Supplementary material for: Gaussian Process Classification and Active Learning with Multiple Annotators

Filipe Rodrigues (fmpr@dei.uc.pt)
Francisco C. Pereira (camara@smart.mit.edu)
Bernardete Ribeiro (bribeiro@dei.uc.pt)

In this extra material, we provide more details on deriving the moments of the product of the cavity distribution with the exact likelihood term $\sum_{z_i \in \{0,1\}} p(\mathbf{y}_i|z_i)p(z_i|f_i)$, which constitutes the step 2 of EP referred in the paper. This derivation was ommited from the main text due to lack of space.

1 Moments derivation

Recall that the product of the cavity distribution with the exact likelihood term is given by:

$$\begin{aligned}\hat{q}(f_i) &\triangleq \hat{Z}_i \mathcal{N}(\hat{\mu}_i, \hat{\sigma}_i^2) \\ &\simeq q_{-i}(f_i) \sum_{z_i \in \{0,1\}} p(\mathbf{y}_i|z_i)p(z_i|f_i)\end{aligned}$$

which, by making of the definitions of the different probabilities, can be manipulated to give:

$$\hat{q}(f_i) = b_i \mathcal{N}(f_i|\mu_{-i}, \sigma_{-i}^2) + (a_i - b_i)\Phi(f_i)\mathcal{N}(f_i|\mu_{-i}, \sigma_{-i}^2) \tag{1}$$

whose moments we wish to compute for moment matching.

In order to make the notation simpler and the derivation easier to follow, we will derive the moments using a "generic" distribution $q(x)$

$$q(x) = \frac{1}{Z}\Big[b\mathcal{N}(x|\mu, \sigma^2) + (a+b)\Phi(x)\mathcal{N}(x|\mu, \sigma^2)\Big] \tag{2}$$

The normalization constant Z is given by:

$$\begin{aligned}Z &= \int_{-\infty}^{+\infty} b\mathcal{N}(x|\mu, \sigma^2) + (a-b)\Phi(x)\mathcal{N}(x|\mu, \sigma^2)dx \\ &= b + (a-b)\underbrace{\int_{-\infty}^{+\infty} \Phi(x)\mathcal{N}(x|\mu, \sigma^2)dx}_{=\Phi(\eta)} \\ &= b + (a-b)\Phi(\eta)\end{aligned} \tag{3}$$

where

$$\eta = \frac{\mu}{\sqrt{1+\sigma^2}}$$

Differentiating both sides with respect to μ gives

$$\begin{aligned}
&\frac{\partial Z}{\partial \mu} = \frac{\partial \Big[b + (a-b)\Phi(\eta) \Big]}{\partial \mu} \\
&\Leftrightarrow b \int \frac{x-\mu}{\sigma^2} \mathcal{N}(x|\mu,\sigma^2)dx + (a-b) \int \frac{x-\mu}{\sigma^2} \Phi(x)\mathcal{N}(x|\mu,\sigma^2)dx = \frac{(a-b)\mathcal{N}(\eta)}{\sqrt{1+\sigma^2}} \\
&\Leftrightarrow \frac{b}{\sigma^2} \int x\mathcal{N}(x|\mu,\sigma^2)dx - \frac{b\mu}{\sigma^2} \int \mathcal{N}(x|\mu,\sigma^2)dx \\
&\quad + \frac{(a-b)}{\sigma^2} \int x\Phi(x)\mathcal{N}(x|\mu,\sigma^2)dx - \frac{(a-b)\mu}{\sigma^2} \int \Phi(x)\mathcal{N}(x|\mu,\sigma^2)dx = \frac{(a-b)\mathcal{N}(\eta)}{\sqrt{1+\sigma^2}} \\
&\Leftrightarrow \int x \Big[b\mathcal{N}(x|\mu,\sigma^2) + (a-b)\Phi(x)\mathcal{N}(x|\mu,\sigma^2) \Big] dx \\
&\quad - \mu \underbrace{\int b\mathcal{N}(x|\mu,\sigma^2) + (a-b)\Phi(x)\mathcal{N}(x|\mu,\sigma^2)dx}_{=Z} = \frac{(a-b)\sigma^2\mathcal{N}(\eta)}{\sqrt{1+\sigma^2}}
\end{aligned}$$

where we made use of the fact that $\partial\Phi(\eta)/\partial\mu = \mathcal{N}(\eta)\partial\eta/\partial\mu$.

We recognise the first term on the left hand side to be Z times the first moment of q which we are seeking. Dividing through by Z gives

$$\begin{aligned}
\mathbb{E}_q[x] &= \mu + \frac{(a-b)\sigma^2\mathcal{N}(\eta)}{Z\sqrt{1+\sigma^2}} \\
&= \mu + \frac{(a-b)\sigma^2\mathcal{N}(\eta)}{\Big[b + (a-b)\Phi(\eta) \Big]\sqrt{1+\sigma^2}} \qquad (4)
\end{aligned}$$

Similarly, the second moment can be obtained by differentiating both sides of eq. 3 twice:

$$\begin{aligned}
&\frac{\partial^2 Z}{\partial^2 \mu} = \frac{\partial^2 \Big[b + (a-b)\Phi(\eta) \Big]}{\partial^2 \mu} \\
&\Leftrightarrow b \int \Big[\frac{x^2}{\sigma^4} - \frac{2\mu x}{\sigma^4} + \frac{\mu^2}{\sigma^4} - \frac{1}{\sigma^2} \Big] \mathcal{N}(x|\mu,\sigma^2)dx \\
&\quad + (a-b) \int \Big[\frac{x^2}{\sigma^4} - \frac{2\mu x}{\sigma^4} + \frac{\mu^2}{\sigma^4} - \frac{1}{\sigma^2} \Big] \Phi(x)\mathcal{N}(x|\mu,\sigma^2)dx = -\frac{(a-b)\eta\mathcal{N}(\eta)}{1+\sigma^2}
\end{aligned}$$

Multiplying through σ^4 and re-arranging gives

$$
\begin{aligned}
\Leftrightarrow b &\int \Big[x^2 - 2\mu x + \mu^2 - \sigma^2\Big]\mathcal{N}(x|\mu,\sigma^2)dx \\
&+ (a-b)\int \Big[x^2 - 2\mu x + \mu^2 - \sigma^2\Big]\Phi(x)\mathcal{N}(x|\mu,\sigma^2)dx = -\frac{(a-b)\sigma^4\eta\mathcal{N}(\eta)}{1+\sigma^2} \\
\Leftrightarrow b &\int x^2\mathcal{N}(x|\mu,\sigma^2)dx - 2\mu b\int x\mathcal{N}(x|\mu,\sigma^2)dx \\
&+ \mu^2 b\int \mathcal{N}(x|\mu,\sigma^2)dx - \sigma^2 b\int \mathcal{N}(x|\mu,\sigma^2)dx \\
&+ (a-b)\int x^2\Phi(x)\mathcal{N}(x|\mu,\sigma^2)dx - 2\mu(a-b)\int x\Phi(x)\mathcal{N}(x|\mu,\sigma^2)dx \\
&+ \mu^2(a-b)\int \Phi(x)\mathcal{N}(x|\mu,\sigma^2)dx - \sigma^2(a-b)\int \Phi(x)\mathcal{N}(x|\mu,\sigma^2)dx = -\frac{(a-b)\sigma^4\eta\mathcal{N}(\eta)}{1+\sigma^2} \\
\Leftrightarrow &\int x^2\Big[b\mathcal{N}(x|\mu,\sigma^2) + (a-b)\Phi(x)\mathcal{N}(x|\mu,\sigma^2)\Big]dx \\
&- 2\mu\int x\Big[b\mathcal{N}(x|\mu,\sigma^2) + (a-b)\Phi(x)\mathcal{N}(x|\mu,\sigma^2)\Big]dx \\
&+ \mu^2\underbrace{\int b\mathcal{N}(x|\mu,\sigma^2) + (a-b)\Phi(x)\mathcal{N}(x|\mu,\sigma^2)dx}_{=Z} \\
&- \sigma^2\underbrace{\int b\mathcal{N}(x|\mu,\sigma^2) + (a-b)\Phi(x)\mathcal{N}(x|\mu,\sigma^2)dx}_{=Z} = -\frac{(a-b)\sigma^4\eta\mathcal{N}(\eta)}{1+\sigma^2} \\
\Leftrightarrow &\; Z\mathbb{E}_q[x^2] - 2\mu Z\mathbb{E}_q[x] + \mu^2 Z - \sigma^2 Z = -\frac{(a-b)\sigma^4\eta\mathcal{N}(\eta)}{1+\sigma^2}
\end{aligned}
$$

Dividing through Z gives

$$
\begin{aligned}
&\Leftrightarrow \mathbb{E}_q[x^2] - 2\mu\mathbb{E}_q[x] + \mu^2 - \sigma^2 = -\frac{(a-b)\sigma^4\eta\mathcal{N}(\eta)}{Z(1+\sigma^2)} \\
&\Leftrightarrow \mathbb{E}_q[x^2] = 2\mu\mathbb{E}_q[x] - \mu^2 + \sigma^2 - \frac{(a-b)\sigma^4\eta\mathcal{N}(\eta)}{Z(1+\sigma^2)} \qquad (5)
\end{aligned}
$$

The second moment is then given by

$$
\begin{aligned}
\mathbb{E}_q[(x-\mathbb{E}_q[x])^2] &= \mathbb{E}_q[x^2]-\mathbb{E}_q[x]^2 \\
&= 2\mu\mathbb{E}_q[x]-\mu^2+\sigma^2-\frac{(a-b)\sigma^4\eta\mathcal{N}(\eta)}{Z(1+\sigma^2)}-\left(\mu+\frac{(a-b)\sigma^2\mathcal{N}(\eta)}{Z\sqrt{1+\sigma^2}}\right)^2 \\
&= 2\mu\left(\mu+\frac{(a-b)\sigma^2\mathcal{N}(\eta)}{Z\sqrt{1+\sigma^2}}\right) \\
&\quad -\mu^2+\sigma^2-\frac{(a-b)\sigma^4\eta\mathcal{N}(\eta)}{Z(1+\sigma^2)}-\left(\mu+\frac{(a-b)\sigma^2\mathcal{N}(\eta)}{Z\sqrt{1+\sigma^2}}\right)^2 \\
&= 2\mu^2+\frac{2\mu(a-b)\sigma^2\mathcal{N}(\eta)}{Z\sqrt{1+\sigma^2}} \\
&\quad -\mu^2+\sigma^2-\frac{(a-b)\sigma^4\eta\mathcal{N}(\eta)}{Z(1+\sigma^2)}-\left(\mu+\frac{(a-b)\sigma^2\mathcal{N}(\eta)}{Z\sqrt{1+\sigma^2}}\right)^2 \\
&= \mu^2+\frac{2\mu(a-b)\sigma^2\mathcal{N}(\eta)}{Z\sqrt{1+\sigma^2}}+\sigma^2-\frac{(a-b)\sigma^4\eta\mathcal{N}(\eta)}{Z(1+\sigma^2)}-\left(\mu+\frac{(a-b)\sigma^2\mathcal{N}(\eta)}{Z\sqrt{1+\sigma^2}}\right)^2
\end{aligned}
$$

Manipulating this expression further, we arrive at

$$
\mathbb{E}_q[(x-\mathbb{E}_q[x])^2]=\sigma^2-\frac{\sigma^4}{1+\sigma^2}\left(\frac{\eta\mathcal{N}(\eta)(a-b)}{b+(a-b)\Phi(\eta)}+\frac{\mathcal{N}(\eta)^2(a-b)^2}{(b+(a-b)\Phi(\eta))^2}\right) \tag{6}
$$

By making use of the moments derived above, the moments of the distribution in eq. 1 are then given by

$$
\begin{aligned}
\hat{Z}_i &= b_i+(a_i-b_i)\Phi(\eta_i) \\
\hat{\mu}_i &= \mu_{-i}+\frac{(a_i-b_i)\sigma_{-i}^2\mathcal{N}(\eta_i)}{\Big[b_i+(a_i-b_i)\Phi(\eta_i)\Big]\sqrt{1-\sigma_{-i}^2}} \\
\hat{\sigma}_i &= \sigma_{-i}^2-\frac{\sigma_{-i}^4}{1+\sigma_{-i}^2}\left(\frac{\eta_i\mathcal{N}(\eta_i)(a_i-b_i)}{b_i+(a_i-b_i)\Phi(\eta_i)}+\frac{\mathcal{N}(\eta_i)^2(a_i-b_i)^2}{(b_i+(a_i-b_i)\Phi(\eta_i))^2}\right)
\end{aligned}
$$

where

$$
\eta_i=\frac{\mu_{-i}}{\sqrt{1+\sigma_{-i}^2}}.
$$