Statistical Techniques in Robotics (16-831)

Project 1 : Weighted Majority Algorithm

2.1 Realizability

a) Realizability is a condition in online learning where it is assumed that :

1. All answers are generated by a target mapping

$$h^* : \mathcal{X} \longrightarrow \mathcal{Y}$$

where $\mathcal{X}$ is the instance domain and $\mathcal{Y}$ is the target domain.

2. And there exists a perfect mapping h^* in the hypothesis domain $\mathcal{H}$

$$h^* \in \mathcal{H}$$

b) Realizability is important in online learning as it helps a mistake bound for the algorithm in situations where a consistent /perfect hypothesis exists.

In unrealizable cases, as no perfect hypothesis exists a bound on the relative regret is required to be calculated.

2.2 Hypothesis Class

a) Hypothesis class $\mathcal{H}$ is the set of target mapping functions. The mapping functions can be hypothesis/ classifiers /regressors /predictors.

$$\mathcal{H} = \{ h : \mathcal{X} \rightarrow \mathcal{Y} \}$$

b) Hypothesis class is ~~finitie~~ finite for the cases/algorithms discussed in class.

c) Hypothesis class is assumed finite so that the process of chosing a consistent hypothesis uniformly ~~at random~~ is well defined.

In the unrealizable cases, if the class is infinite, the weight update process for the hypothesis is again not well defined. It also increases the ~~mistake~~ bound and the regret bound infinitely

2.3 Regret

a) Yes, regret can be negative.

Example

In cases, when ~~regret~~ learner's loss is compared to the loss of the worst expert, the regret might be negative.

Assume, A, B, C are experts and L is learner.

A	B	C	L	Truth
0	1	0	0	1
1	0	0	0	0
0	1	1	1	1

Here after 3 time-steps, Loss(L) = 1 Loss(A) = 3

$\therefore$ Regret of L wrt A = ~~2~~ -2

b) ~~No No~~ Yes, regret vs the best expert in hindsight can be negative as learner's prediction is a cummulative result of the observations from the multiple experts and not just a single best expert ~~for~~ till that time step.

Example :

Suppose there are 2 experts A & B. ~~and~~ ~~have deteminis~~ If A is consistent for the first 20 time steps and B is consistent for the next ~~20~~ 100 time steps, then after 40 time steps, the learner's cummulative loss would be better than the best expert in hindsight (ie B).

2.4 Consistent Algorithm Regret Bound

To prove : $M \leq N \cdot m^* + (N-1)$

where m^* = mistakes of best expert

N = number of experts

M = number of mistakes of learner.

Proof →

let $m^* = 0$

This implies that ~~there~~ best expert is consistent

∴ for the realizable consistent algorithm

$M \leq (N-1)$ (given)

if $m^*_t = 1$

ie, The best expert commits 1 mistake

∴ for that time step t

$M_t \leq N$ (bounded by total number of experts as in worst case all can be wrong)

∴ Till time step t $M_t \leq N + (N-1)$

By induction,

if $m^*_t = 2$

$$M_t \leq 2N + (N-1)$$

$\vdots$

$\therefore$ $\cancel{m}$ $\cancel{M_t}$ for m^*_t mistakes of the ~~best~~ expert

$$\boxed{M_t \leq Nm^* + (N-1)}$$

2.5 Understanding the η parameter

$$E(R) \leq \eta m^* + \frac{\ln N}{\eta}$$

(a) Optimal value of η ⇒

When $E(R)$ ~~bound~~ is minimized, η is optimal

Differentiating $E(R)$ wrt η

$$\frac{dE(R)}{d\eta} = 0$$

$$m^* - \frac{\ln N}{\eta^2} = 0$$

$$m^* = \frac{\ln N}{\eta^2}$$

$$\boxed{\eta = \sqrt{\frac{\ln N}{m^*}}}$$ (Optimal value of η)

(b) When $N >> T$, η should be a bigger value.

Intuitively, when the number of hypothesis are much higher than the number of trials, the learner needs to converge the regret in less time. So the penalty parameter η on experts should be high at each step.

Mathematically,

Avg. regret $\frac{E(R)}{T} \to 0$ as $T \to \infty$

$$\frac{E(R)}{T} \leq \frac{\eta m^*}{T} + \frac{\ln N}{\eta T}$$

as $N >> T$

$$\frac{\ln N}{T} >> 1$$

(as $\ln N$ sublinearly increases with N)

To create a tighter bound $\frac{E(R)}{T}$, η should be high enough to make $\frac{\ln N}{\eta T} < \cancel{0}1$

ie $\eta \geq \frac{\ln N}{T}$.

Also, η should not ~~that~~ be high enough to make $\frac{\eta m^*}{T}$ very high. This is taken care by the linear nature of $\ln N$.

c) If $m^* = O(T)$, η should be chosen to make $E(R)$ ~~ηm^*~~ sublinear in T.

$\therefore$ when $\eta = \cancel{0}\frac{1}{T}$

$\eta m^* = O(1)$ (sub-linear) but $\frac{\ln N}{\eta} = O(T)$ (linear)

when $\eta = \frac{1}{\sqrt{T}}$

$\eta m^* = O(\sqrt{T})$ (sub-linear) but $\frac{\ln N}{\eta} = O(\sqrt{T})$ (sub-linear)

$$\therefore \boxed{\eta = \frac{1}{\sqrt{T}}}$$

If m^* is sub-linear $(O(\sqrt{T}))$

then

$\eta = k$ (const)

$\eta m^* = O(\sqrt{T})$ and $\frac{\ln N}{\eta} = O(1)$

(sub-linear) (sub-linear)

$\therefore \eta$ should be chosen a const k (independent of T)

2.6 Understanding Adversarial Environments

~~(1) For WMA, ~~to maximise~~~~

~~mistakes of the Learner $M^{(T)} \leq 2(1+\eta)m_n^{(T)} + \frac{2\ln N}{\eta}$~~

~~To maximise $M^{(T)}$, $m_n^{(T)}$, ie, mistakes of expert n~~

~~should be maximised in the upper bound.~~

(1) Maximising net loss for WMA requires the nature to give adversarial true labels or outcomes.

As in deterministic algorithm, the learner chooses the experts in majority based on their weights. Thus, the expert with maximum wt will be more deterministic for the Learner's prediction.

$\therefore$ A good strategy for the adversary would be to choose an expert with max weight at each time step and report an outcome opposite to that expert.

(2) Intuitively, RWMA has a stochastic learner while WMA has a deterministic learner that aligns with the nature's approach to give adversarial outcomes. Even in the max loss case for RWMA, the loss would be better than WMA due to the randomized ~~sampling~~ selection of experts.

Mathematically,

in the worst adversarial case,

the mistakes of the learner would be maximum

for WMA

$$M^{(T)} \leq 2(1+\eta)\, m_n^{(T)} + \frac{2\ln N}{\eta}$$

for RWMA

$$M^{(T)} \leq (1+\eta)\, m_n^{(T)} + \frac{\ln N}{\eta}$$

as can be seen, $(M^{(T)}_{WMA})_{max} = 2\,(M^{(T)}_{RWMA})_{max}$

$\therefore$ for worst adversary RWMA is strictly better than WMA.

Collaborators:

Richa Varma

Keerthana Mannivannan

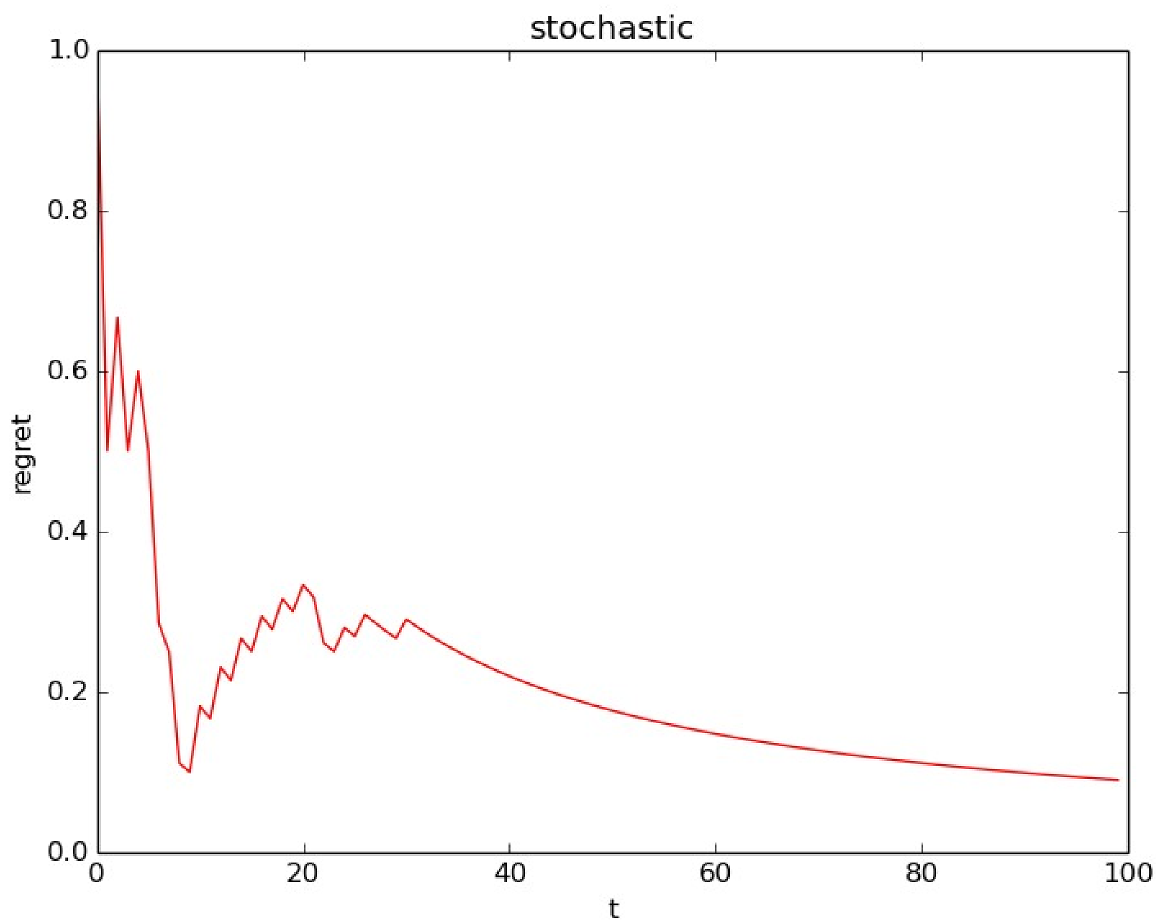
Jimit Gandhi

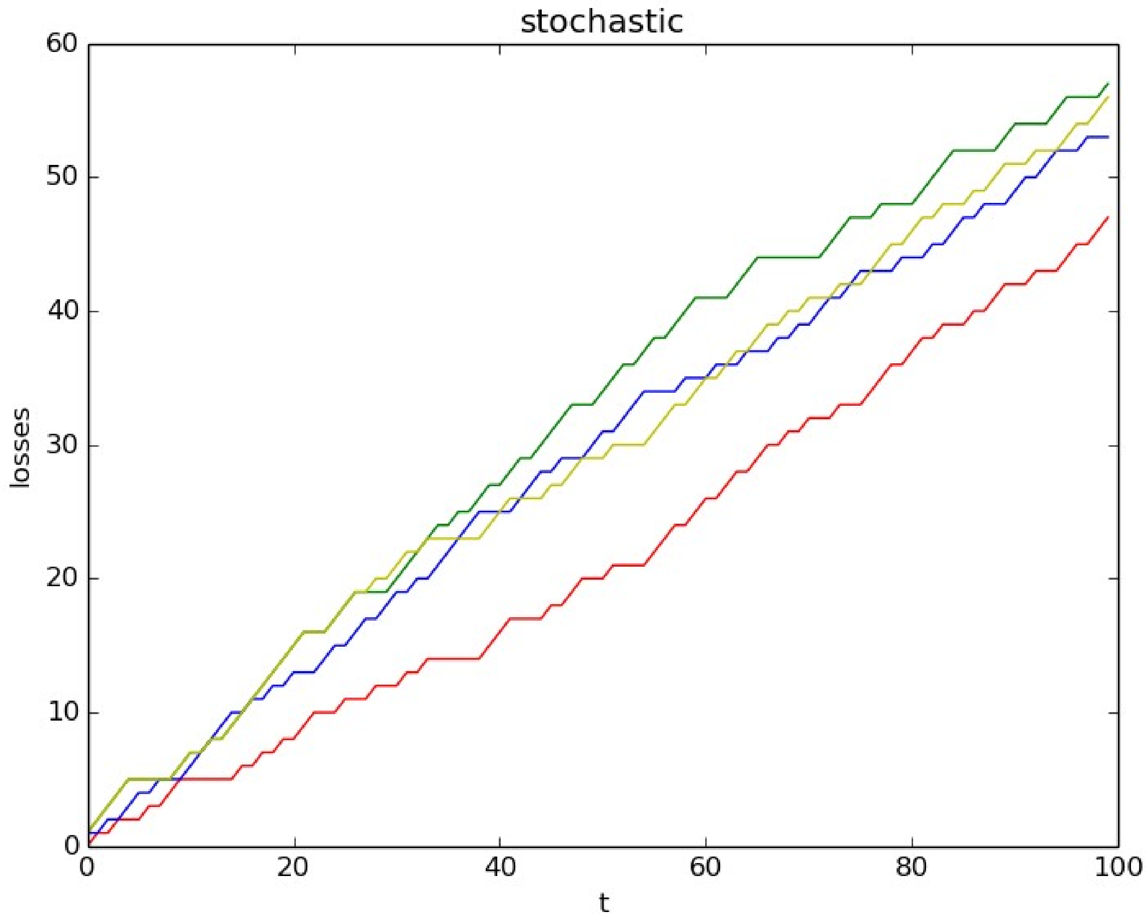
Rohan Thakker

Explanation of plots:
For all the plots of losses, following legend has been used:

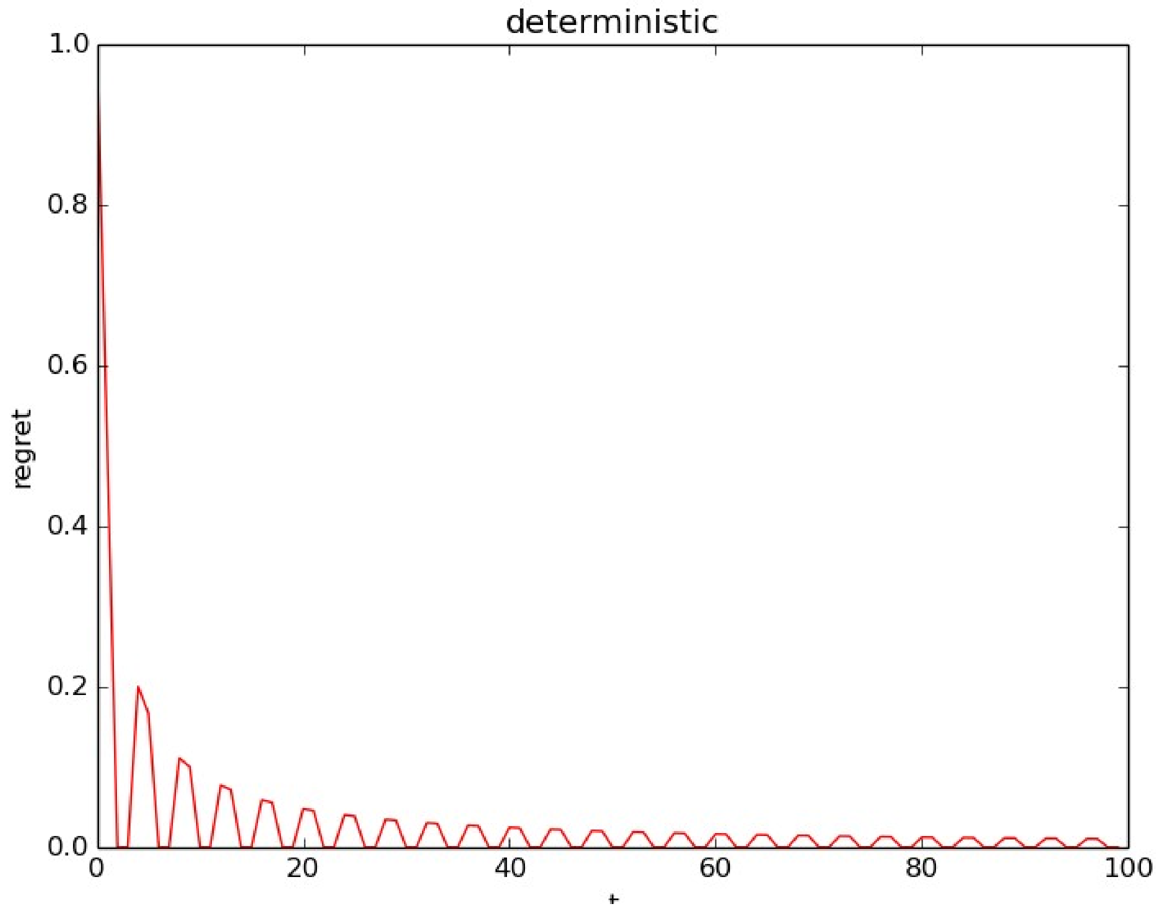
Expert 1 : Red line
Expert 2: Blue line
Expert 3: Green line
Expert 4: Magenta line
Learner: Yellow line

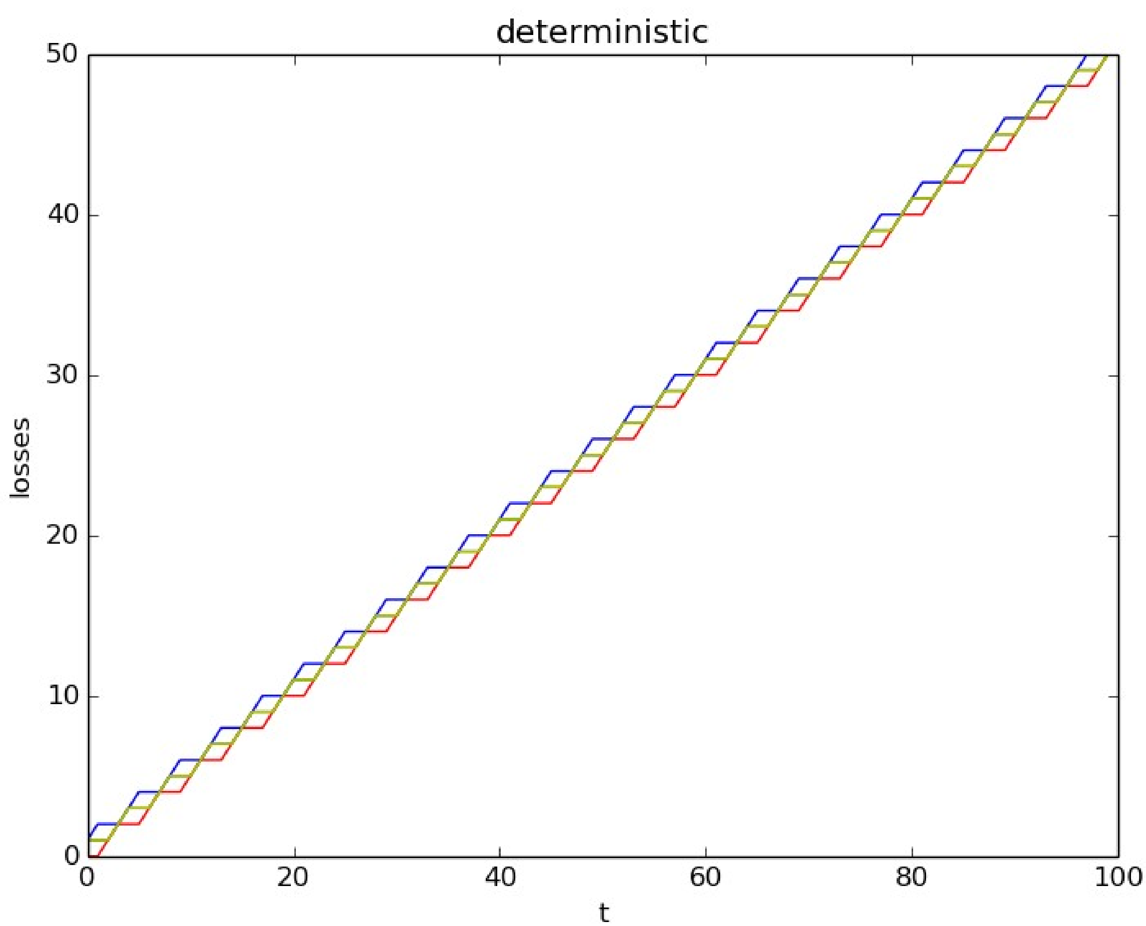
WMA(3 experts):
T=100 eta=0.1
1. Stochastic: The average regret was converging to 0 after 100 obs



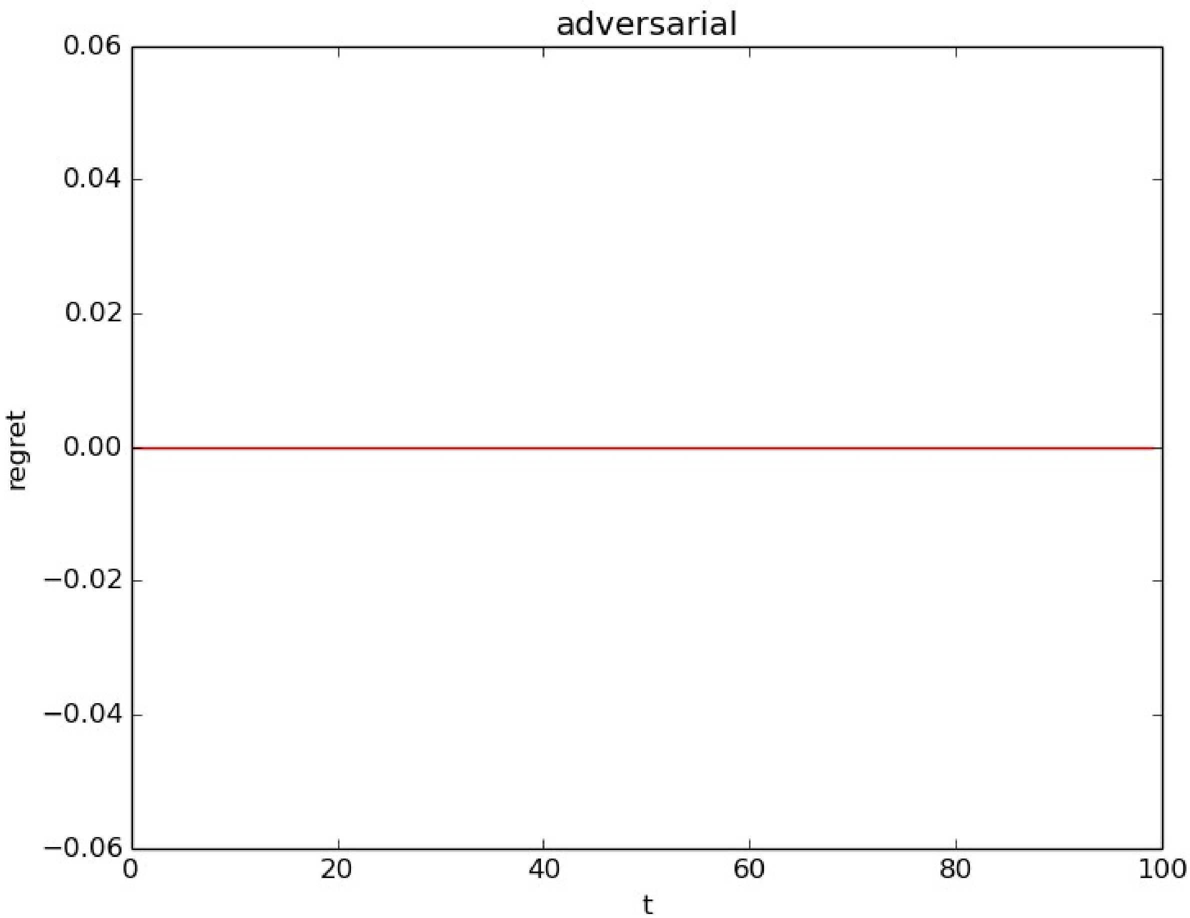


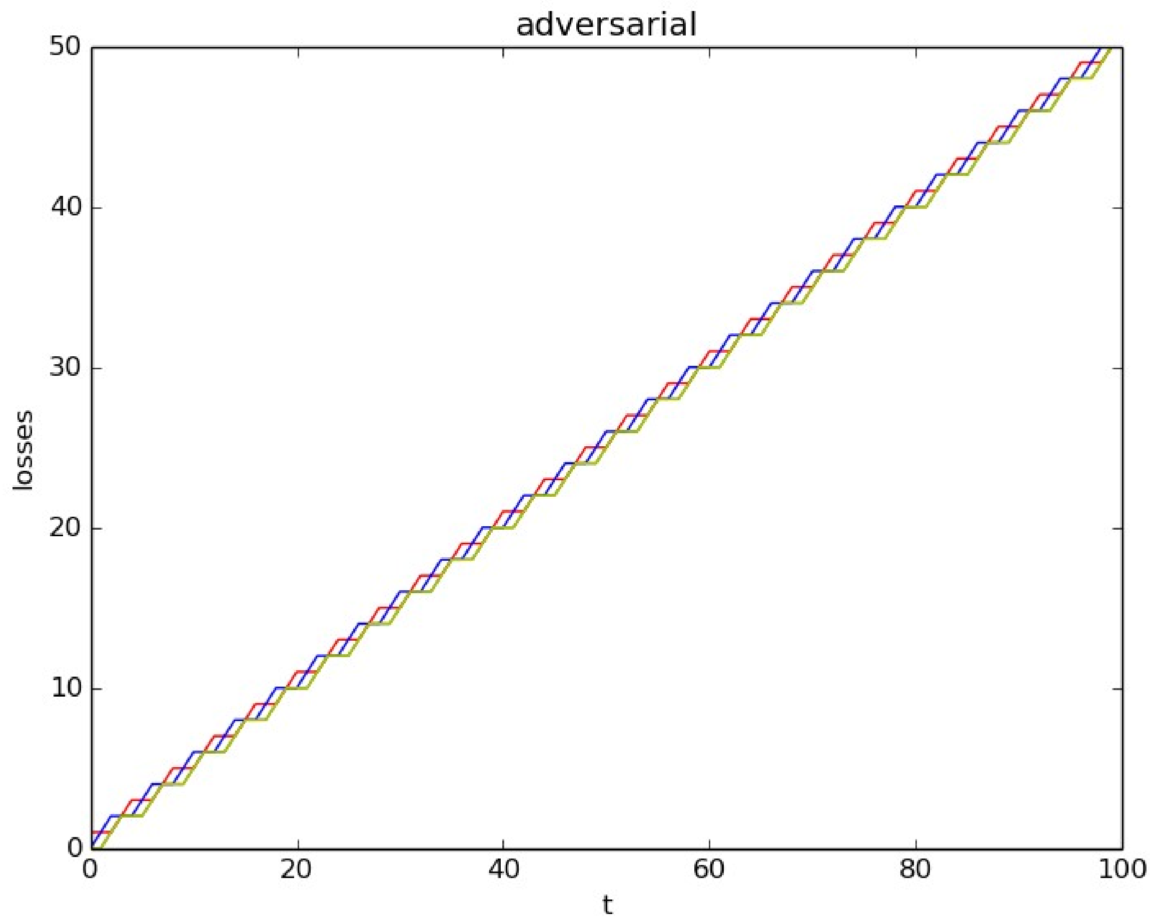
2. Deterministic: The losses increased linearly with time; regret sinosoidally converged to zero


deterministic
regret
t
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
100

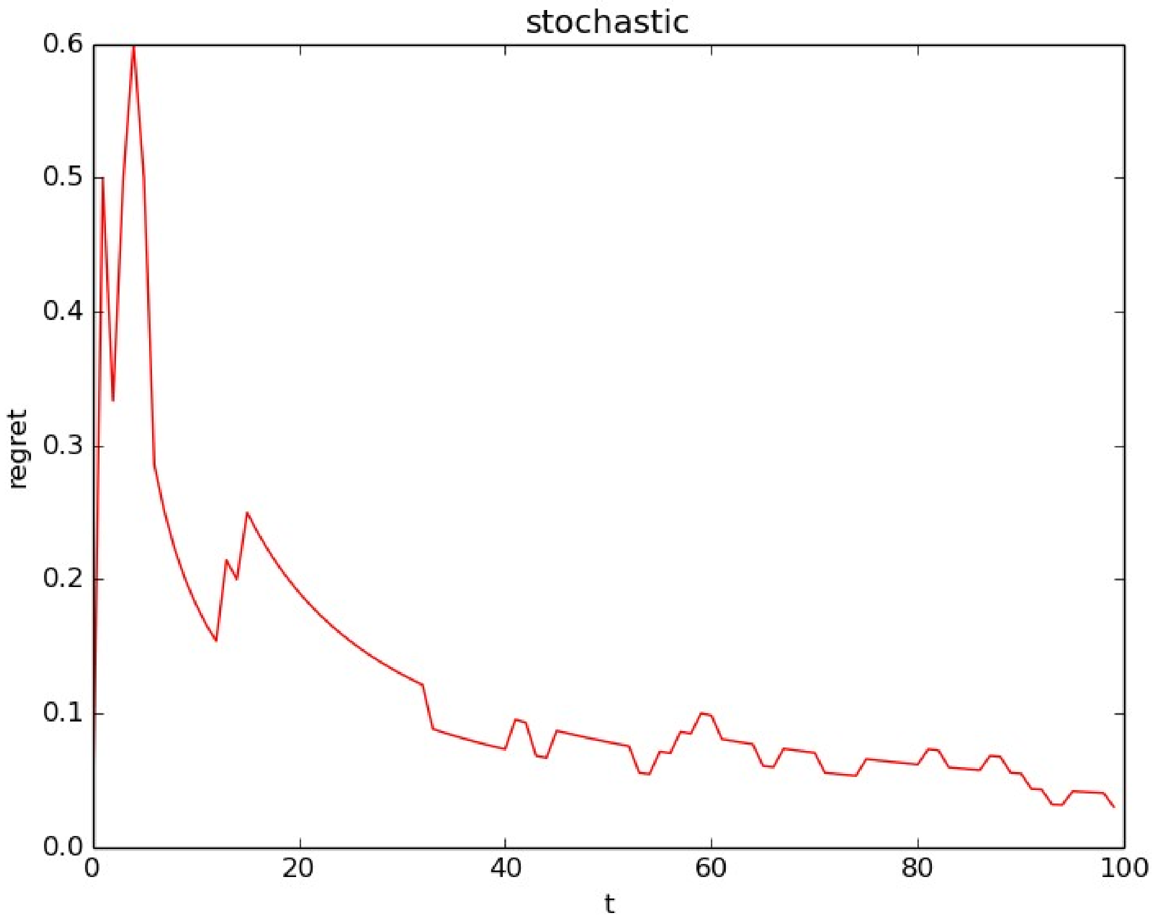

deterministic
losses
t
0
10
20
30
40
50
0
20
40
60
80
100

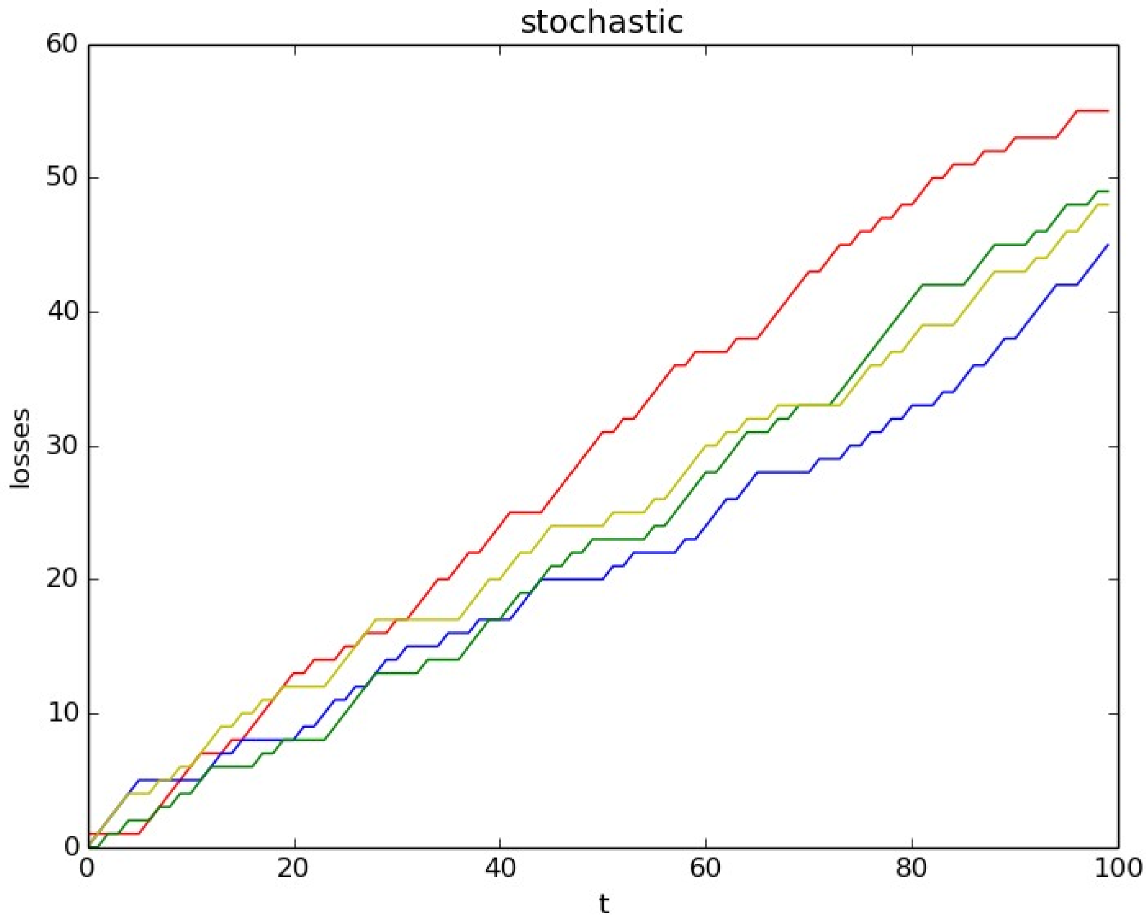
3. Adversarial: Losses increased linearly; regret was 0



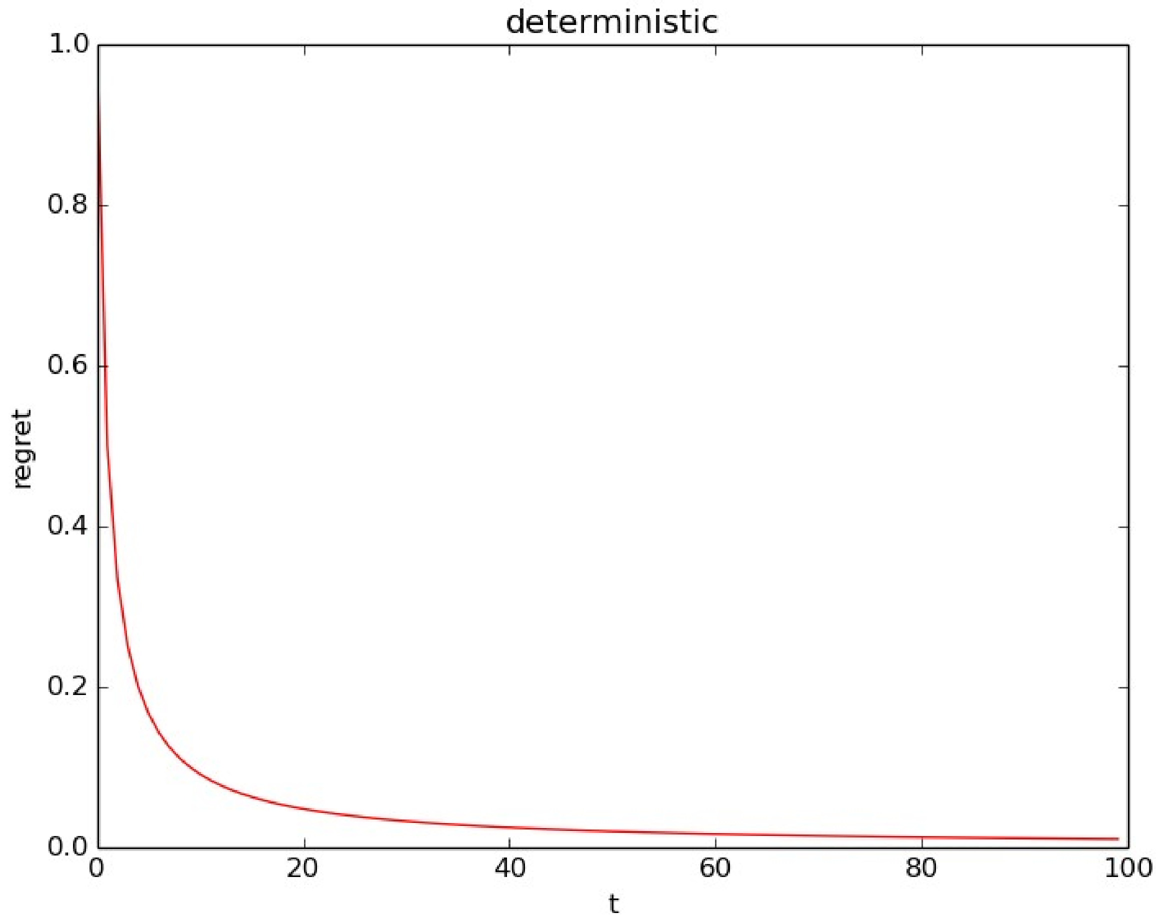


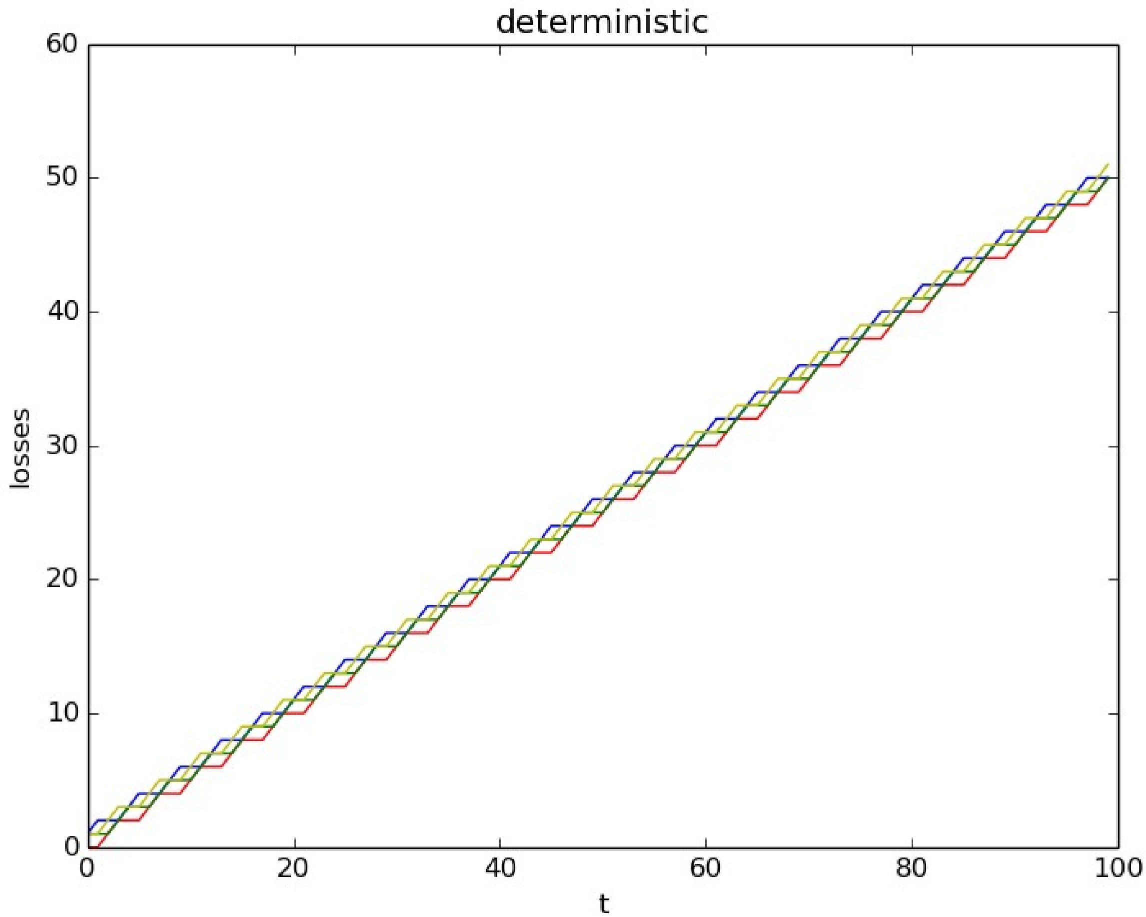
RWMA(3 experts):
T=100 eta=1/sqrt(T)
1. Stochastic: Learner's loss was less compared to wma; avg regret converged to 0


stochastic
regret
t
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
20
40
60
80
100

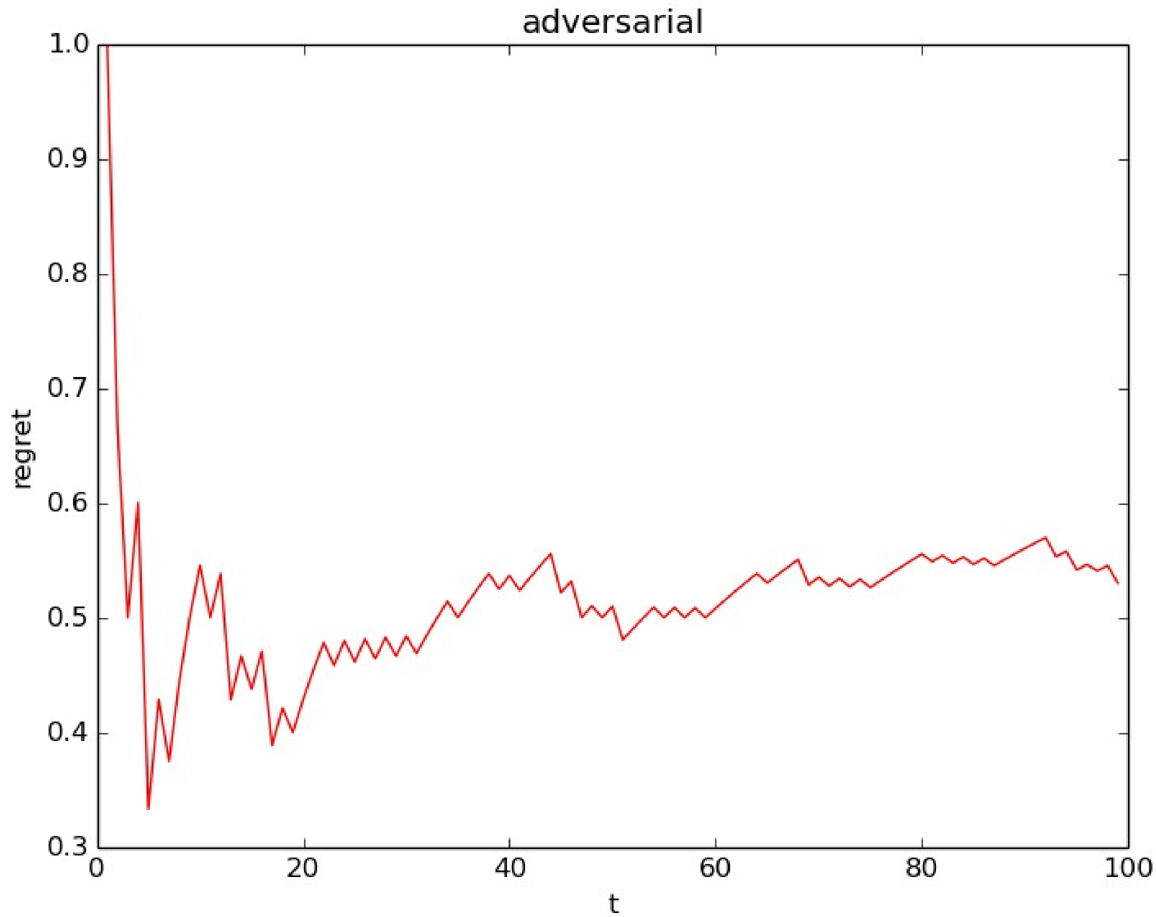


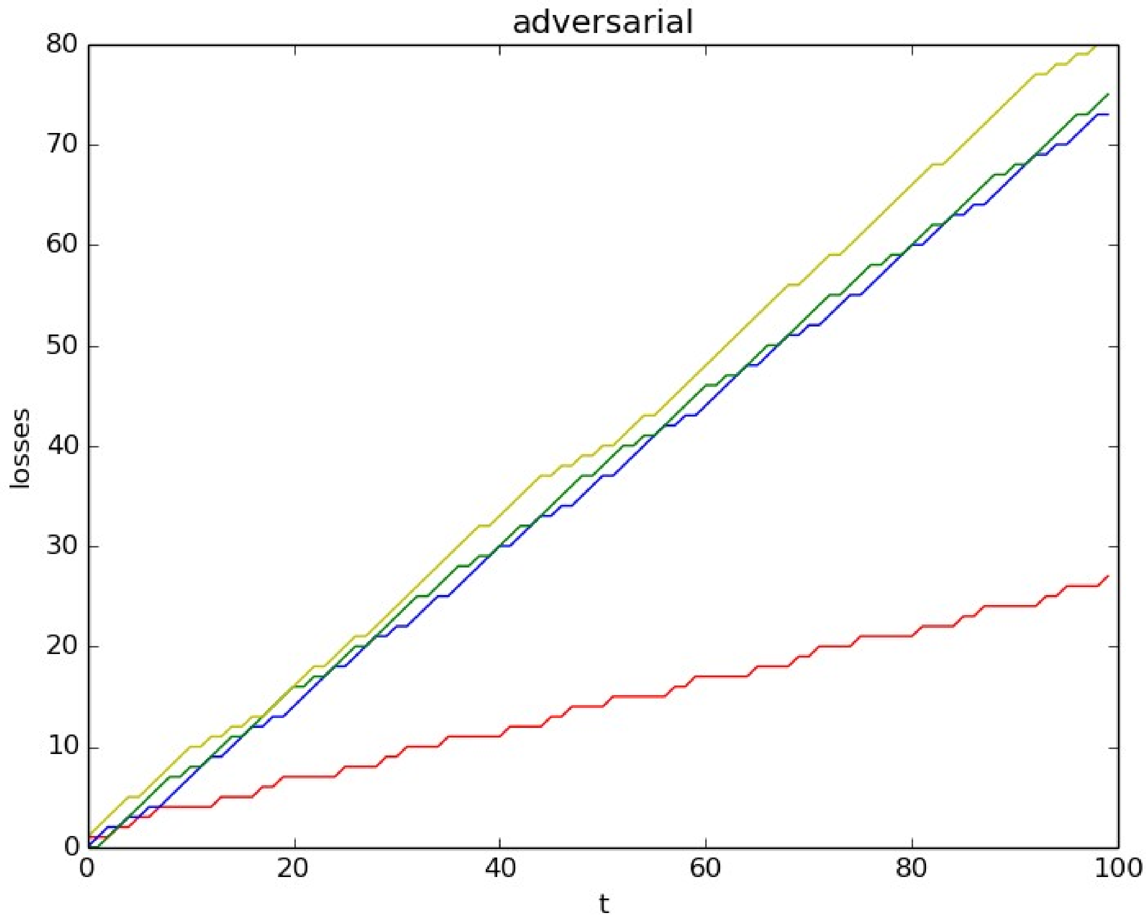
2. Deterministic: The losses increased linearly with time; regret converged to zero


deterministic
regret
t
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100

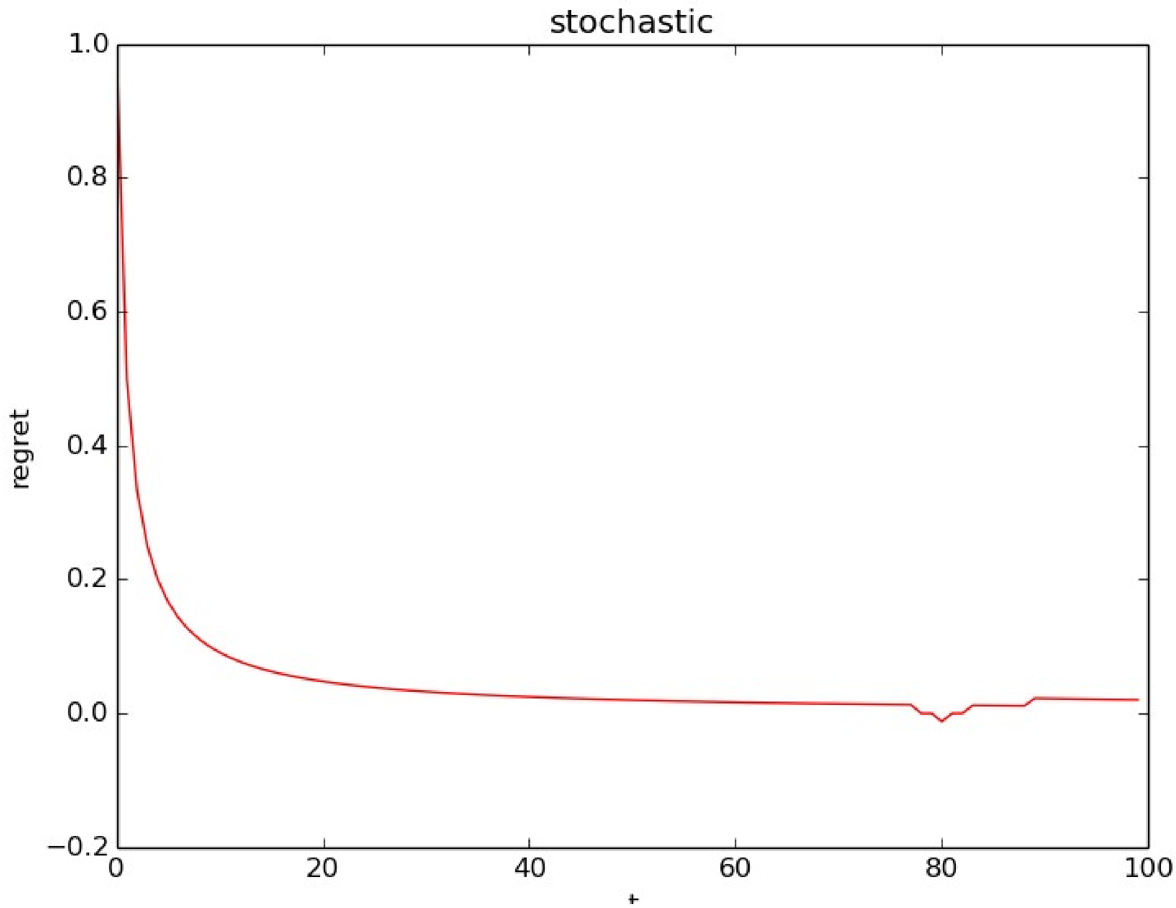


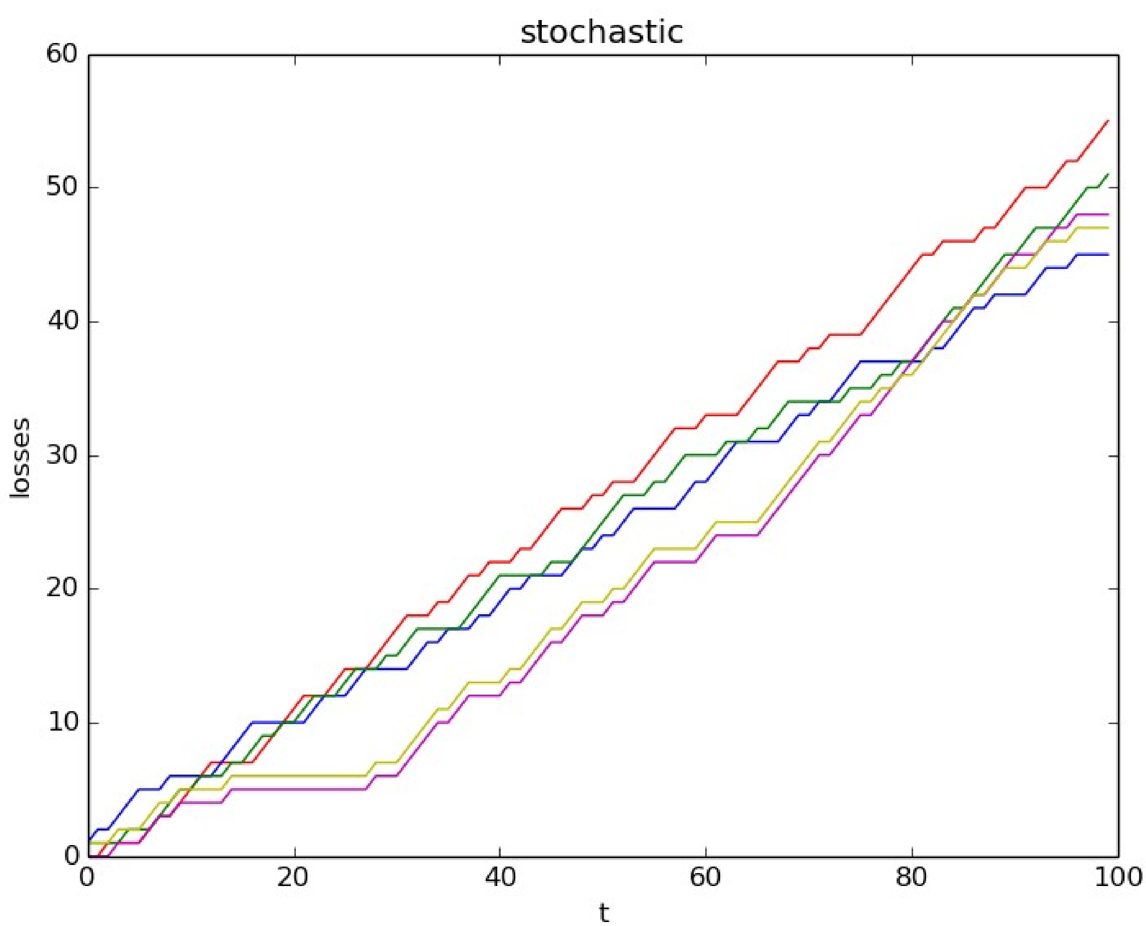
3. Adversarial: Regret converged at 0.5


adversarial
regret
t
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0
20
40
60
80
100

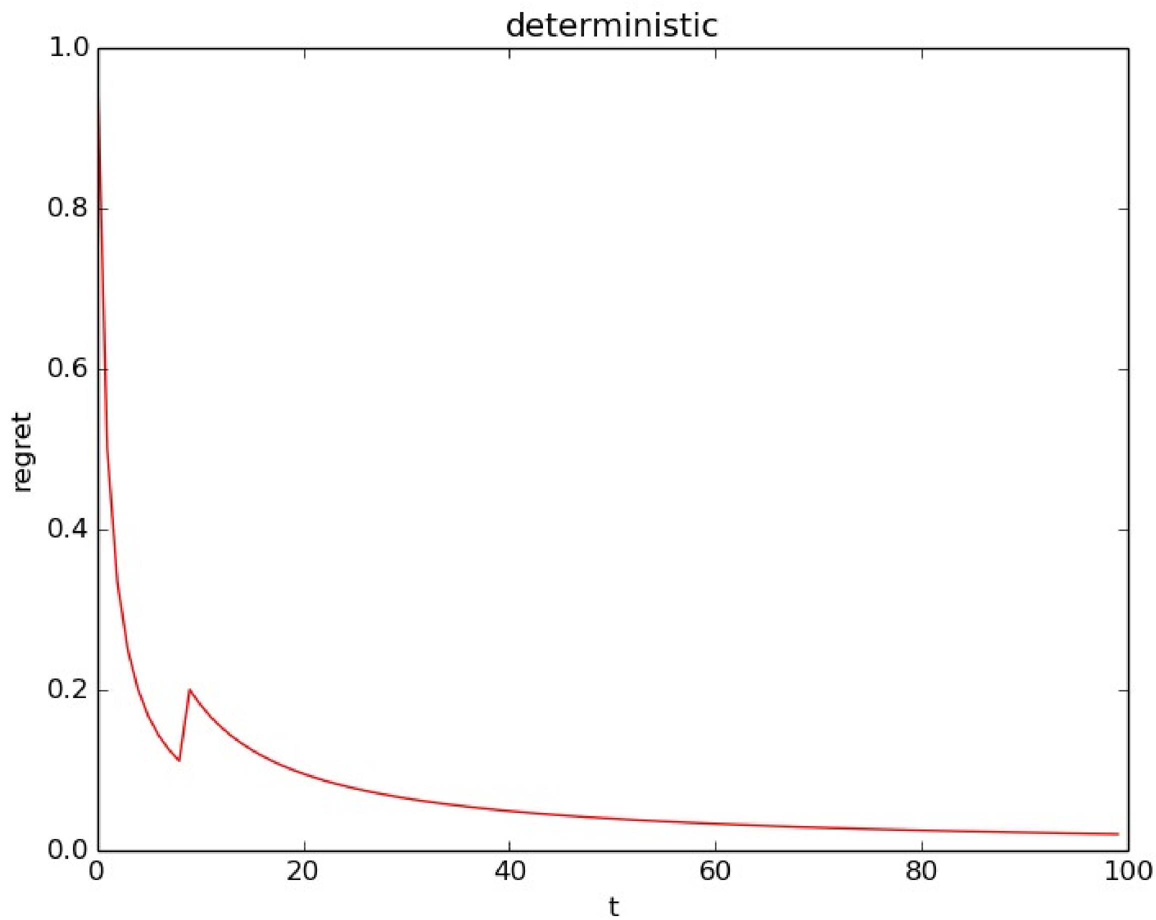


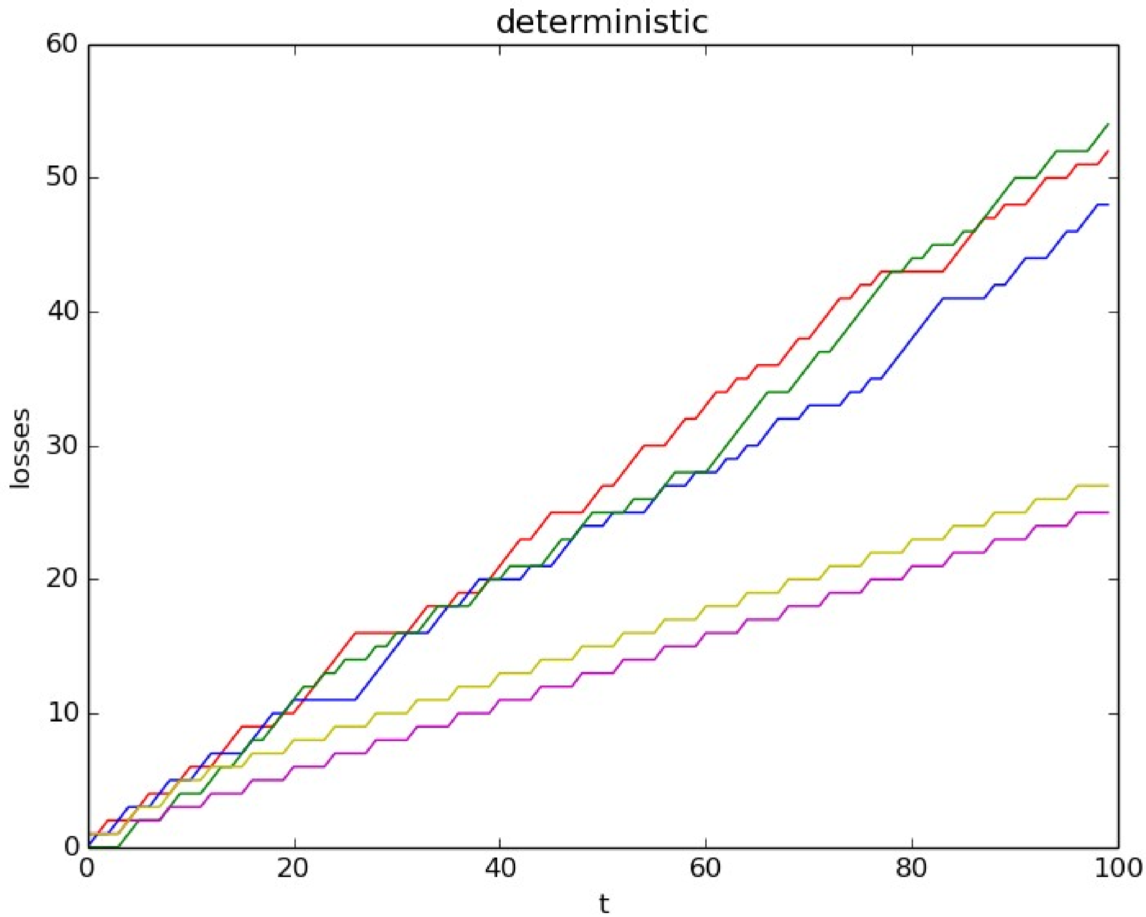
WMA(4 experts):
T=100 eta=0.1
1. Stochastic:


stochastic
regret
t
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
0
20
40
60
80
100

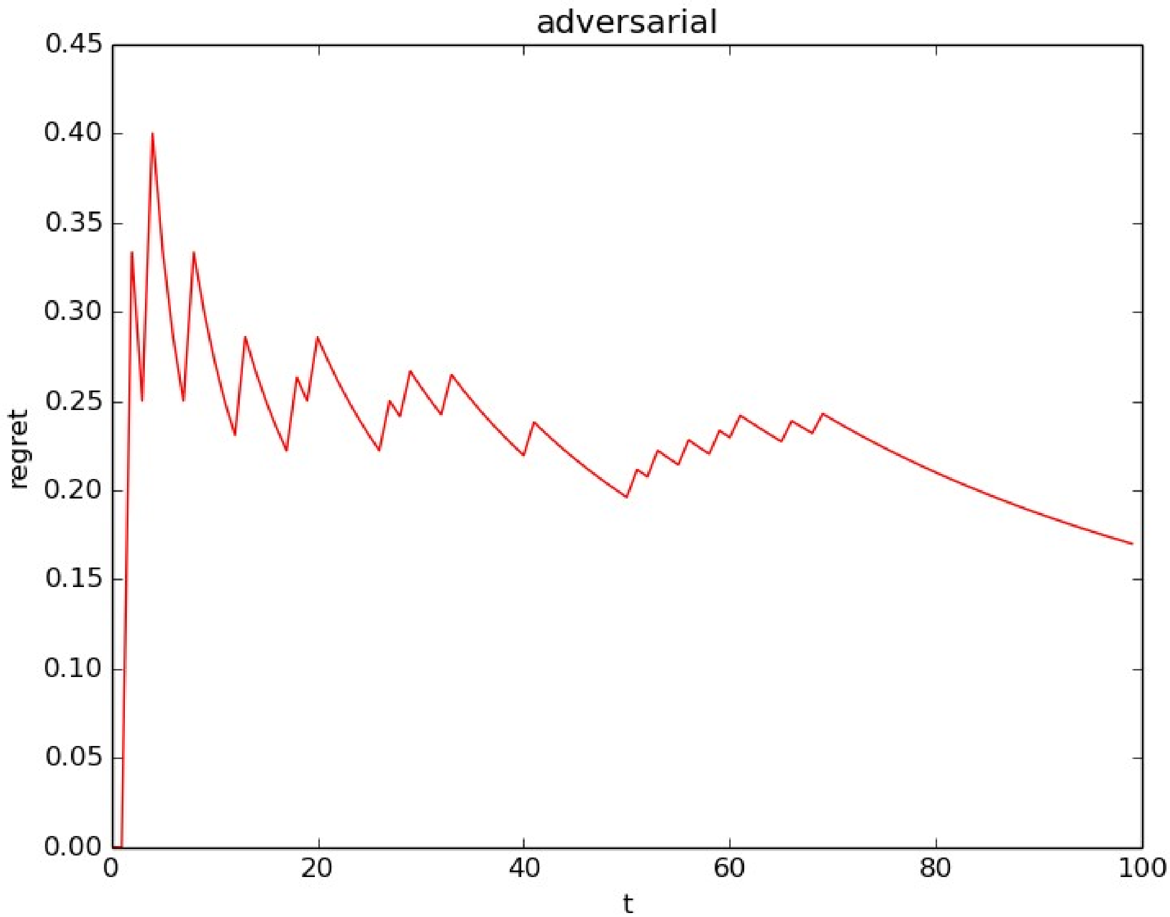

stochastic
losses
t
60
50
40
30
20
10
0
0
20
40
60
80
100

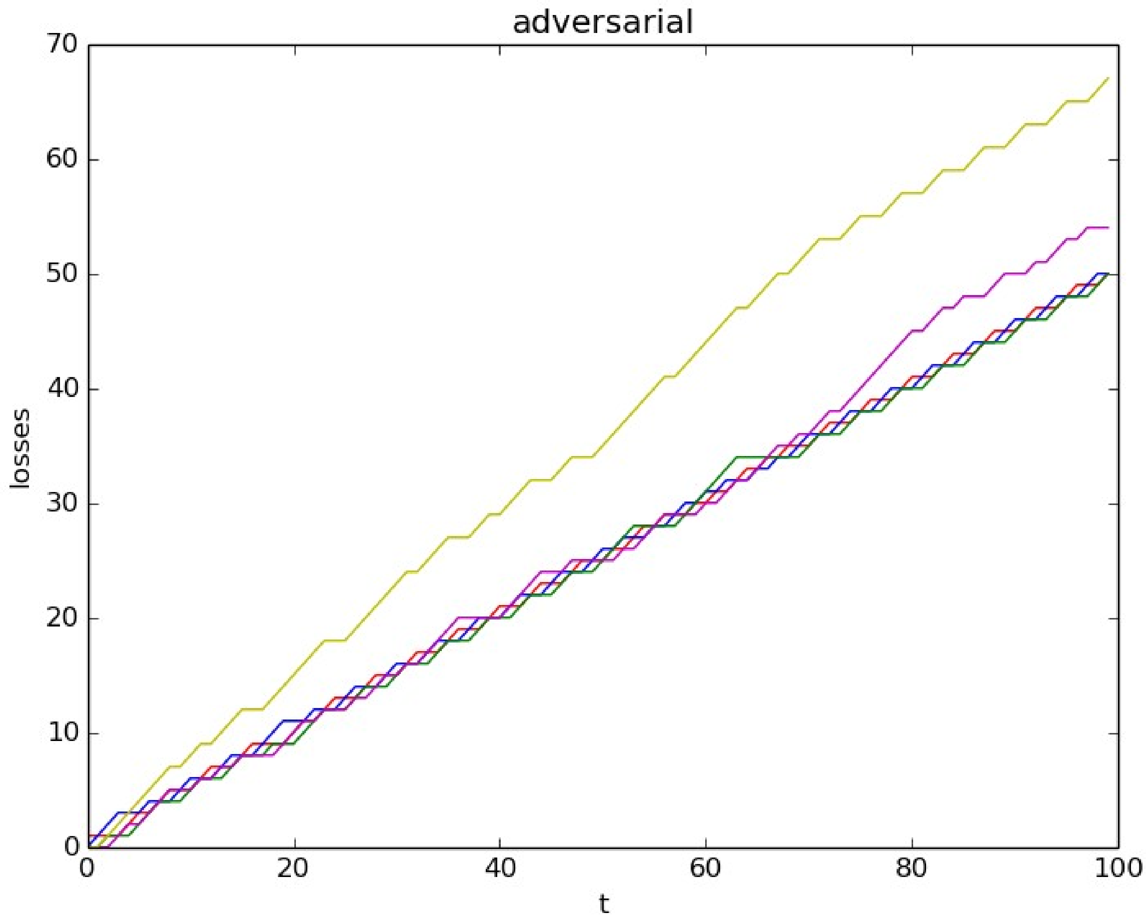
2. Deterministic:





3. Adversarial:


adversarial
regret
t
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
20
40
60
80
100



adversarial
losses
t
0
10
20
30
40
50
60
70
0
20
40
60
80
100