Notes by: Anubhav Singh

Corrected KernelUCB algorithm

October 8, 2018

We chose RBF kernel for our tests.

Using the following definitions (clears up inconsistencies):

K(t) is Kernel Matrix at time t.Gauss

- $k_{x^\star,t} = [k(x^\star, x_1), \dots k(x^\star, x_t)]^\intercal$
- $y_t = [r_1, \dots, r_t]^\intercal$
- $\Phi_t = [\phi(x_1)^\intercal, \dots, \phi(x_t)^\intercal]^\intercal$

-Here phi(x) is the vector transformation o

- $K_t = \Phi_t^\intercal \Phi_t$ Wrong! Swap the matrices
- $\hat{\sigma}_{a,t+1} = \left[\phi(x_{a,t+1})^\intercal (\Phi_t^\intercal \Phi_t + \gamma I)^{-1} \phi(x_{a,t+1})\right]^{\frac{1}{2}}$
- $\hat{\mu}_{a,t+1} = k_{x_{a,t+1},t}^\intercal (K_t + \gamma I)^{-1} y_t$

In mu (mean) formula, we are assuming that phi_upper is square

---

Algorithm 1 KernelUCB with online updates

Input: N the number of actions, T the number of pulls, γ, η regularization and exploration parameters, $k(\cdot,\cdot)$ kernel function

```
for t ∈ 1, ..., T do
    Receive contexts {x_{1,t}, ..., x_{N,t}}
    if t = 1 then
        u_t ← [1, 0, ..., 0]^T
    else
        for n ∈ {1, ... N} do
            σ_{n,t} ← [k(x_{n,t}, x_{n,t}) − k^T_{x_{n,t},t−1} K^{−1}_{t−1} k_{x_{n,t},t−1}]^{1/2}
            u_{n,t} ← k^T_{x_{n,t},t−1} K^{−1}_{t−1} y_{t−1} + (η/√γ) σ_{n,t}
        end for
    end if
    Choose action a_t ← arg max u_t and receive reward r_t
    Store context for action a_t: x_t ← x_{a,t}
    Update reward history: y_t ← [r_1, ..., r_t]^T
    if t = 1 then                                   ▷ initialise kernel matrix inverse
        K_t^{−1} ← (k(x_t, x_t) + γ)^{−1}
    else                                            ▷ online update of kernel matrix inverse
        b ← k_{x_t,t−1}
        K_22 ← (k(x_t, x_t) + γ − b^T K^{−1}_{t−1} b)^{−1}
        K_11 ← K^{−1}_{t−1} + K_22 K^{−1}_{t−1} b b^T K^{−1}_{t−1}
        K_12 ← −K_22 K^{−1}_{t−1} b
        K_21 ← −K_22 b^T K^{−1}_{t−1}
        K_t^{−1} ← [K_11  K_12; K_21  K_22]
    end if
end for
```

These can be proved to be equal using Wo

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

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