Advanced Topics in Machine Learning - Assignment 6

Christoffer Thrysøe - dfv107

October 24, 2017

1 Feature space dimensionality

1.1

Using a feature map, we define our decision function as followed:

$$f(x) = \langle \Phi(x), w \rangle + b \tag{1}$$

thus, we describe a hyperplane using the parameters (w, b), where w is the weights of the hyperplane and b is the bias parameter. The hyperplane must exist in the same space as the feature space, thus w must have the same number of weights as number of features in $\Phi(x)$. The kernel is defined as followed: $k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$ in the case of the linear transform, we have that $\Phi(x) = x$, therefore the feature transform is defined as followed: $\Phi : \mathbb{R}^d \to \mathbb{R}^d$. Thus in order to define a solution for the SVM using a linear kernel, we require the same number of parameters as the dimension of the feature space, for defining w, plus one parameter for the bias b. Combined. we need $d + 1$, where d is the dimension of the feature space. Because the feature space, using the linear kernel, will also be d we will need $d + 1$ parameters to define the solution, when using the linear kernel.

1.2

The following is the primal and dual solution to the SVM optimization problem for Linear hard margin. I will use these examples to show how the sufficient number of parameters, needed to solve the optimization problem, differs: The primal solution is given by:

$$\text{Minimize}_{w,b} \quad \frac{1}{2}\langle w, w \rangle \tag{2}$$

$$\text{subject to} \quad y_i(\langle w, x_i \rangle + b) \geq 1, i = 1, ..., N \tag{3}$$

The dual form is given by:

$$\text{Maximize}_{\alpha} \quad \sum_{i=0}^{N} \alpha_i - \frac{1}{2}\sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle \tag{4}$$

$$\text{subject to} \quad \sum_{i=1}^{N} \alpha_i y_i = 0 \tag{5}$$

$$\alpha_i \geq 0, i = 1, ..., N \tag{6}$$

For the primal problem, we need to learn $d+1$ parameters, and for the dual problem we need to learn $N+1$ parameters (N α and one bias b). As we saw from the previous assignment, we need as many parameters as the dimension of the feature space d' plus one for the bias: $d'+1$. If we have that $d' > N$, then we can instead use the dual problem and solve for α rather than w, only requiring $N+1$ number of parameters.

1.3

We wish to show that the $n \times n$ gram matrix K is positive definite i.e. not strictly positive. That is we wish to prove that for an arbitrary vector $z = (z_1, \dots, z_n)^T \in \mathbb{R}^n$, we have $z^T K z \geq 0$

$$\begin{aligned} z^T K z &= \sum_{i=1}^{n}\sum_{j=1}^{n} k(x_j, x_i) z_i z_j && (7) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \langle \Phi(x_j), \Phi(x_i) \rangle z_i z_j && (8) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \langle z_j \Phi(x_j), z_i \Phi(x_i) \rangle && (9) \\ &= \left\langle \sum_{j=1}^{n} z_j \Phi(x_j), \sum_{i=1}^{n} z_i \Phi(x_i) \right\rangle && (10) \\ &= \left\| \sum_{i=1}^{n} z_i \Phi(x_i) \right\|^2 \geq 0 && (11) \end{aligned}$$

The above holds for any arbitrary vector z, we now wish to show that it is not strictly greater than zero. That is, if $z \neq 0$, $z^T K z$ will be equal to zero. Because we have that $d' < n$, we know that the matrix K will not have full rank. Therefore a subset of the vectors $\Phi(x)$ will be linear dependent. For linear dependent vectors $\vec{v_1}, \vec{v_2}, ..., \vec{v_n}$ and non-zero scalars $a_1, a_2, ..., a_n$ we can express the following, due to the fact that linear dependent vectors can be written as a combination of each other:

$$a_1 \vec{v_1} + a_2 \vec{v_2} + ... + a_n \vec{v_n} = \vec{0} \qquad (12)$$

Thus in our case, it means that we can find a vector $z_j \neq 0$ s.t $z\Phi(x) = 0$, in which we can also have that $z^T K z = 0$. Therefore it can be concluded that the gram matrix K is semi positive definite.

1.4

To show an example of a Gaussian kernel I will quickly show the feature space of the RBF kernel, which is defined as:

$$K(x, x') = exp(-\gamma ||x - x'||^2) \qquad (13)$$

To see the transformed feature space, we can pick $\gamma = 1$ and write out the transformation:

$$\begin{aligned} K(x, x') &= exp(||x - x'||^2) && (14) \\ &= exp(-(x)^2) \cdot exp(2xx') \cdot exp(-(x')^2) && (15) \\ &= exp(-(x)^2) \cdot \sum_{k=0}^{\infty} \frac{2^k (x)^k (x')^k}{k!} \cdot exp(-(x')^2) && (16) \end{aligned}$$

Thus the non linear transform can be written as followed:

$$\Phi(x) = exp(-x^2) \cdot \left(1, \sqrt{\frac{2^1}{1!}} x, \sqrt{\frac{2^2}{2!}} x^2, \dots \right) \qquad (17)$$

We have that the transformation Φ is an infinite dimensional feature transformation. We would now have to search for an optimal SVM in the infinite dimensional space, which is not possible, solving the primal problem. Here we can use the representer theorem to re-formulate our problem as regularized risk minimization, which allows us to search for a solution in the infinite space, but using a problem, reduced to finite-dimensional. The infinite to finite dimensional solution comes from the usage of the regularization term. The representer theorem states that a solution of the regularized loss admits a representation of the form:

$$f(x) = \sum_{i=1}^{N} \alpha_i k(x_i, x) \tag{18}$$

Thus the number of parameters has been reduced from infinite to the number of training points N.

2 The doubling trick

2.1

We wish to prove that for any $T = 2^m - 1$, the overall expected regret, given time periods $1, \dots, T$, is bounded by:

$$\mathbb{E}[R_T] \leq \frac{1}{\sqrt{2} - 1} \sqrt{\frac{1}{2} T \ln N} \tag{19}$$

We have given that for a single period, that is $(2^m, \dots, 2^{m+1} - 1)$ the expected regret is bounded by:

$$\mathbb{E}[R_T] \leq \sqrt{\frac{1}{2} 2^m \ln N} \tag{20}$$

thus we wish to prove the following:

$$\sum_{i=0}^{m-1} \sqrt{\frac{1}{2} 2^i \ln N} \leq \frac{1}{\sqrt{2} - 1} \sqrt{\frac{1}{2} T \ln N} \tag{21}$$

First we remove $\sqrt{\frac{1}{2} \ln N}$ on both sides of (21) and replace T with $2^m - 1$, resulting in the following:

$$\sum_{i=0}^{m-1} \sqrt{2^i} \leq \frac{1}{\sqrt{2} - 1} \sqrt{2^m - 1} \tag{22}$$

re-writing the left hand side of (22) yields:

$$\sum_{i=0}^{m-1} \sqrt{2^i} = \sum_{i=0}^{m-1} 2^{\frac{1}{2} i} \tag{23}$$

This forms a geometric series, which we can write out as followed:

$$\sum_{i=0}^{m-1} 2^{\frac{1}{2} i} = 2^{\frac{1}{2} 0} + 2^{\frac{1}{2} 1} + \cdots + 2^{\frac{1}{2} m - 1} = S \tag{24}$$

where we denote the sequence S. A geometric sequence consists of a sequence $\{a, ar, a^2, ar^3, \dots\}$. In our sequence S, $r = 2^{\frac{1}{2}}$. We can derive the sequence S by first multiplying r and then subtracting the sequence S again, note that when we multiply the sequence S with r we get the following sequence:

$$S \cdot r = 2^{\frac{1}{2} 1} + 2^{\frac{1}{2} 2} \cdots + 2^{\frac{1}{2} m} \tag{25}$$

thus $S \cdot r - S$, gives us the following:

$$S \cdot r - S = 2^{\frac{1}{2}m} - 2^{\frac{1}{2}0} \Rightarrow S = \frac{2^{\frac{1}{2}m} - 1}{2^{\frac{1}{2}} - 1} \tag{26}$$

now if we plug this value of the sequence back into inequality (22) we get the following:

$$\frac{2^{\frac{1}{2}m} - 1}{2^{\frac{1}{2}} - 1} \leq \frac{1}{\sqrt{2} - 1}\sqrt{2^m - 1} \Rightarrow \tag{27}$$

$$0 \leq \sqrt{2^m} - 1 \leq \sqrt{2^m - 1} \Rightarrow \tag{28}$$

$$2^m - 2\sqrt{2^m} + 1 \leq 2^m - 1 \Rightarrow \tag{29}$$

$$0 \leq 2\sqrt{2^m} - 2 \Rightarrow \tag{30}$$

$$0 \leq \sqrt{2^m} - 1 \tag{31}$$

which holds, because $m = \{0, 1, 2.....\}$. Thus proving that the combined expected regret over all the time periods is bounded by:

$$\mathbb{E}[R_T] \leq \frac{1}{\sqrt{2} - 1}\sqrt{\frac{1}{2}T \ln N} \tag{32}$$

2.2

Using the same forecasting strategy, we wish to prove that for any arbitrary time T, the expected regret bounded by:

$$\mathbb{E}[R_T] \leq \frac{\sqrt{2}}{\sqrt{2} - 1}\sqrt{\frac{1}{2}T \ln N} \tag{33}$$

I will prove this by stating that our arbitrary T is located within two time periods, and then bound the lower time period + some constant C, with the upper time period, using that we have just derived a bound on the combined expected regret for a given time period. Thus we have the following time periods, where T is included:

$$(1), (2, 3), (4, ..., 7), ..., (2^{m-1}, ..., 2^m - 1), (2^m, ...[T = 2^m + C]..., 2^{m+1} - 1) \tag{34}$$

We have the following in (35): The first term is the bound on the expected regret of the time period before the arbitrary T, which we obtained from the previous assignment, the middle term is the regret at the arbitrary T and the last term is the regret at the end of the time period, where T is included:

$$\frac{1}{\sqrt{2} - 1}\sqrt{\frac{1}{2}(2^m - 1) \ln N} \leq \frac{1}{\sqrt{2} - 1}\sqrt{\frac{1}{2}(2^m + C) \ln N} \leq \frac{1}{\sqrt{2} - 1}\sqrt{\frac{1}{2}(2^{m+1} - 1) \ln N} \tag{35}$$

again for each term, we remove $\sqrt{\frac{1}{2} \ln N}$. We wish to provide a bound on the middle term of (35), which we can verify. Therefore we will try to rewrite the last term of (35) to show that the inequality holds and provide an upper bound for the regret at time T. We multiply the last term by $\sqrt{2}$, which is allowed as the inequality will still hold.

$$\frac{1}{\sqrt{2} - 1}\sqrt{2^{m+1} - 1} \leq \frac{\sqrt{2}}{\sqrt{2} - 1}\sqrt{2^m + C} = \frac{1}{\sqrt{2} - 1}\sqrt{2^{m+1} + 2C} \tag{36}$$

Note that the middle term in (36) is the bound we want to prove. We now have:

$$\frac{1}{\sqrt{2} - 1}\sqrt{2^m + C} \leq \frac{1}{\sqrt{2} - 1}\sqrt{2^{m+1} + 2C} \Rightarrow \tag{37}$$

$$\sqrt{2^m + C} \leq \sqrt{2^{m+1} + 2C} \tag{38}$$

which clearly holds because C is a positive integer. Thus we have proven that for and arbitrary $T = 2^m + C$, we have the following bound on the expected regret:

$$\mathbb{E}[R_T] \leq \frac{\sqrt{2}}{\sqrt{2}-1}\sqrt{\frac{1}{2}T \ln N} \tag{39}$$

3 Empirical evaluation of UCB1 and EXP3 algorithms

We wish to compare the performance of the two algorithms UCB1 and EXP3. The algorithms will be compared by taking time horizon $T = 10000$ and number of suboptimal arms: $1, 3, 7, 15$, and one optimal arm. This experiment is done three times, for the following settings of μ for the suboptimal arms:

$$\mu = \mu^* - \frac{1}{4}, \mu = \mu^* - \frac{1}{8}, \mu = \mu^* - \frac{1}{16}$$

where $\mu^* = 0.5$. Each experiment is done with K arms, that is $K - 1$ suboptimal arms and 1 optimal arm. For each arm, a sequence of rewards is generated, where μ defines the probability of returning a reward, that is, the probability of a reward at time t is given by:

$$\mathbb{P}\{r_t = 1\} = \mu \tag{40}$$

$$\mathbb{P}\{r_t = 0\} = 1 - \mu \tag{41}$$

The UCB1 algorithm is shown below:

Algorithm 1 UCB1

```
    Initialization: Play each action once
for $t = K + 1, K + 2, ...$ do
    Play $A_t = \arg\max_a \hat{\mu}_{t-1}(a) + \sqrt{\frac{3 \ln t}{2N_{t-1}(a)}}$
end for
```

The program starts by playing each action once, once an arm a is played $N_t(a)$ is incremented, the reward of the sequence for that specific arm is then fetched and the reward is logged. The potential rewards of the other arms is unknown as we are in bandit setting and therefore only know the reward of the arm being pulled. $\hat{\mu}(a)$ is calculated by taking the number of rewards returned from a, i.e. the number of times we pulled arm a and got a reward, divided by the number of times the arm has been pulled. The upper confidence bound is then taken for each arm and the arm with the highest upper confidence bound is played. In case of a tie, the played arm is picked randomly amongst the tied arms. The EXP3 algorithm is defined as followed:

Algorithm 2 EXP3

```
    Input: Learning rates: $\eta_1 \geq \eta_2 \geq \dots > 0$
    $\forall a : L_0(a) = 0$
for $t = 1, 2, ...$ do
    $\forall a : p_t(a) = \frac{e^{-\eta_t L_{t-1}(a)}}{\sum_{a'} e^{-\eta_t L_{t-1}(a')}}$
    Sample $A_t$ according to $p_t$ and play it
    Observe and suffer $l_t^{A_t}$
    Set $\tilde{l}_t^a = \begin{cases} \frac{l_t^a}{p_t(a)} & \text{if } A_t = a \\ 0 & \text{Otherwise} \end{cases}$
    $\forall a : L_t(a) = L_{t-1} + \tilde{l}_t^a$
end for
```

Again, we only know the reward of the arm being played. If a pulled arm has zero reward, we count it as a loss, and therefore update the probability distribution p for that single arm only. The predicted arm is sampled according to p and played.
The performance of the algorithms is given by the pseudo regret, which is defined as:

$$\bar{R}_T = \sum_{s=1}^{t} \mathbb{E}\left[N_t(a)\right] \Delta(a) \tag{42}$$

That is, each time we play a suboptimal arm we accumulate the gap. The pseudo regret is taken for each iteration, and plotted as a function of T. The pseudo regret is plotted for each experiment, i.e. for each setting of μ. The below plots show the result of each experiment, where each experiment is run 10 times , and the average is taken plus the standard deviation of the runs. For each algorithm, the result is shown for $K = \{2, 4, 8, 16\}$, where the pseudo regret is shown as a line, and the pseudo regret + the standard deviation over the 10 runs is shown as a scribbled line (The color of the scribbled line identifies, what pseudo regret it belongs to).

$\mu^* - \frac{1}{4}$

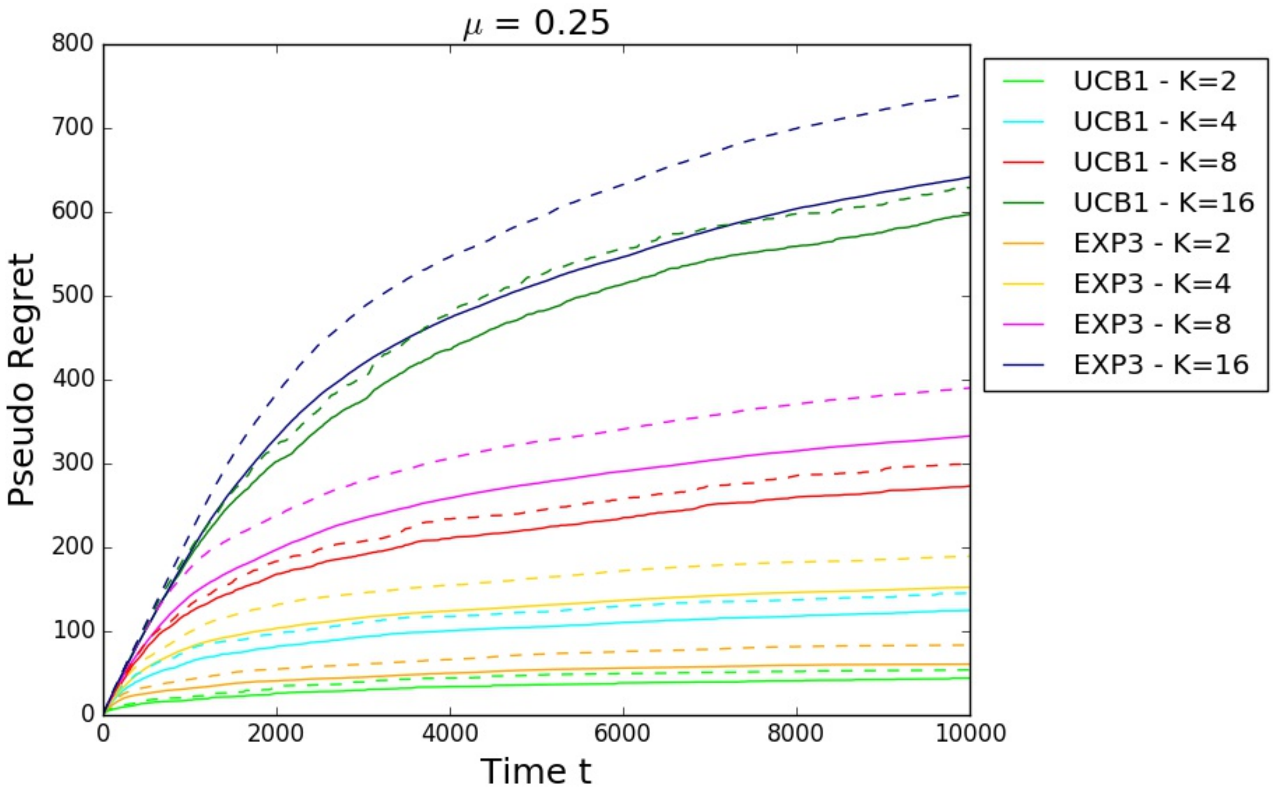

Figure 1: Average pseudo regret and average pseudo regret + standard deviation (scribbled line) over 10 runs with $\mu = 0.25$

$\mu^* - \frac{1}{8}$

Figure 1 shows the experiment for $\mu = 0.25$. As evident, the UCB1 and EXP3 algorithm has a very similar performance, however the UCB1 seems to outperform EXP3 slightly on each setting of K. It is also clear that the regret increases with the number of arms. The regret increases rapidly for the first few iterations and converges afterwards, as the optimal arm is identified.

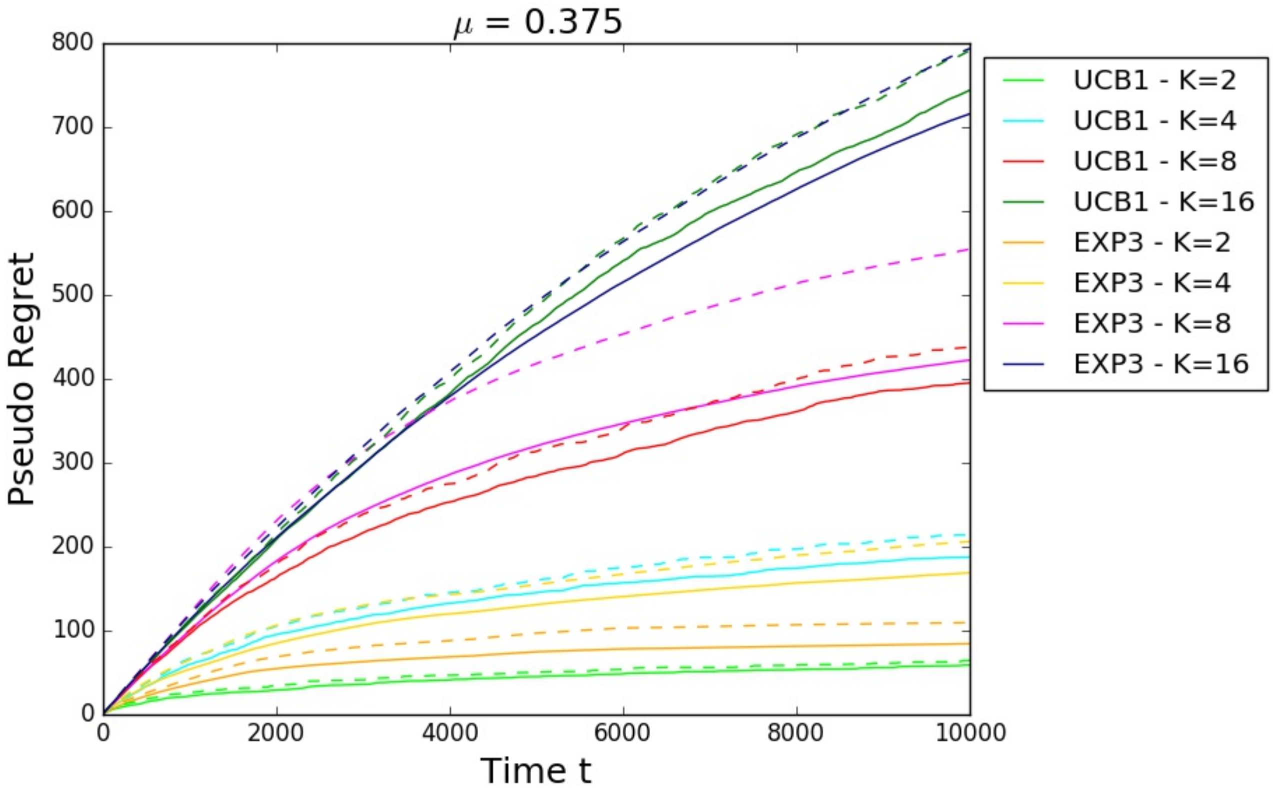

Figure 2: Average pseudo regret and average pseudo regret + standard deviation (scribbled line) over 10 runs with $\mu = 0.375$

$\mu^* - \frac{\mathbf{1}}{\mathbf{16}}$

Figure 2 shows the experiment for $\mu = 0.375$. The pseudo regret is higher than for $\mu = 0.25$, also for larger values of K, the regret does not converge as fast. It is also noticeable, that for $K = 16$ the pseudo regret is higher for the UCB1 algorithm.

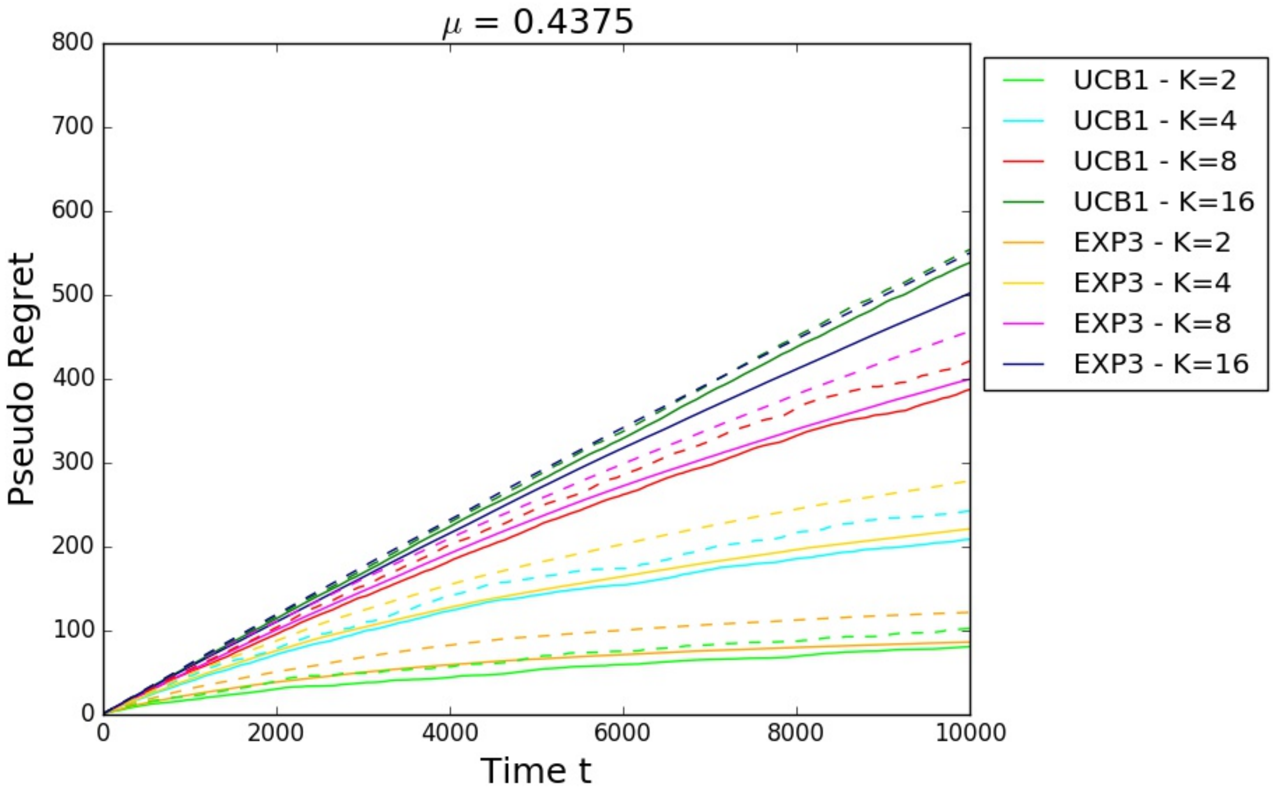

Figure 3: Average pseudo regret and average pseudo regret + standard deviation (scribbled line) over 10 runs with $\mu = 0.4375$

Figure 3 shows the experiment for $\mu = 0.4375$. Interestingly the pseudo regret for $K = 16$ is smaller than for the previous settings of μ.