binom{\beta}{i} s^i(-1)^{\beta-i} \int_0^s (\alpha)^{\beta-i}\alpha^{-1/2}d\alpha \tag{42}$$

$$p_{k+1}(Y=s) \propto \exp\left(-\frac{s}{2}\right) \sum_{i=0}^{\beta} \binom{\beta}{i} s^i(-1)^{\beta-i} \int_0^s (\alpha)^{\beta-i-1/2}d\alpha \tag{43}$$

$$p_{k+1}(Y=s) \propto \exp\left(-\frac{s}{2}\right) \sum_{i=0}^{\beta} \binom{\beta}{i} s^i(-1)^{\beta-i} \left[\frac{1}{\beta-i-1/2+1}\alpha^{\beta-i-1/2+1}\right]_0^s \tag{44}$$

$$p_{k+1}(Y=s) \propto \exp\left(-\frac{s}{2}\right) \sum_{i=0}^{\beta} \binom{\beta}{i} (-1)^{\beta-i} \frac{1}{\beta-i-1/2+1} s^{\beta-i-1/2+1} \times s^i \tag{45}$$

$$p_{k+1}(Y=s) \propto \sum_{i=0}^{\beta} \binom{\beta}{i} (-1)^{\beta-i} \frac{1}{\beta-i-1/2+1} s^{\beta-1/2+1} \exp\left(-\frac{s}{2}\right) \tag{46}$$

We found finaly that the term in s doesn't depend anymore on i and so on the sum. The sum is just a factor. We rewrite with $\beta = k/2 - 1$ and we obtain that:

$$p_{k+1}(Y=s) \propto \left[\sum_{i=0}^{k/2-1} \binom{\beta}{i} (-1)^{\beta-i} \frac{1}{k/2-1-i-1/2+1}\right] \times s^{k/2-1-1/2+1} \exp\left(-\frac{s}{2}\right) \tag{47}$$

$$\boxed{p_{k+1}(Y=s) \propto s^{(k+1)/2-1} \exp\left(-\frac{s}{2}\right)} \tag{48}$$

so we found the good proportionality and the induction demonstration is finished. We still have to calculate the factor of proportionality to obtain the exact PDF. We know that all the $p_k(Y=S)$ are probabilities so that the integral of them have to be equal to 1. We just have to find the value of the proportionality ψ_k defined as:

$$\int_0^\infty p_k(Y=s)ds = 1 = \psi_k \times \int_0^\infty s^{k/2-1} \exp\left(-\frac{s}{2}\right) ds \tag{49}$$

We can recognize the form of the gamma funtion:

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt, \tag{50}$$

but we have to make a variable substitution to have the good expression in the exponential term. So we do the following variable substitution:

$$z = s/2 \quad \Rightarrow \quad \frac{dz}{ds} = 1/2 \quad \Rightarrow \quad ds = 2 \times dz, \tag{51}$$

so we have that the integral become:

$$\begin{aligned}
1 &= \psi_k \times \int_0^\infty s^{k/2-1} \exp\left(-\frac{s}{2}\right) ds && (52)\\
1 &= \psi_k \times \int_0^\infty (2z)^{k/2-1} e^{(-z)} 2dz && (53)\\
1 &= \psi_k \times 2^{k/2-1} \times 2 \times \int_0^\infty z^{k/2-1} e^{(-z)} dz && (54)\\
1 &= \psi_k \times 2^{k/2} \times \Gamma(k/2) && (55)\\
\psi_k &= \frac{1}{2^{k/2} \times \Gamma(k/2)} && (56)
\end{aligned}$$

So finally we obtain that:

$$\boxed{p_k(Y=s) = \frac{1}{2^{k/2} \times \Gamma(k/2)} s^{k/2-1} \exp\left(-\frac{s}{2}\right)} \tag{57}$$

We can finally derive the cumulative distribution function which is $P_k(Y \leq s)$ so by definition:

$$P_k(Y \leq s) = \int_0^s p_k(Y=x)dx. \tag{58}$$

This integral can be re-write using the incomplete gamma function $\gamma(k/2, s/2)$. The definition of the incomplete gamma function is:

$$\gamma(k,s) = \int_0^s x^{k-1} e^{-x} dx \tag{59}$$

where one more time we have to subistitute the variable in the exponential because we have $-x/2$ and not $-x$. So one more time we use:

$$z = x/2 \quad \Rightarrow \quad \frac{dz}{dx} = 1/2 \quad \Rightarrow \quad dx = 2 \times dz, \tag{60}$$

and we can write

$$\begin{aligned}
P_k(Y \leq s) &= \int_0^s \frac{1}{2^{k/2} \times \Gamma(k/2)} x^{k/2-1} \exp\left(-\frac{x}{2}\right) dx && (61)\\
P_k(Y \leq s) &= \frac{1}{2^{k/2} \times \Gamma(k/2)} \int_0^s (2z)^{k/2-1} \exp(-z)(2dz) && (62)\\
P_k(Y \leq s) &= \frac{2^{k/2-1} \times 2}{2^{k/2} \times \Gamma(k/2)} \underbrace{\int_0^s z^{k/2-1} \exp(-z) dz}_{=\gamma(k/2, s/2)} && (63)
\end{aligned}$$

So finally we have the χ^2 cumulative distribution function for k degree of freedom up to the value $Y = s$ as:

$$\boxed{P_k(Y \leq s) = \frac{\gamma(k/2, s/2)}{\Gamma(k/2)}} \tag{64}$$

2 Fisher Matrix and $\Delta\chi^2$ law

The idea of the Fisher Matrix is to assume a fiducial cosmology (i.e a set of parameters assumed to be the good one) and see how well we are able to constrain these parameters using the second derivatien of the log-Likelihood.

In our context, we will use the relation between the Likelihood and the χ^2 estimator when the number of degree of freedom is large (when the χ^2 pdf is similar to a gaussian). In this case, we have $\mathcal{L} = \exp -\chi^2/2$

As we saw in the project, the general estimation for the χ^2 is:

$$\chi^2(\vec{\theta}|\vec{y}^D) = \left(\vec{y}^{mod}(\vec{\theta}) - \vec{y}^D\right)^t Cov^{-1} \left(\vec{y}^{mod}(\vec{\theta}) - \vec{y}^D\right), \tag{65}$$

where $\vec{y}^D$ is the vector of data, $\vec{y}^{mod}(\vec{\theta})$ is the vector of points from the model for the parameters $\vec{\theta}$ and Cov^{-1} is invert covariance matrix.

We derive in the first section the p-value from the χ^2 distribution. This p-value tell us the probability to produce the set of data $\vec{y}^D$ assuming a model $\vec{\theta}$ and gaussian errors.

However, the p-values contours we obtain are not equivalent to the ones obtained using the MCMC approach. The MCMC approach is only sensitive on the difference of the likelihood (i.e χ^2 for most of the cosmological observations where $\mathcal{L} = \exp -\chi^2/2$) value in the parameter space. So the constrain on parameters depends on the derivative of the likelihood (χ^2) value arround the best model.

We will express the Fisher matrix and then develop the $\Delta\chi^2$ at the second order in Taylor serie.

2.1 $\Delta\chi^2$ developpment

We consider first the case where we have 2 parameters $\vec{\theta} = (\theta_1, \theta_2)$. After mesure the values for the $\chi^2(\theta_1, \theta_2)$ we found minimum value which correspond to the best-fit parameters ($\vec{\theta^b} = (\theta_1^b, \theta_2^b)$) .

We do the Taylor expansion, to the second order, arround the best parameter values (θ_1^b, θ_2^b) :

$$\begin{aligned} \chi^2(\theta_1^b + \Delta\theta_1, \theta_2^b + \Delta\theta_2) &= \chi^2(\theta_1^b, \theta_2^b) + \Delta\theta_1 \overbrace{\frac{\partial\chi^2(\theta_1^b, \theta_2^b)}{\partial\theta_1}}^{\equiv 0} + \Delta\theta_2 \overbrace{\frac{\partial\chi^2(\theta_1^b, \theta_2^b)}{\partial\theta_2}}^{\equiv 0} + \frac{\Delta\theta_1^2}{2}\frac{\partial^2\chi^2(\theta_1^b, \theta_2^b)}{\partial\theta_1^2} + & (66) \\ &+ \frac{\Delta\theta_2^2}{2}\frac{\partial^2\chi^2(\theta_1^b, \theta_2^b)}{\partial\theta_2^2} + 2 \times \frac{\Delta\theta_1\Delta\theta_2}{2}\frac{\partial^2\chi^2(\theta_1^b, \theta_2^b)}{\partial\theta_1\partial\theta_2} + \mathcal{O}(\Delta\theta^3), & (67) \end{aligned}$$

where the 2 terms of first derivative are null by definition of the minimum of a function. Here we assumed that the second derivative between θ_1 and θ_2 commute So we can re-write the calculation as:

$$\chi^2(\theta_1^b + \Delta\theta_1, \theta_2^b + \Delta\theta_2) \simeq \chi^2(\theta_1^b, \theta_2^b) + \frac{\Delta\theta_1^2}{2}\frac{\partial^2\chi^2(\theta_1^b, \theta_2^b)}{\partial\theta_1^2} + \frac{\Delta\theta_2^2}{2}\frac{\partial^2\chi^2(\theta_1^b, \theta_2^b)}{\partial\theta_2^2} + 2 \times \frac{\Delta\theta_1\Delta\theta_2}{2}\frac{\partial^2\chi^2(\theta_1^b, \theta_2^b)}{\partial\theta_1\partial\theta_2} \tag{68}$$

$$\chi^2(\theta_1^b + \Delta\theta_1, \theta_2^b + \Delta\theta_2) \simeq \chi^2(\theta_1^b, \theta_2^b) + \frac{\Delta\theta_1^2}{2}\frac{\partial^2\chi^2(\theta_1^b, \theta_2^b)}{\partial\theta_1^2} + \frac{\Delta\theta_2^2}{2}\frac{\partial^2\chi^2(\theta_1^b, \theta_2^b)}{\partial\theta_2^2} + 2 \times \frac{\Delta\theta_1\Delta\theta_2}{2}\frac{\partial^2\chi^2(\theta_1^b, \theta_2^b)}{\partial\theta_1\partial\theta_2} \tag{69}$$

$$\Delta\chi^2(\Delta\theta_1, \Delta\theta_2) \equiv \chi^2(\theta_1^b + \Delta\theta_1, \theta_2^b + \Delta\theta_2) - \chi^2(\theta_1^b, \theta_2^b). \tag{70}$$

$$\Delta\chi^2(\Delta\theta_1, \Delta\theta_2) \simeq \frac{\Delta\theta_1^2}{2}\frac{\partial^2\chi^2(\theta_1^b, \theta_2^b)}{\partial\theta_1^2} + \frac{\Delta\theta_2^2}{2}\frac{\partial^2\chi^2(\theta_1^b, \theta_2^b)}{\partial\theta_2^2} + 2 \times \frac{\Delta\theta_1\Delta\theta_2}{2}\frac{\partial^2\chi^2(\theta_1^b, \theta_2^b)}{\partial\theta_1\partial\theta_2} \tag{71}$$

We can use the matrix formalism to rewrite the same equation as :

$$\boxed{\Delta\chi^2(\Delta\theta_1, \Delta\theta_2) \simeq \frac{1}{2}\begin{pmatrix}\Delta\theta_1 & \Delta\theta_2\end{pmatrix}\begin{pmatrix}\frac{\partial^2\chi^2(\theta_1^b,\theta_2^b)}{\partial\theta_1^2} & \frac{\partial^2\chi^2(\theta_1^b,\theta_2^b)}{\partial\theta_1\partial\theta_2} \\ \frac{\partial^2\chi^2(\theta_1^b,\theta_2^b)}{\partial\theta_1\partial\theta_2} & \frac{\partial^2\chi^2(\theta_1^b,\theta_2^b)}{\partial\theta_2^2}\end{pmatrix}\begin{pmatrix}\Delta\theta_1 \\ \Delta\theta_2\end{pmatrix}} \tag{72}$$

One time the elements of this matrix are evaluated, we can apply the equaion 72 and generate the approximate value for each value of $\Delta\theta_1$ and $\Delta\theta_2$. Then we can apply the $\Delta\chi^2$ table values for 2 parameters in order to find the contours. That's the real way to use the $\Delta\chi^2$ table values, not using the χ^2 array and substracting the χ^2_{min}. The result will be similar only if the χ^2 values follow perfectly a second degree polynomial.

The advantage with the matrix formalism is that we can write the second order derivative for N parameters $\vec{\theta} = (\theta_1 \dots \theta_N)$

$$\boxed{\Delta\chi^2(\Delta\theta_1 \dots \Delta\theta_N) \simeq \frac{1}{2}\begin{pmatrix}\Delta\theta_1 & \Delta\theta_2 & \dots & \Delta\theta_N\end{pmatrix}\begin{pmatrix}\frac{\partial^2\chi^2(\theta_1^b \dots \theta_N^b)}{\partial\theta_1^2} & \frac{\partial^2\chi^2(\theta_1^b \dots \theta_N^b)}{\partial\theta_1\partial\theta_2} & \dots & \frac{\partial^2\chi^2(\theta_1^b \dots \theta_N^b)}{\partial\theta_1\partial\theta_N} \\ \frac{\partial^2\chi^2(\theta_1^b \dots \theta_N^b)}{\partial\theta_1\partial\theta_2} & \frac{\partial^2\chi^2(\theta_1^b \dots \theta_N^b)}{\partial\theta_2^2} & \dots & \frac{\partial^2\chi^2(\theta_1^b \dots \theta_N^b)}{\partial\theta_2\partial\theta_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2\chi^2(\theta_1^b \dots \theta_N^b)}{\partial\theta_1\partial\theta_N} & \frac{\partial^2\chi^2(\theta_1^b \dots \theta_N^b)}{\partial\theta_2\partial\theta_N} & \dots & \frac{\partial^2\chi^2(\theta_1^b \dots \theta_N^b)}{\partial\theta_N^2}\end{pmatrix}\begin{pmatrix}\Delta\theta_1 \\ \Delta\theta_2 \\ \vdots \\ \Delta\theta_N\end{pmatrix}} \tag{73}$$

and we can apply the $\Delta\chi^2$ table values for N parameters in order to compute the confidence contours. **The important point here is that we interpret the $\Delta\chi^2$ as the posterior distribution (when the χ^2 as the likelihood) so the number of degree of freedom changes! For the χ^2 we consider the errors over the data points, so the number of degree of freedom is given by the number of data points minus the number of parameters of the model. For the $\Delta\chi^2$ we assume that the error is due to the incertitude of the parameters and so the number of degree of freedom become the number of parameters of the model.**

2.2 Fisher Matrix

The Fisher information (i.e the Fisher Matrix) is a relation between the espectation of the log Likelihood $\log\mathcal{L}$

The usual idea is to use the Fisher Matrix before to have data in order to forecast the typical constraints on parameters assuming a model as the reality and errors (so the covariance matrix). Doing that, the vector of data is given by the fiducial model and we want see how the likelihood change when we move the parameters around. We can so do the same exercise we did for the $\Delta\chi^2$

The definition of the Fisher Matrix is :

$$F_{ij} = -\left\langle \frac{\partial^2 \log(\mathcal{L}(\theta_1^b \dots \theta_N^b))}{\partial\theta_i\partial\theta_j} \right\rangle \tag{74}$$

In the case where the likelihood is $\mathcal{L} = \exp -\chi^2/2$ then we have :

$$\log(\mathcal{L}(\theta_1^b \dots \theta_N^b)) = -\frac{1}{2}\chi^2(\theta_1^b \dots \theta_N^b), \tag{75}$$

and so that :

$$F_{ij} = \frac{1}{2}\left\langle \frac{\partial^2\chi^2(\theta_1^b \dots \theta_N^b)}{\partial\theta_i\partial\theta_j} \right\rangle, \tag{76}$$

where we assume $\vec{y}^D \equiv \vec{y}(\theta_1^b \dots \theta_N^b)$, so where the χ^2 here is:

$$\chi^2(\vec{\theta}|\vec{y}^D) = \chi^2(\vec{\theta}|\vec{y}^{mod}(\vec{\theta^b})) = \left(\vec{y}^{mod}(\vec{\theta}) - \vec{y}^{mod}(\vec{\theta^b})\right)^t Cov^{-1}\left(\vec{y}^{mod}(\vec{\theta}) - \vec{y}^{mod}(\vec{\theta^b})\right). \tag{77}$$

2.2.1 How to use it

The information we have here is a sketch of the likelihood surface we evaluate in the various examples. However, it is not the real result but a second order approximation of the result assuming the "best values of the set of parameters". With this tools, we are not evaluatng the best parameters, but we evaluate the quality of the constraints (contours) assuming the best parameters and the error over measurements (i.e the covariance matrix).

That's important when you want to estimate the utility of a future experiment that you are designing before to do it. On the other hand, that's a perfect tool to estimate the accuracy you have to target on your measurmements in order to constrain the cosmological parameters as well as needed.

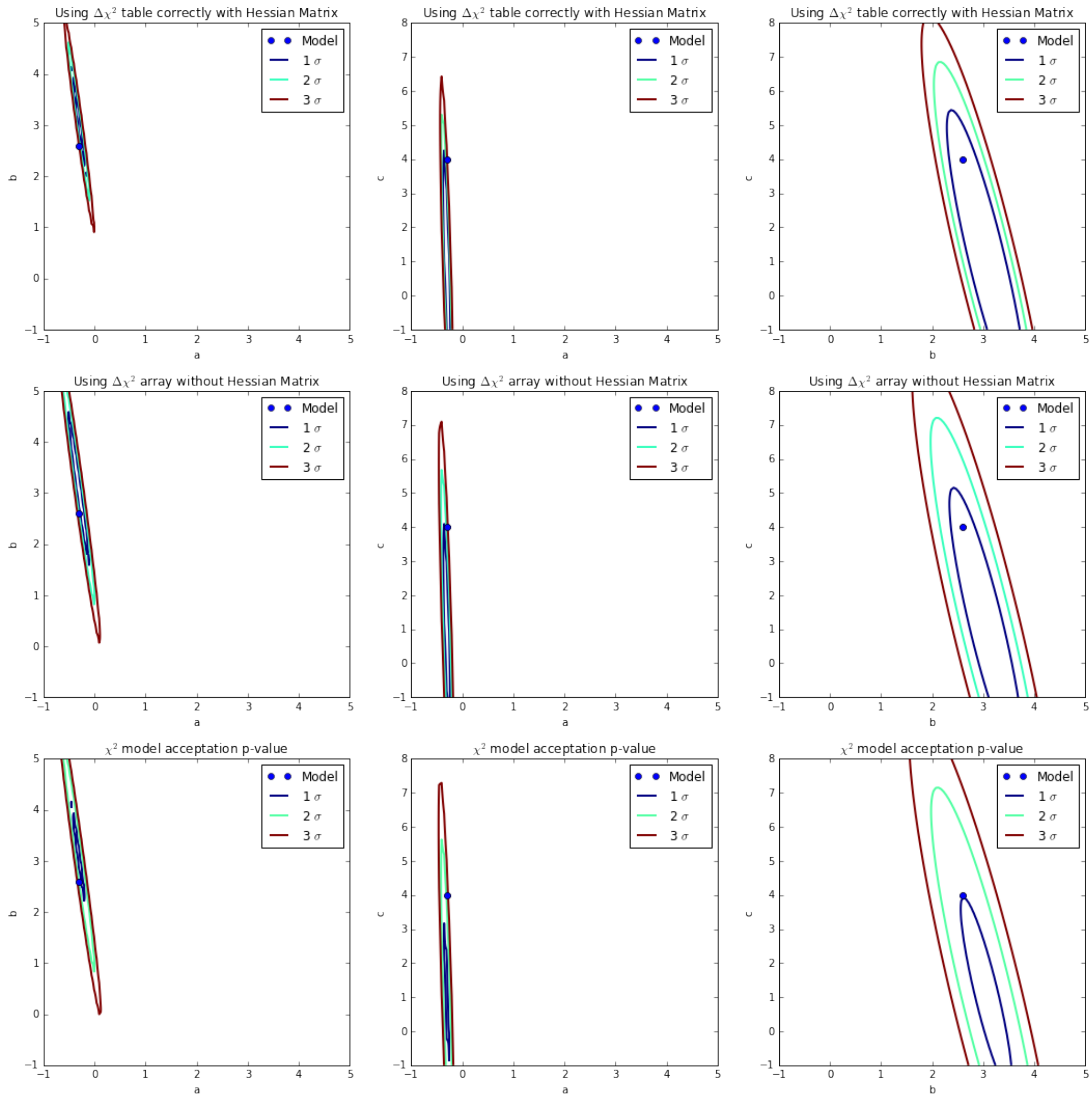

Figure 1: Comparaison of contours at 1σ, 2σ and 3σ for the p-values (from χ^2), the $\Delta\chi^2$ array confidence (wrong way to do) and the confidence contours from the Hessian Matrix of the χ^2 array evaluated arround the best-fit parameters.