Jannis Metrikat, Philipp Hendel Probabilistic Machine Learning

Assignment 5

Sommersemester 2024

Part 2 - Fast Bayesian Linear Regression:

Proof the multiplication of two Normal distributions:

$$\mathcal{N}\left(w_i;\mu_i,\sigma_i^2\right)\cdot\mathcal{N}\left(w_i;\phi_i^{-1}\cdot(y-\mu^T\phi+\mu_i\phi_i),\phi_i^{-2}\cdot(\beta^2+\sum_j\phi_j^2\sigma_j^2-\phi_i^2\sigma_i^2)\right)$$

$$=\mathcal{N}\left(w_i;\frac{\mu_i\cdot(\phi_i^{-2}\cdot(\beta^2+\sum_j\phi_j^2\sigma_j^2-\phi_i^2\sigma_i^2))+(\phi_i^{-1}\cdot(y-\mu^T\phi+\mu_i\phi_i))\cdot\sigma_i^2}{\sigma_i^2+(\phi_i^{-2}\cdot(\beta^2+\sum_j\phi_j^2\sigma_j^2-\phi_i^2\sigma_i^2))},\right.$$
$$\left.\frac{\sigma_i^2\cdot(\phi_i^{-2}\cdot(\beta^2+\sum_j\phi_j^2\sigma_j^2-\phi_i^2\sigma_i^2))}{\sigma_i^2+(\phi_i^{-2}\cdot(\beta^2+\sum_j\phi_j^2\sigma_j^2-\phi_i^2\sigma_i^2))}\right)$$

First, we can simplify the mean:

$$\begin{aligned}
\mu' &= \frac{\mu_i\cdot(\phi_i^{-2}\cdot(\beta^2+\sum_j\phi_j^2\sigma_j^2-\phi_i^2\sigma_i^2))+(\phi_i^{-1}\cdot(y-\mu^T\phi+\mu_i\phi_i))\cdot\sigma_i^2}{\sigma_i^2+(\phi_i^{-2}\cdot(\beta^2+\sum_j\phi_j^2\sigma_j^2-\phi_i^2\sigma_i^2))}\\
&= \frac{\mu_i\cdot(\phi_i^{-2}\cdot(\beta^2+\sum_j\phi_j^2\sigma_j^2-\phi_i^2\sigma_i^2))+(\phi_i^{-1}\cdot(y-\mu^T\phi+\mu_i\phi_i))\cdot\sigma_i^2}{\phi_i^{-2}\cdot(\phi_i^2\sigma_i^2+\left(\beta^2+\sum_j\phi_j^2\sigma_j^2-\phi_i^2\sigma_i^2\right))}\\
&= \frac{\mu_i\cdot(\beta^2+\sum_j\phi_j^2\sigma_j^2-\phi_i^2\sigma_i^2)+\phi_i\cdot(y-\mu^T\phi+\mu_i\phi_i)\cdot\sigma_i^2}{\beta^2+\sum_j\phi_j^2\sigma_j^2}\\
&= \frac{\mu_i\cdot(\beta^2+\sum_j\phi_j^2\sigma_j^2)-\mu_i\phi_i^2\sigma_i^2+\phi_i\cdot(y-\mu^T\phi)\sigma_i^2+\mu_i\phi_i^2\sigma_i^2}{\beta^2+\sum_j\phi_j^2\sigma_j^2}\\
&= \mu_i+\frac{(y-\mu^T\phi)\cdot\phi_i\sigma_i^2}{\beta^2+\sum_j\phi_j^2\sigma_j^2}\\
&= \mu_i+\frac{y-\mu^T\phi}{\phi_i}\cdot\frac{\phi_i^2\sigma_i^2}{\beta^2+\sum_j\phi_j^2\sigma_j^2}
\end{aligned}$$

Then, we simplify the variance:

$$
\begin{aligned}
(\sigma')^2 &= \frac{\sigma_i^2 \cdot (\phi_i^{-2} \cdot (\beta^2 + \sum_j \phi_j^2 \sigma_j^2 - \phi_i^2 \sigma_i^2))}{\sigma_i^2 + (\phi_i^{-2} \cdot (\beta^2 + \sum_j \phi_j^2 \sigma_j^2 - \phi_i^2 \sigma_i^2))} \\
&= \frac{\phi_i^{-2} \cdot \sigma_i^2 \cdot (\beta^2 + \sum_j \phi_j^2 \sigma_j^2 - \phi_i^2 \sigma_i^2)}{\phi_i^{-2} \cdot (\phi_i^2 \sigma_i^2 + (\beta^2 + \sum_j \phi_j^2 \sigma_j^2 - \phi_i^2 \sigma_i^2))} \\
&= \sigma_i^2 \cdot \frac{\beta^2 + \sum_j \phi_j^2 \sigma_j^2 - \phi_i^2 \sigma_i^2}{\phi_i^2 \sigma_i^2 + (\beta^2 + \sum_j \phi_j^2 \sigma_j^2 - \phi_i^2 \sigma_i^2)} \\
&= \sigma_i^2 \cdot \frac{(\beta^2 + \sum_j \phi_j^2 \sigma_j^2 - \phi_i^2 \sigma_i^2)}{\beta^2 + \sum_j \phi_j^2 \sigma_j^2} \\
&= \sigma_i^2 \cdot \left(1 - \frac{\phi_i^2 \sigma_i^2}{\beta^2 + \sum_j \phi_j^2 \sigma_j^2 y}\right)
\end{aligned}
$$

Ultimately, we end up with the assertion:

$$
\mathcal{N}\left(w_i; \mu', (\sigma')^2\right) = \mathcal{N}\left(w_i; \mu_i + \frac{y - \mu^T \phi}{\phi_i} \cdot \frac{\phi_i^2 \sigma_i^2}{\beta^2 + \sum_j \phi_j^2 \sigma_j^2}, \sigma_i^2 \cdot \left(1 - \frac{\phi_i^2 \sigma_i^2}{\beta^2 + \sum_j \phi_j^2 \sigma_j^2}\right)\right)
$$

Part 3 - Sherman-Morrison Formula:

Verify that the inverse is correct, i.e., that the following identity holds:

$$\left(\mathbf{A}+\mathbf{u}\mathbf{v}^{\mathbf{T}}\right)\left(\mathbf{A}^{-1}-\frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^{\mathbf{T}}\mathbf{A}^{-1}}{1+\mathbf{v}^{\mathbf{T}}\mathbf{A}^{-1}\mathbf{u}}\right)=\mathbf{I}$$

Proof:

$$\begin{aligned}
&\left(A+uv^T\right)\left(A^{-1}-\frac{A^{-1}uv^TA^{-1}}{1+v^TA^{-1}u}\right)\\
&=AA^{-1}+uv^TA^{-1}-\frac{(A+uv^T)\cdot(A^{-1}uv^TA^{-1})}{1+v^TA^{-1}u}\\
&=I+uv^TA^{-1}-\frac{AA^{-1}uv^TA^{-1}+uv^TA^{-1}uv^TA^{-1}}{1+v^TA^{-1}u}\\
&=I+uv^TA^{-1}-\frac{Iuv^TA^{-1}+(v^TA^{-1}u)\cdot uv^TA^{-1}}{1+v^TA^{-1}u}\\
&=I+uv^TA^{-1}-\frac{(1+v^TA^{-1}u)\cdot uv^TA^{-1}}{1+v^TA^{-1}u}\\
&=I+uv^TA^{-1}-uv^TA^{-1}\\
&=I
\end{aligned}$$