9A

Machine Learning II

ID: 5684926 Tristan Scheidemann

8A-2

Auxiliary calculation

Let $a, b \in \mathbb{R}^n$ be elements of an inner product space over the field $\mathbb{R}$. If the inner product is a valid bilinear form then we have for scalars c:

$$\begin{aligned} c \cdot \langle a, b \rangle^k &= \prod_{i=1}^{k} \sqrt[k]{c} \langle a, b \rangle \\ &= \langle \sqrt[k]{c} a, b \rangle^k . \end{aligned}$$

Furthermore, the multinomial theorem asserts that

$$\left[\sum_{i=1}^{n} x_i \right]^k = \sum_{m_1 + \cdots + m_n = k}^{k} \binom{k}{m_1, \dots, m_n} \prod_{t=1}^{n} x_t^{m_t},$$

where $\binom{k}{m_1, \dots, m_n}$ represents the binomial coefficient.

Starting off with McLaurin:

$$
\begin{aligned}
k(\mathbf{x},\boldsymbol{\mu}) &= e^{-\frac{1}{2\sigma^2}(\mathbf{x}-\boldsymbol{\mu})^T(\mathbf{x}-\boldsymbol{\mu})} \\
&= e^{-\frac{1}{2\sigma^2}\langle\mathbf{x}-\boldsymbol{\mu},\mathbf{x}-\boldsymbol{\mu}\rangle} \\
&= \underbrace{e^{-\frac{1}{2\sigma^2}\langle\mathbf{x},\mathbf{x}\rangle}}_{a(\mathbf{x})} e^{-\frac{1}{2\sigma^2}[-2\langle\mathbf{x},\boldsymbol{\mu}\rangle]} \underbrace{e^{-\frac{1}{2\sigma^2}[-2\langle\boldsymbol{\mu},\boldsymbol{\mu}\rangle]}}_{a(\boldsymbol{\mu})} \\
&= a(\mathbf{x})a(\boldsymbol{\mu})\left[\sum_{k=0}^{\infty}\left(\frac{1}{\sigma^2}\right)^k \frac{(\langle\mathbf{x},\boldsymbol{\mu}\rangle)^k}{k!}\right] \\
&= \sum_{k=0}^{\infty}\left(\langle\underbrace{\frac{\sqrt[k]{a(\mathbf{x})}}{\sqrt[2k]{k!}\,\sigma}}_{b(\mathbf{x})}\mathbf{x},\underbrace{\frac{\sqrt[k]{a(\boldsymbol{\mu})}}{\sqrt[2k]{k!}\,\sigma}}_{b(\boldsymbol{\mu})}\boldsymbol{\mu}\rangle\right)^k \\
&= \sum_{k=0}^{\infty}(\langle b(\mathbf{x})\mathbf{x}, b(\boldsymbol{\mu})\boldsymbol{\mu}\rangle)^k \\
&= \sum_{k=0}^{\infty}\left(\sum_{i=0}^{n} b(\mathbf{x})x_i \cdot b(\boldsymbol{\mu})\mu_i\right)^k \\
&= \sum_{k=0}^{\infty}\left(\sum_{m_1+\cdots+m_n=k}^{k}\binom{k}{m_1,\ldots,m_n}\prod_{t=1}^{n}[b(\mathbf{x})x_t \cdot b(\boldsymbol{\mu})\mu_{\mathrm{t}}]^{m_t}\right) \\
&= \sum_{k=0}^{\infty}\left(\sum_{|\alpha|=k}\phi_{k,\alpha}(\mathbf{x})\phi_{k,\alpha}(\boldsymbol{\mu})\right),
\end{aligned}
$$

utilizing, that

$$
\phi_{k,\alpha}(\mathbf{x}) = \sqrt{\binom{k}{m_1,\ldots,m_n}}\prod_{t=1}^{n}[b(\mathbf{x})x_t]^{m_t}
$$

$$
\phi_{k,\alpha}(\boldsymbol{\mu}) = \sqrt{\binom{k}{m_1,\ldots,m_n}}\prod_{t=1}^{n}[b(\boldsymbol{\mu})\mu_{\mathrm{t}}]^{m_t}.
$$

This can be decomposed into a dot product again, which leads to:

$$
\sum_{k=0}^{\infty}\left(\sum_{|\alpha|=k}\phi_{k,\alpha}(\mathbf{x})\phi_{k,\alpha}(\boldsymbol{\mu})\right) \quad = \sum_{k=0}^{\infty}\boldsymbol{\Phi}_k(\mathbf{x})^{\mathrm{T}}\boldsymbol{\Phi}_k(\boldsymbol{\mu}),
$$

where

$$
\boldsymbol{\Phi}_k(\mathbf{x})_i = \phi_{\mathrm{k,f(i)}}(\mathbf{x}),
$$

with index set f: $\mathbb{N} \to \{\alpha : m_1 + \cdots + m_n = k, m_j \in \mathbb{N}\}$.