MRN078 Maximum Entropy in Inverse Problems June 21, 2010

Proof that Normal Distribution maximizes entropy S over all normalized distributions with mean $\hat{m}$ and variance σ^2.

$$S = \int p(m) \ln(p(m)) \, dm$$

minimize/maximize S with constraints using Euler/Lagrange eqns

$$\int p(m)\, dm = 1 \qquad \int p(m)(m-\hat{m})\, dm = 0$$

$$\int p(m)(m-\hat{m})^2 dm = \sigma^2$$

$$\Phi = p(m) \ln p(m) + \lambda_1 p(m) + \lambda_2 p(m)(m-\hat{m}) + \lambda_3 p(m)(m-\hat{m})^2$$

$$\frac{\partial \Phi}{\partial p} = \ln p(m) + 1 + \lambda_1 + \lambda_2(m-\hat{m}) + \lambda_3(m-\hat{m})^2 = 0$$

$$\ln p(m) = (-1-\lambda_1) - \lambda_2(m-\hat{m}) - \lambda_3(m-\hat{m})^2$$

constraints require $(-1-\lambda_1) = \ln \frac{1}{\sqrt{2\pi}\,\sigma}$ $\quad \lambda_2 = 0 \quad \lambda_3 = \frac{1}{2\sigma^2}$

so $p(m) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{ -\frac{(m-\hat{m})^2}{2\sigma^2} \right\}$

now let $S = \int p(m) \ln p(m)/q(m)\, dm$ with $q(m)$ prescribed, and minimize with constraints

$\int p(m) dm = 1 \qquad \int p(m)(\underline{d} - \underline{\underline{G}}\,\underline{m}) = 0$

$$\Phi = p(m) \ln p(m)/q(m) + \lambda_1 p(m) + \underline{\lambda}_2^T p(m)(\underline{d} - \underline{\underline{G}}\,\underline{m})$$

$\hookrightarrow p\ln p - p \ln q$

$$0 = \frac{\partial \Phi}{\partial p} = \ln p + 1 - \ln q + \lambda_1 + \underline{\lambda}_2^T(\underline{d} - \underline{\underline{G}}\,\underline{m})$$

$$p(m) = q(m) \exp\{-(1+\lambda_1) - \underline{\lambda}_2^T(\underline{d} - \underline{\underline{G}}\,\underline{m})\}$$

now suppose $q(m) = \exp\{-\tfrac{1}{2}(\underline{m}-\underline{s})^T \underline{\underline{C}}_m^{-1}(\underline{m}-\underline{s})\}$

$p(m) = \exp\{-J(m)\}$ with

$$J(m) = \tfrac{1}{2}(\underline{m}-\underline{s})^T \underline{\underline{C}}_m^{-1}(\underline{m}-\underline{s}) + (1+\lambda) - \underline{\lambda}_2^T(d - Gm)$$

now mean $\hat{m}$ also max. likelihood point satisfies $\left.\frac{\partial J}{\partial m}\right|_{m=\hat{m}} = 0$

$$\left.\frac{\partial J}{\partial m}\right|_{m=\hat{m}} = 0 = \underline{\underline{C}}_m^{-1}(\underline{\hat{m}}-\underline{s}) + \underline{\underline{G}}^T\underline{\lambda}$$

assume C_m exists, so $(\hat{m}-s) = -C_m G^T\lambda$

premultiply by G to get $\underline{\underline{G}}\underline{\hat{m}} - \underline{\underline{G}}\underline{s} = -\underline{\underline{C}}_m\underline{\underline{G}}^T\underline{\lambda}$

and use $\underline{\underline{G}}\underline{\hat{m}} = \underline{d}$ to get $(\underline{d} - \underline{\underline{G}}\underline{s}) = -\underline{\underline{G}}\underline{\underline{C}}_m\underline{\underline{G}}^T\underline{\lambda}$

assume $\underline{\underline{G}}\underline{\underline{C}}_m\underline{\underline{G}}^T$ has inverse and solve for $\underline{\lambda}$

$$\lambda = -(\underline{\underline{G}}\underline{\underline{C}}_m\underline{\underline{G}}^T)^{-1}(\underline{d} - \underline{\underline{G}}\underline{s})$$

and now plug back into $\partial J/\partial m = 0$ to get

$$(\hat{m}-s) = -\underline{\underline{C}}_m\underline{\underline{G}}^T\underline{\lambda}$$

$$(\hat{m}-s) = \underline{\underline{C}}_m\underline{\underline{G}}^T(\underline{\underline{G}}\underline{\underline{C}}_m\underline{\underline{G}}^T)^{-1}(\underline{d} - \underline{\underline{G}}\underline{s})$$

which is a "minimum length" type solution

Thus a minimum $\left\{ \begin{array}{l} \text{length} \\ \text{"ML"} \end{array} \right.$ solution can be derived from "minimum relative entropy" MRE considerations.