The probability of observing the value $y_i \equiv y(t_i)$ when the expectation value is $\mu(t_i)$ and the error is gaussian is

$$f(y_i|\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\nu}} \exp\left(-\frac{1}{2}\frac{(y_i - \mu(t_i))^2}{\nu}\right),$$
$$\boldsymbol{\theta} = (\nu, \omega, \phi),$$
$$\mu(t_i) = N_0\left(1 + P\sin(\omega t_i + \phi)\right).$$

The likelihood of observing a set of observations $\mathbf{y} = (y_1, \ldots, y_K)$, under the i.i.d. assumption, is the product of propabilities taken as a function of the parameters:

$$\mathcal{L}(\boldsymbol{\theta}|\mathbf{y}) = \prod_i f(y_i|\boldsymbol{\theta}),$$

and the log-likelihood

$$\ell(\boldsymbol{\theta}|\mathbf{y}) = -\frac{K}{2}\log 2\pi - \frac{K}{2}\log\nu - \frac{1}{2\nu}\sum_i \epsilon_i^2,\ \ \epsilon_i = y_i - \mu(t_i).$$

The usual assumptions for the error term are zero expectation and strict exogeneity

$$\mathrm{E}\left[\epsilon_i|\ \boldsymbol{\theta}_0\right] = \mathrm{E}\left[t_i\epsilon_i|\ \boldsymbol{\theta}_0\right] = 0,$$

and the relations between the mean's derivatives are

$$\mu'_\phi = N_0 P\cos(\omega t + \phi),$$
$$\mu'_\omega = t\cdot\mu'_\phi, \epsilon'_\xi = -\mu'_\xi.$$

The log-likelihood derivatives:

$$\ell'_\nu = -\frac{K}{2\nu} + \frac{1}{2\nu^2}\sum_i \epsilon_i^2;$$

$$\ell'_\omega = \frac{1}{\nu}\sum_i \mu'_\phi(t_i)t_i\epsilon_i;$$

$$\ell'_\phi = \frac{1}{\nu}\sum_i \mu'_\phi(t_i)\epsilon_i;$$

$$\ell''_{\nu^2} = \frac{K}{2\nu^2} - \frac{1}{\nu^3}\sum_i \epsilon_i^2, \qquad -\mathrm{E}\left[\ell''_{\nu^2}|\ \boldsymbol{\theta}_0\right] = \frac{K}{2\nu^2} - \frac{1}{\nu^3}\sum_i \nu = \frac{K}{2\nu^2};$$

$$\ell''_{\nu\omega} = -\frac{1}{\nu^2}\sum_i \mu'_\phi(t_i)t_i\epsilon_i, \qquad -\mathrm{E}\left[\ell''_{\nu\omega}|\ \boldsymbol{\theta}_0\right] = \frac{1}{\nu^2}\sum_i \mu'_\phi(t_i)\mathrm{E}\left[t_i\epsilon_i|\ \boldsymbol{\theta}_0\right] = 0;$$

$$\ell''_{\nu\phi} = -\frac{1}{\nu^2}\sum_i \mu'_\phi(t_i)\epsilon_i, \qquad -\mathrm{E}\left[\ell''_{\nu\phi}|\ \boldsymbol{\theta}_0\right] = \frac{1}{\nu^2}\sum_i \mu'_\phi(t_i)\mathrm{E}\left[\epsilon_i|\ \boldsymbol{\theta}_0\right] = 0;$$

$$\ell''_{\phi^2} = \frac{1}{\nu}\sum_i \left(\mu''_{\phi^2}(t_i)\epsilon_i - \left(\mu'_\phi(t_i)\right)^2\right), \qquad -\mathrm{E}\left[\ell''_{\phi^2}|\ \boldsymbol{\theta}_0\right] = \frac{1}{\nu}\sum_i \left(\left(\mu'_\phi(t_i)\right)^2 - \mu''_{\phi^2}(t_i)\mathrm{E}\left[\epsilon_i|\ \boldsymbol{\theta}_0\right]\right) = \frac{1}{\nu}\sum_i \left(\mu'_\phi(t_i)\right)^2;$$

$$\ell''_{\phi\omega} = \frac{1}{\nu}\sum_i \left(\mu''_{\phi^2}(t_i)t_i\epsilon_i - \left(\mu'_\phi(t_i)\right)^2 t_i\right), \qquad -\mathrm{E}\left[\ell''_{\phi\omega}|\ \boldsymbol{\theta}_0\right] = \frac{1}{\nu}\sum_i \left(t_i\left(\mu'_\phi(t_i)\right)^2 - \mu''_{\phi^2}(t_i)\mathrm{E}\left[t_i\epsilon_i|\ \boldsymbol{\theta}_0\right]\right) = \frac{1}{\nu}\sum_i t_i\left(\mu'_\phi(t_i)\right)^2;$$

$$\ell''_{\omega^2} = \frac{1}{\nu}\sum_i \left(\mu''_{\phi^2}(t_i)t_i^2\epsilon_i - \left(\mu'_\phi(t_i)t_i\right)^2\right), \qquad -\mathrm{E}\left[\ell''_{\omega^2}|\ \boldsymbol{\theta}_0\right] = \frac{1}{\nu}\sum_i \left(\left(t_i\mu'_\phi(t_i)\right)^2 - \mu''_{\phi^2}(t_i)\mathrm{E}\left[t_i^2\epsilon_i|\ \boldsymbol{\theta}_0\right]\right) = \frac{1}{\nu}\sum_i \left(t_i\mu'_\phi(t_i)\right)^2.$$

1 Variances

The Fisher matrix

$$I(\boldsymbol{\theta}_0) = \begin{pmatrix} K/2\nu & 0 & 0 \\ 0 & 1/\nu\sum\left(t_i\mu'_\phi(t_i)\right)^2 & 1/\nu\sum t_i\left(\mu'_\phi(t_i)\right)^2 \\ 0 & 1/\nu\sum t_i\left(\mu'_\phi(t_i)\right)^2 & 1/\nu\sum\left(\mu'_\phi(t_i)\right)^2 \end{pmatrix}.$$

The determinant

$$|I(\boldsymbol{\theta}_0)| = \frac{K}{2\nu^4}\underbrace{\left(\sum\left(t_i\mu'_\phi(t_i)\right)^2\sum\left(\mu'_\phi(t_i)\right)^2 - \left(\sum t_i\left(\mu'_\phi(t_i)\right)^2\right)^2\right)}_{\Omega}.$$

The variance-covariance matrix

$$vcov = \begin{pmatrix} 2\nu^2/K & 0 & 0 \\ 0 & \nu\frac{\sum\left(\mu'_\phi(t_i)\right)^2}{\Omega} & \nu\frac{\sum t_i\left(\mu'_\phi(t_i)\right)^2}{\Omega} \\ 0 & \nu\frac{\sum t_i\left(\mu'_\phi(t_i)\right)^2}{\Omega} & \nu\frac{\sum\left(t_i\mu'_\phi(t_i)\right)^2}{\Omega} \end{pmatrix}.$$

The variance of the frequency estimate

$$\mathrm{var}\left[\hat{\omega}\right] = \nu\frac{\sum\left(\mu'_\phi(t_i)\right)^2}{\sum\left(t_i\mu'_\phi(t_i)\right)^2\sum\left(\mu'_\phi(t_i)\right)^2 - \left(\sum t_i\left(\mu'_\phi(t_i)\right)^2\right)^2}. \tag{1}$$

Cross-check. Let $\mu(t_i) = \phi + \omega t_i$. In that case $\mu'_\phi(t_i) = 1$, $\mu'_\omega(t_i) = t_i = t_i \cdot \mu'_\phi(t_i)$, the determinant of the Fisher matrix simplifies to

$$\begin{aligned} |I(\boldsymbol{\theta}_0)| &= \frac{K}{2\nu^4}\left(K\sum_i t_i^2 - \left(\sum t_i\right)^2\right) \\ &= \frac{K^3}{2\nu^4}\left(\frac{1}{K}\sum t_i^2 - \langle t\rangle^2\right) \\ &= \frac{K}{2\nu^4}\cdot\underbrace{K\sum\left(t_i - \langle t\rangle\right)^2}_{\Omega} \end{aligned}$$

and the variance-covariance matrix becomes

$$vcov = \begin{pmatrix} 2\nu^2/K & 0 & 0 \\ 0 & \frac{\nu}{\sum(t_i-\langle t\rangle)^2} & \nu\frac{\sum t_i}{K\sum(t_i-\langle t\rangle)^2} \\ 0 & \nu\frac{\sum t_i}{K\sum(t_i-\langle t\rangle)^2} & \nu\frac{\sum t_i^2}{K\sum(t_i-\langle t\rangle)^2} \end{pmatrix},$$

with the well-known expression for the slope variance

$$\mathrm{var}\left[\hat{\omega}\right] = \frac{\nu}{\sum\left(t_i - \langle t\rangle\right)^2}.$$

Let us denote $\left(\mu'_\phi(t_i)\right)^2 = (N_0P)^2\cos^2(\omega t_i + \phi) \equiv x_i$. Eq. (1) can be rewritten in the following form:

$$\begin{aligned} \mathrm{var}\left[\hat{\omega}\right] &= \frac{\nu}{\sum_j x_j\left(\sum_i t_i^2\frac{x_i}{\sum_j x_j} - \left(\sum_i t_i\frac{x_i}{\sum_j x_j}\right)^2\right)} \\ &= \frac{\nu}{\sum_j x_j\sum_i w_i\left(t_i - \langle t\rangle_w\right)^2} \\ &= \frac{\nu}{\sum_j x_j\cdot\mathrm{var}_w\left[t\right]}. \end{aligned} \tag{2}$$

In matrix form, the frequency variance is written as

$$\mathrm{var}\left[\hat{\omega}\right] = \frac{\nu}{\left(\underline{\mathrm{T}}'\mathcal{D}_\mu^2\underline{\mathrm{T}}\right) - \left(\underline{\mathrm{T}}'\mathcal{D}_\mu\underline{\mathrm{M}}\right)^2/\left(\underline{\mathrm{M}}'\underline{\mathrm{M}}\right)},$$

with

$$\underline{\mathrm{T}} = \left(t_0, \ldots, t_{K-1}\right)',\ \underline{\mathrm{M}} = \left(\mu'_\phi(t_0), \ldots, \mu'_\phi(t_{K-1})\right)',$$

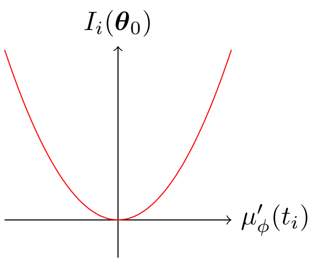
$$\mathcal{D}_\mu = \begin{pmatrix} \mu'_\phi(t_0) & 0 & \cdots & 0 \\ 0 & \mu'_\phi(t_1) & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu'_\phi(t_{K-1}) \end{pmatrix}.$$

2 Sampling modulation

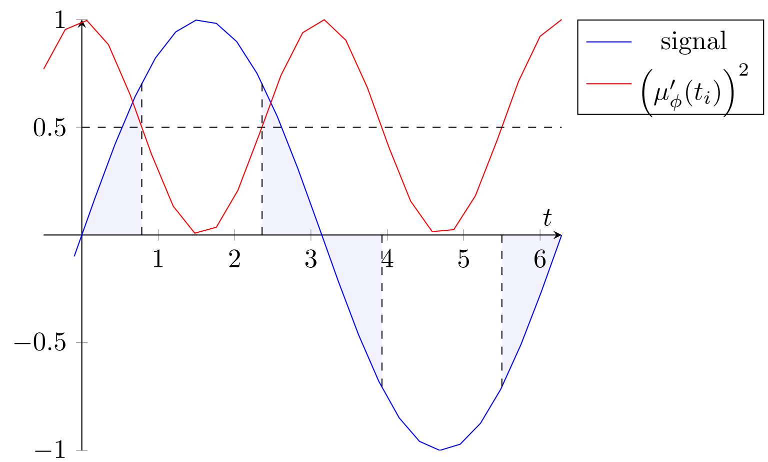
Suppose we write the Fisher matrix as a sum:

$$I(\boldsymbol{\theta}_0) = \sum_i I_i(\boldsymbol{\theta}_0);\; I_i(\boldsymbol{\theta}_0) = \frac{1}{\nu}\begin{pmatrix} \left(\sqrt{2}\cdot\mu'_\phi(t_i)\right)^{-2} & 0 & 0 \\ 0 & t_i^2 & t_i \\ 0 & t_i & 1 \end{pmatrix} \cdot \left(\mu'_\phi(t_i)\right)^2. \tag{3}$$

$I_i(\boldsymbol{\theta}_0) = -\mathrm{E}\left[\frac{\partial^2}{\partial\boldsymbol{\theta}^2}\log f(y_i|\boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}|\ \boldsymbol{\theta}_0\right]$ could be[1] interpreted as the information about the parameter that's carried in y_i.



(a) Fisher information of a point is a parabola of the signal derivative.



(b) Filled areas are where the points are more informative.

If we attribute each point a weight proportional to its Fisher information, i.e. $w_i = \cos^2(\omega t_i + \phi)$, [2] the weight of a region where $\left(\mu'_\phi(t_i)\right)^2 \geq 1/2$ is greater than that of an equivalent region with $\left(\mu'_\phi(t_i)\right)^2 < 1/2$ by the factor:

$$\int_{t_0}^{t_1}\cos^2(\omega t+\phi)\mathrm{d}t = \frac{1}{\omega}\int_{\omega t_0}^{\omega t_1}\cos^2\theta\mathrm{d}\theta = \frac{\Delta t}{2} + \frac{1}{2\omega}\sin\omega\Delta t\cos\omega\Sigma t \approx 1.9.$$

The implication is that increasing the number of points measured during the signal's rise and fall is roughly twice as beneficial as doing so during the peaks and troughs.

3 The dependence of $\mathrm{var}\left[\hat{\omega}\right]$ on ω

The variance in eq. (2) is conditional on the frequency ω and phase ϕ of the signal. The conditioning happens because the parameters determine the sampling range of the distribution of the x_i weights.

That distribution is

$$f(x_i) = \frac{1}{\pi}\cdot\left[\frac{x_i}{N_0P}\left(1-\frac{x_i}{N_0P}\right)\right]^{-1/2}.$$

$$\mathrm{E}\left[\sum_i x_i|\ (\omega,\phi)\right] = \sum_i\int_{\cos^2\phi}^{\cos^2(\omega t_i+\phi)} x_i f(x_i)\mathrm{d}x_i;$$

[1] The t_i in the structural matrix in eq. (3) worries me, because it appears that a point carries more information simply by virtue of it being measured later in time; but as far as I can tell the reason for it is that it is assumed that the point labeled as i is the i-th point in a series, and so a later point is more informative than a point closer to the origin, all other things being equal. And it's nothing new; in linear regression we also want our predictors to be as spread out as possible.

[2] The variance of ω is proportional to the (2,2)-minor, in which time doesn't figure, only the squared cosine.