(7.18) を考える。

そもそも、ガウス過程の仮定 (と、その平均が 0 という仮定 (P.189 の脚注)) より、$X, Y \in \mathcal{R}^N$ として、

$$p(Y|X) = \mathcal{N}(Y|0, K_\theta) = \frac{1}{\sqrt{(2\pi)^N |K_\theta|}} exp\left\{-\frac{1}{2}Y^T {K_\theta}^{-1} Y\right\} \tag{1}$$

なお、K_θ は $k_\theta(x, x')$ を要素とする行列になっている。よって、(7.17) のように、

$$ln\, p(Y|X) = -\frac{1}{2}Y^T {K_\theta}^{-1} Y - \frac{1}{2} ln\, |K_\theta| - \frac{N}{2} ln\, 2\pi \tag{2}$$

これを微分するが、PRML の (C.21),(C.22)(同様の話は、参考文献 [100](https://www2.imm.dtu.dk/pubdb/edoc/imm3274.pc
にある、(59) と,(46) を計算して、(C.22) 相当を求めるところにある。) を参考に、それぞれ、θ_i で微分する。
K_θ が対称行列で、その逆行列も対称行列になるので、

$$\begin{aligned}
\frac{\partial}{\partial \theta_i} ln\, p(Y|X, \theta) &= \frac{1}{2}Y^T {K_\theta}^{-1} \frac{\partial K_\theta}{\partial \theta_i} {K_\theta}^{-1} Y - \frac{1}{2} Tr({K_\theta}^{-1} \frac{\partial K_\theta}{\partial \theta_i}) = \frac{1}{2}\alpha^T \frac{\partial K_\theta}{\partial \theta_i} \alpha - \frac{1}{2} Tr({K_\theta}^{-1} \frac{\partial K_\theta}{\partial \theta_i}) \\
&= \frac{1}{2} Tr(\alpha^T \frac{\partial K_\theta}{\partial \theta_i} \alpha) - \frac{1}{2} Tr({K_\theta}^{-1} \frac{\partial K_\theta}{\partial \theta_i}) = \frac{1}{2} Tr(\alpha\alpha^T \frac{\partial K_\theta}{\partial \theta_i}) - \frac{1}{2} Tr({K_\theta}^{-1} \frac{\partial K_\theta}{\partial \theta_i}) \\
&= \frac{1}{2} Tr(\alpha\alpha^T \frac{\partial K_\theta}{\partial \theta_i} - {K_\theta}^{-1} \frac{\partial K_\theta}{\partial \theta_i}) = \frac{1}{2} Tr((\alpha\alpha^T - {K_\theta}^{-1}) \frac{\partial K_\theta}{\partial \theta_i})
\end{aligned} \tag{3}$$

3 個目の等号はスカラの Tr をとっても変わらないことを、4 個目の等号は PRML の (C.8) を使った。