1 Nonlinear regression/kernel regression exercises

1.1 Definitions

Let

$$
\begin{aligned}
X &= (x_1, \ldots, x_N) \\
T &= (t_1, \ldots, t_N) \\
\phi(x) &= (\phi_1(x), \ldots, \phi_M(x)) \\
\Phi(X) &= \begin{pmatrix} \phi_1(x_1) & \cdots & \phi_M(x_1) \\ \cdots & \cdots & \cdots \\ \phi_1(x_N) & \cdots & \phi_M(x_N) \end{pmatrix}
\end{aligned} \tag{1}
$$

1.2 Problem 1: Isotropic nonlinear regression

Let $w \sim \mathcal{N}\left(0, \alpha^{-1} I\right)$, $t|w, x \sim \mathcal{N}\left(\langle \phi(x), w \rangle, \beta^{-1}\right)$.

(a) Write out (log) of $P(T, w|X)$
(b) Using completing the square of other means, derive an expression for $P(w|T, X) = \mathcal{N}\left(\mu_{w|T}, \Sigma_{w|T}\right)$
(c) Continue completing the square with (b), and marginalize out w to come up with an expression for $P(T|X)$. You can normalization factors for now (so just focus on the exponential term).
(d) In (c), the expression for Σ_T will be very complicated. Use Woodbury's identity to prove that $\Sigma_T = \Phi \alpha^{-1} I \Phi^\intercal + \beta^{-1} I$

1.3 Problem 2: More general nonlinear regression

More generally, let $w \sim \mathcal{N}(\mu_w, \Sigma_w)$, $T|w, X \sim \mathcal{N}\left(\Phi w, \Sigma_{T|X}\right)$.

(a) Write out (log) of $P(T, w|X)$
(b) Using completing the square of other means, derive an expression for $P(w|T, X) = \mathcal{N}\left(\mu_{w|T}, \Sigma_{w|T}\right)$
(c) Continue completing the square with (b), and marginalize out w to come up with an expression for $P(T|X)$. You can normalization factors for now (so just focus on the exponential term).
(d) In (c), the expression for Σ_T will be very complicated. Use Woodbury's identity to prove that $\Sigma_T = \Phi \Sigma_w \Phi^\intercal + \Sigma_{T|X}$

1.4 Problem 3: GP regression

Suppose $y \sim GP\left(\mu_y(\cdot), k(\cdot, \cdot)\right)$
Let $Y = (y(x_1), \ldots, y(x_N))$, $\mu_y = (\mu_y(x_1), \ldots, \mu_y(x_N))$,

$$
K = \begin{pmatrix} k(x_1, x_1) & \cdots & k(x_1, x_N) \\ \cdots & \cdots & \cdots \\ k(x_N, x_1) & \cdots & k(x_N, x_N) \end{pmatrix} \tag{2}
$$

By definition, $Y \sim \mathcal{N}(\mu_y, K)$ and $T|Y, X \sim \mathcal{N}\left(Y, \Sigma_{T|Y}\right)$

(a) Write out (log) of $P(T, Y|X)$
(b) Using completing the square of other means, derive an expression for $P(Y|T, X)$
(c) Continue completing the square with (b), and marginalize out Y to come up with an expression for $P(T|X)$. You can normalization factors for now (so just focus on the exponential term).
(d) In (c), the expression for Σ_T will be very complicated. Use Woodbury's identity to prove that $\Sigma_T = K + \Sigma_{T|Y}$

1.5 Problem 4: Multivariate Gaussians

Note our random variables are unrelated to the previous problems.

$$
\begin{aligned}
X &= \begin{pmatrix} X_A \\ X_B \end{pmatrix} \\
\mu &= \begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix} \\
\Sigma &= \begin{pmatrix} \Sigma_{AA} & \Sigma_{AB} \\ \Sigma_{BA} & \Sigma_{BB} \end{pmatrix} \\
\Lambda &:= \Sigma^{-1} \\
\Lambda &= \begin{pmatrix} \Lambda_{AA} & \Lambda_{AB} \\ \Lambda_{BA} & \Lambda_{BB} i \end{pmatrix} \\
X &\sim \mathcal{N}(\mu, \Sigma)
\end{aligned} \tag{3}
$$

(a) Write out the exponential term of $P(X)$ (ignore the normalizations constants).
(b) Complete the square for the x_A terms.
(c) Marginalize out x_A using (b) to get an expression for the exponential term of $P(x_B|x_A)$ in terms of $x_B - \mu_B$ and $(x_B - \mu_B)(x_b - \mu_B)^\intercal$
(d) The quadratic coefficient for $x_B - \mu_B$ is written in terms of precision matrices. Use the Schur complement to convert it into covariance matrices.
(e) Use Woodbury's identity to simplify the expressions for $\mu_{x_B|x_A}, \Sigma_{x_b|x_A}$ ($x_B|x_A \sim \mathcal{N}\left(\mu_{x_B|x_A}, \Sigma_{x_b|x_A}\right)$).

1.6 Problem 5: Kernels