Draft

6. maj 2014

1 Randomised algorithm

Uncertianty based on asymptotic liklihood and $\overline{w}$-distribution
Let $\mathcal{L}_\infty$ be the log-likelihood function for a distribution, now let $\mathcal{L}_N$ denote the log-likelihood function based on N observations from this distribution. Furthermore, let N be a large number, for which $L_N \approx L_\infty$.

$$\mathcal{L}_\mathcal{N} = \frac{1}{N}\sum_{n=1}^{N} \ell_n \qquad \overline{w}\, s.t. \frac{\delta \mathcal{L}}{\delta \overline{w}} = \overline{0} \tag{1.1}$$

Where ℓ_n is the log-likelihood of the n^{th} observation. And $\overline{w}$ is the optimal **(true?)** weights for the distribution, then we combine the expressions from (1.1), such that for the optimal weights the following must be fulfilled:

$$\frac{1}{N}\sum_{n=1}^{N} \frac{\delta \ell_n}{\delta \overline{w}} = 0 \tag{1.2}$$

(Skal vi lige skrive lidt om at $\Delta w = w - w_0$ og er en lille forskydelse i vægtene? Eller er det en lille forskydelse?) For each of the N observations, we can write the log-likelihood of the n^{th} observation as:

$$\ell_n(\Delta\overline{w}) = \ell_n\left(\overline{w}_0\right) + \left.\frac{\delta \ell_n}{\delta \overline{w}}\right|_{\overline{w}_0} \Delta\overline{w} + \frac{1}{2} Tr\left[\left.\frac{\delta \ell_n}{\delta \overline{w} \delta \overline{w}^T}\right|_{\overline{w}_0} \Delta\overline{w}\Delta\overline{w}^T\right] \tag{1.3}$$

Or for the entire log-likelihood function:

$$\mathcal{L}_N(\Delta\overline{w}) = \mathcal{L}_\infty\left(\overline{w}_0\right) + \left(\left.\frac{\delta \mathcal{L}_N}{\delta \overline{w}}\right|_{\overline{w}_0}\right)^T \cdot \Delta\overline{w} + \frac{1}{2}\Delta\overline{w}^T \left(\left.\frac{\delta^2 \mathcal{L}_N}{\delta \overline{w} \delta \overline{w}^T}\right|_{\overline{w}_0}\right) \Delta\overline{w} + R \tag{1.4}$$

Where R is the error of the approximation. Furthermore, we define the functions $\overline{\overline{H}}_N = \left.\frac{\delta^2 \mathcal{L}_N}{\delta \overline{w} \delta \overline{w}^T}\right|_{\overline{w}_0}$, and $\overline{g} = \left.\frac{\delta \mathcal{L}_N}{\delta \overline{w}}\right|_{\overline{w}_0}$. And evaluate the condition on $\overline{w}$, stated in (1.1):

$$\frac{\delta \mathcal{L}_\mathcal{N}}{\delta \overline{w}} = \overline{g}_N + \overline{\overline{H}}_N \Delta\overline{w} = \overline{0} \tag{1.5}$$

We replace $\Delta\overline{w}$ with $\hat{\Delta\overline{w}}$ as N is a finite number, thus only approximating $\Delta\overline{w}$. Isolating $\hat{\Delta\overline{w}}$, and using Ljung [REFERENCE?], we get:

$$\hat{\Delta\overline{w}} = -\overline{\overline{H}}_N^{-1} \cdot \overline{g}_N \overbrace{=}^{Ljung} -\overline{\overline{H}}_0^{-1} \cdot \overline{g}_{\overline{w}}\left(\overline{w}_0\right) \tag{1.6}$$

(Forklaring af at H_0 er uafhængig af datasæt, mens g nu er afhængig af w evalueret i w_0) Besides getting an estimate for $\hat{\Delta\overline{w}}$, we can find the mean of the distribution:

$$\left\langle \hat{\Delta\overline{w}} \right\rangle = -\overline{\overline{H}}_0^{-1} \left\langle \overline{g}_w \right\rangle \left(\overline{w}_0\right)$$

Nu skal vi prøve at kombinere flere datasets