Proof of Standard Deviation formula, in Kernel-UCB algorithm, using Woodbury Matrix Identity

Anubhav Singh

January 30, 2019

Proof

According to the Woodbury Matrix Identity:

$$(\mathbf{A} + \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} + \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} \tag{1}$$

Now, assign the following to matrices A, B, C and D

$$\mathbf{A} = \gamma\mathbf{I_d},\ \mathbf{B} = \mathbf{\Phi^T_{dxt}},\ \mathbf{C} = \mathbf{\Phi_{txd}},\ \mathbf{D} = \mathbf{I_t} \tag{2}$$

Using equation (1),

$$\begin{aligned}
(\gamma\mathbf{I_d} + \mathbf{\Phi^T\Phi})^{-1} &= (\gamma\mathbf{I_d} + \mathbf{\Phi^T I_t^{-1}\Phi})^{-1} \\
&= \gamma^{-1}\mathbf{I_d^{-1}} - \gamma^{-1}\mathbf{I_d^{-1}\Phi^T}(\mathbf{I_t} + \mathbf{\Phi}(\gamma^{-1}\mathbf{I_d^{-1}})\mathbf{\Phi^T})^{-1}\mathbf{\Phi}\gamma^{-1}\mathbf{I_d^{-1}} \\
&= \gamma^{-1}\mathbf{I_d} - \gamma^{-1}\mathbf{I_d\Phi^T}(\mathbf{I_t} + \mathbf{\Phi}(\gamma^{-1}\mathbf{I_d})\mathbf{\Phi^T})^{-1}\mathbf{\Phi}\gamma^{-1}\mathbf{I_d} \\
&= \gamma^{-1}\mathbf{I_d} - \gamma^{-2}\mathbf{\Phi^T}\gamma(\gamma\mathbf{I_t} + \mathbf{\Phi\Phi^T})^{-1}\mathbf{\Phi} \\
&= \gamma^{-1}(\mathbf{I_d} - \mathbf{\Phi^T}(\gamma\mathbf{I_t} + \mathbf{\Phi\Phi^T})^{-1}\mathbf{\Phi}) \\
&= \gamma^{-1}(\mathbf{I_d} - \mathbf{\Phi^T}(\gamma\mathbf{I_t} + \mathbf{K})^{-1}\mathbf{\Phi})
\end{aligned} \tag{3}$$

Applying the results obtained in (3) to the Standard Deviation Formula, we get the formula used in Algorithm-1

$$\begin{aligned}
\sigma_{\mathbf{n,t}} &= [\phi^{\mathbf{T}}_{\mathbf{x_{n,t}}}(\gamma\mathbf{I} + \mathbf{\Phi^T_{t-1}\Phi_{t-1}})^{-1}_{\mathbf{t-1}}\phi_{\mathbf{x_{n,t}}}]^{1/2} \\
&= [\gamma^{-1}\phi^{\mathbf{T}}_{\mathbf{x_{n,t}}}(\mathbf{I} - \mathbf{\Phi^T_{t-1}}(\gamma\mathbf{I} + \mathbf{K_{t-1}})^{-1}\mathbf{\Phi_{t-1}})\phi_{\mathbf{x_{n,t}}}]^{1/2} \\
&= [\gamma^{-1}\{(\phi^{\mathbf{T}}_{\mathbf{x_{n,t}}}\phi_{\mathbf{x_{n,t}}}) - (\phi^{\mathbf{T}}_{\mathbf{x_{n,t}}}\mathbf{\Phi^T_{t-1}})(\gamma\mathbf{I} + \mathbf{K_{t-1}})^{-1}(\mathbf{\Phi_{t-1}}\phi_{\mathbf{x_{n,t}}})\}]^{1/2} \\
&= \gamma^{-1/2}[\mathbf{k}(\mathbf{x_{n,t}}, \mathbf{x_{n,t}}) - \mathbf{k^T_{x_{n,t},t-1}}(\gamma\mathbf{I} + \mathbf{K_{t-1}})^{-1}\mathbf{k_{x_{n,t},t-1}}]^{1/2}
\end{aligned} \tag{4}$$