Game Theory

Lecture notes for MATH11090 & MATH09002

Peter Richtárik

University of Edinburgh

November 16, 2010

Learning in Games

So far we have

- studied one-shot games (**static games**) only
  - no repetition
  - no chance to **learn** what the other person is doing
  - players just choose their pure (mixed) strategies once, get a payoff (expected payoff) and GAME OVER
- assumed that **strategies and payoffs are known** to all players in advance

These assumptions (no repetition, known payoffs) may be **unrealistic**

- Many games are played repeatedly
- Payoffs are not known

We will study repeated game-playing with unknown payoffs using a very intuitive, flexible and powerful algorithm:

- **Weighted Majority Algorithm (WM),** and its generalized version
- **Multiplicative Weights Update Algorithm (MWU)**

We will apply the results we get to show that **natural game-playing strategies in matrix games converge to the value of the game**

Example 1: Premier League Betting (1)

You want to bet on football games in the Premier League.

There are 20 teams, each is playing 38 matches:

$$T = (38 \times 20)/2 = 380 \text{ matches}$$

Current League Table

LEAGUE TABLE BARCLAYS PREMIER LEAGUE

Updated 18:03 14th November 2010

			Home						Away						
POS		NAME	P	W	D	L	F	A	W	D	L	F	A	GD	PTS
1	-	Chelsea	13	6	0	1	17	3	3	1	2	11	5	+20	28
2	▲	Arsenal	13	4	0	2	15	6	4	2	1	11	6	+14	26
3	▼	Man Utd	13	5	1	0	15	5	1	6	0	11	10	+11	25
4	-	Manchester City	13	3	3	1	7	5	3	1	2	8	5	+5	22
5	-	Bolton	13	2	3	1	10	8	2	4	1	11	11	+2	19
6	▲	Sunderland	13	3	3	0	7	3	1	4	2	8	10	+2	19
7	▼	Tottenham	13	3	3	1	11	7	2	1	3	7	10	+1	19
8	▼	Newcastle	13	2	2	3	15	9	3	1	2	6	7	+5	18
9	▼	Aston Villa	13	3	4	0	10	5	1	1	4	5	13	-3	17
10	-	Stoke City	13	4	1	2	11	8	1	0	5	4	10	-3	16
11	-	Liverpool	13	3	2	1	9	6	1	2	4	4	11	-4	16
12	-	WBA	13	3	2	1	8	6	1	2	4	8	16	-6	16
13	-	Everton	13	2	3	2	9	8	1	3	2	5	5	+1	15
14	-	Blackburn	13	2	2	2	6	6	2	1	4	9	12	-3	15
15	-	Blackpool	13	1	2	2	9	10	3	1	4	10	16	-7	15
16	-	Fulham	13	2	3	1	8	6	0	5	2	5	7	0	14
17	-	Wigan Athletic	13	2	3	3	6	15	1	2	2	4	6	-11	14
18	-	Birmingham	13	2	3	1	6	5	0	4	3	8	12	-3	13
19	-	Wolves	13	2	2	3	9	11	0	1	5	4	12	-10	9
20	-	West Ham Utd	13	1	3	3	7	11	0	3	3	4	11	-11	9

Secret Subliminal Slide

The Best Team Ever

Example 1: Premier League Betting (2)

Before every match you bet on a team:

- If you are **wrong** you **lose**
- If you are **right** you **win**

You can make your decisions using the **predictions of n experts** (game analysts, coaches):

- Each will tell you who they think will win
- It may be that all of them are very good or very poor, their judgment may be random or correlated, . . .
- You do not know who is the "best"
- In fact, who is the **best expert** will only be known in **hindsight** and will depend on T: someone who was good in the first week ($T = 7$) could be a big loser after 1 year ($T = 365$) and vice versa

Example 2: Buy or Sell?

Example 2: Buy or Sell?

You hold a stock and want to decide whether to buy or sell.

- Every day for T days you predict whether the price of your stock goes **up** or **down**
- You buy or sell depending on your predictions
  - If your prediction is **wrong** you **lose**
  - If your prediction is **right** you **gain**
- You can make your decisions using the predictions of n **experts** (market analysts, computer programs, ...)

How Can You Make Use of the Experts' Predictions?

Possible answers:

- On each day, do what the **majority** of the experts say
  - fails when most "experts" make poor predictions
- Wait some time, observe the experts, and then stick with the one who was **best**
  - it's possible that he/she was just lucky and will give poor predictions in the future
- What experts? These folks surely know much less about football than you do!
  - Of course, this is the way to go!
  - But: we would not have any math to talk about, so let's pretend this option is not available

Weighted Majority solves the problems using a **combination** of the two ideas:

- In each time period we will consider advice of **all** experts BUT
- each expert's advice will be **weighted** according to his past success

Weighted Majority Algorithm

1. At time $t = 1$ assign **unit weight** to all experts: $w_i^1 = 1, i = 1, \ldots, n$
2. Compute
   - 2.1 $p_i^t = w_i^t / \sum_j w_j^t$ for all i (note that $\sum_i p_i^t = 1$)
   - 2.2 $P_a^t =$ sum of p_i^t over experts i predicting outcome a
   - 2.3 $P_b^t =$ sum of p_i^t over experts i predicting outcome b
3. Decide (bet, invest) based on **weighted majority** (breaking ties arbitrarily):
   - 3.1 If $P_a^t > P_b^t$, assume a will happen and act (invest, bet) accordingly
   - 3.2 If $P_b^t > P_a^t$, assume b will happen and act (invest, bet) accordingly
4. Observe the outcome (a happened or b happened).
   - 4.1 If you **guessed wrong,** you incur a **loss** of 1.
   - 4.2 If you **guessed right,** you incur **no loss.**
5. **Shrink the weights** of the experts who were wrong
   - 5.1 $w_i^{t+1} = w_i^t$ if expert i was right
   - 5.2 $w_i^{t+1} = (1 - \epsilon) w_i^t$ if expert i was wrong
6. Proceed to time $t \leftarrow t + 1$ and go to step 2

WM Algorithm: Choice of the Learning Rate (1)

Parameter ϵ can be thought of as a **learning rate:**

- **Small** ϵ (close to 0) = **learning slowly** because
  - $(1 - \epsilon)$ will be close to 1 and hence
  - weights of failing experts will be decreased **very slowly**
- **Large** ϵ (close to 1) = **learning fast** because
  - $(1 - \epsilon)$ will be close to 0 and hence
  - weights of failing experts will be decreased **very fast**

Weighted Majority Algorithm: An Excel Example (1)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1		epsilon=	0.050			T=25											
2		REALITY	EXPERTS						ME								
3		Outcome	Predictions			Loss			Weights (w)			Normalized weights (p)			WA	WM Prediction	Loss
4	t		Exp 1	Exp 2	Exp 3	Exp 1	Exp 2	Exp 3	Exp 1	Exp 2	Exp 3	Exp 1	Exp 2	Exp 3			
5	1	1	1	0	0	0	1	1	1.0000	1.0000	1.0000	0.3333	0.3333	0.3333	0.3333	0	1
6	2	1	1	1	0	0	0	1	1.0000	0.9500	0.9500	0.3448	0.3276	0.3276	0.6724	1	0
7	3	1	0	1	0	1	0	1	1.0000	0.9500	0.9025	0.3506	0.3330	0.3164	0.3330	0	1
8	4	1	1	0	0	0	1	1	0.9500	0.9500	0.8574	0.3445	0.3445	0.3109	0.3445	0	1
9	5	1	1	1	0	0	0	1	0.9500	0.9025	0.8145	0.3562	0.3384	0.3054	0.6946	1	0
10	6	1	1	0	0	0	1	1	0.9500	0.9025	0.7738	0.3617	0.3436	0.2946	0.3617	0	1
11	7	1	1	0	1	0	1	0	0.9500	0.8574	0.7351	0.3737	0.3372	0.2891	0.6628	1	0
12	8	1	1	1	0	0	0	1	0.9500	0.8145	0.7351	0.3801	0.3259	0.2941	0.7059	1	0
13	9	1	1	0	0	0	1	1	0.9500	0.8145	0.6983	0.3857	0.3307	0.2835	0.3857	0	1
14	10	1	1	0	0	0	1	1	0.9500	0.7738	0.6634	0.3980	0.3241	0.2779	0.3980	0	1
15	11	1	0	0	0	1	1	1	0.9500	0.7351	0.6302	0.4103	0.3175	0.2722	0.0000	0	1
16	12	1	1	1	0	0	0	1	0.9025	0.6983	0.5987	0.4103	0.3175	0.2722	0.7278	1	0
17	13	1	1	0	0	0	1	1	0.9025	0.6983	0.5688	0.4160	0.3219	0.2622	0.4160	0	1
18	14	1	1	0	0	0	1	1	0.9025	0.6634	0.5404	0.4285	0.3150	0.2565	0.4285	0	1
19	15	1	1	0	0	0	1	1	0.9025	0.6302	0.5133	0.4411	0.3080	0.2509	0.4411	0	1
20	16	1	1	0	0	0	1	1	0.9025	0.5987	0.4877	0.4538	0.3010	0.2452	0.4538	0	1
21	17	1	1	0	0	0	1	1	0.9025	0.5688	0.4633	0.4665	0.2940	0.2395	0.4665	0	1
22	18	1	1	0	0	0	1	1	0.9025	0.5404	0.4401	0.4793	0.2870	0.2337	0.4793	0	1
23	19	1	0	0	0	1	1	1	0.9025	0.5133	0.4181	0.4921	0.2799	0.2280	0.0000	0	1
24	20	1	1	0	0	0	1	1	0.8574	0.4877	0.3972	0.4921	0.2799	0.2280	0.4921	0	1
25	21	1	1	0	0	0	1	1	0.8574	0.4633	0.3774	0.5049	0.2728	0.2222	0.5049	1	0
26	22	1	1	1	0	0	0	1	0.8574	0.4401	0.3585	0.5177	0.2658	0.2165	0.7835	1	0
27	23	1	1	0	0	0	1	1	0.8574	0.4401	0.3406	0.5234	0.2687	0.2079	0.5234	1	0
28	24	1	1	0	0	0	1	1	0.8574	0.4181	0.3235	0.5362	0.2615	0.2023	0.5362	1	0
29	25	1	1	0	1	0	1	0	0.8574	0.3972	0.3074	0.5489	0.2543	0.1968	0.7457	1	0
30		Total Loss				3	19	23									15

Weighted Majority Algorithm: An Excel Example (2)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1		epsilon=	0.500			T=25											
2		REALITY	EXPERTS						ME								
3		Outcome	Predictions			Loss			Weights (w)			Normalized weights (p)			WA	WM Prediction	Loss
4	t		Exp 1	Exp 2	Exp 3	Exp 1	Exp 2	Exp 3	Exp 1	Exp 2	Exp 3	Exp 1	Exp 2	Exp 3			
5	1	1	1	0	0	0	1	1	1.0000	1.0000	1.0000	0.3333	0.3333	0.3333	0.3333	0	1
6	2	1	1	1	0	0	0	1	1.0000	0.5000	0.5000	0.5000	0.2500	0.2500	0.7500	1	0
7	3	1	0	1	0	1	0	1	1.0000	0.5000	0.2500	0.5714	0.2857	0.1429	0.2857	0	1
8	4	1	1	0	0	0	1	1	0.5000	0.5000	0.1250	0.4444	0.4444	0.1111	0.4444	0	1
9	5	1	1	1	0	0	0	1	0.5000	0.2500	0.0625	0.6154	0.3077	0.0769	0.9231	1	0
10	6	1	1	0	0	0	1	1	0.5000	0.2500	0.0313	0.6400	0.3200	0.0400	0.6400	1	0
11	7	1	1	0	1	0	1	0	0.5000	0.1250	0.0156	0.7805	0.1951	0.0244	0.8049	1	0
12	8	1	1	1	0	0	0	1	0.5000	0.0625	0.0156	0.8649	0.1081	0.0270	0.9730	1	0
13	9	1	1	0	0	0	1	1	0.5000	0.0625	0.0078	0.8767	0.1096	0.0137	0.8767	1	0
14	10	1	1	0	0	0	1	1	0.5000	0.0313	0.0039	0.9343	0.0584	0.0073	0.9343	1	0
15	11	1	0	0	0	1	1	1	0.5000	0.0156	0.0020	0.9660	0.0302	0.0038	0.0000	0	1
16	12	1	1	1	0	0	0	1	0.2500	0.0078	0.0010	0.9660	0.0302	0.0038	0.9962	1	0
17	13	1	1	0	0	0	1	1	0.2500	0.0078	0.0005	0.9679	0.0302	0.0019	0.9679	1	0
18	14	1	1	0	0	0	1	1	0.2500	0.0039	0.0002	0.9837	0.0154	0.0010	0.9837	1	0
19	15	1	1	0	0	0	1	1	0.2500	0.0020	0.0001	0.9918	0.0077	0.0005	0.9918	1	0
20	16	1	1	0	0	0	1	1	0.2500	0.0010	0.0001	0.9959	0.0039	0.0002	0.9959	1	0
21	17	1	1	0	0	0	1	1	0.2500	0.0005	0.0000	0.9979	0.0019	0.0001	0.9979	1	0
22	18	1	1	0	0	0	1	1	0.2500	0.0002	0.0000	0.9990	0.0010	0.0001	0.9990	1	0
23	19	1	0	0	0	1	1	1	0.2500	0.0001	0.0000	0.9995	0.0005	0.0000	0.0000	0	1
24	20	1	1	0	0	0	1	1	0.1250	0.0001	0.0000	0.9995	0.0005	0.0000	0.9995	1	0
25	21	1	1	0	0	0	1	1	0.1250	0.0000	0.0000	0.9997	0.0002	0.0000	0.9997	1	0
26	22	1	1	1	0	0	0	1	0.1250	0.0000	0.0000	0.9999	0.0001	0.0000	1.0000	1	0
27	23	1	1	0	0	0	1	1	0.1250	0.0000	0.0000	0.9999	0.0001	0.0000	0.9999	1	0
28	24	1	1	0	0	0	1	1	0.1250	0.0000	0.0000	0.9999	0.0001	0.0000	0.9999	1	0
29	25	1	1	0	1	0	1	0	0.1250	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	1	0
30		Total Loss				3	19	23									5

Weighted Majority Algorithm: A Theorem

Theorem (Weighted Majority)

Let m_i^t (m^t) denote the number of mistakes that expert i makes (you make) in the first t predictions. Let $0 < \epsilon \le \frac{1}{2}$ be the **learning rate.** *Then for all i and t the following inequality holds*

$$m^t \le \frac{2 \ln n}{\epsilon} + 2(1+\epsilon) m_i^t .$$

In particular, this holds for the best expert i (which has minimum m_i^t).

This means that:

- You **won't make many more mistakes** than the **best expert in hindsight**
- That is, you are able to **learn** which experts are right and wrong

WM Algorithm: Choice of the Learning Rate (2)

Minimizing the number of mistakes: The WM theorem gives the upper bound on the number of mistakes

$$F(\epsilon) = \frac{2 \ln n}{\epsilon} + 2(1+\epsilon) m^t_{i(t)},$$

where $i(t)$ is the best expert at time t (the one with fewest mistakes).

Function $F(\epsilon)$

- is **convex** (since $F''(\epsilon) \geq 0$ for $\epsilon > 0$)
- has a **global minimum** at ϵ' satisfying $F'(\epsilon') = 0$: $\epsilon' = \sqrt{\frac{\ln n}{m^t_{i(t)}}}$
- is **decreasing** on $(0, \epsilon']$ and **increasing** on $[\epsilon', \infty)$

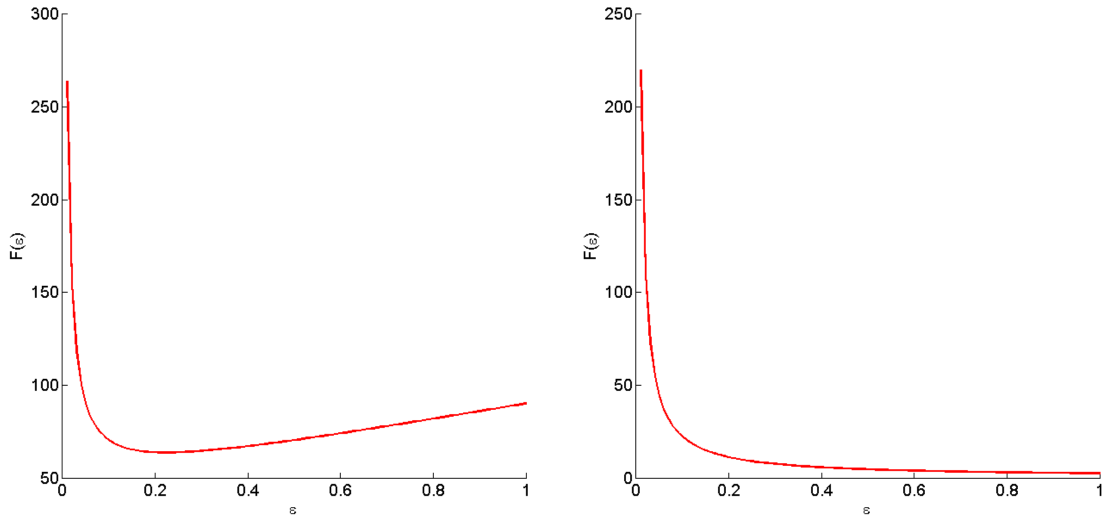
However, we are interested in the minimum ϵ^* of F on $(0, \frac{1}{2}]$:

- If $\epsilon' \leq \frac{1}{2}$ ($4 \ln n \leq m^t_{i(t)}$), then $\epsilon^* = \epsilon'$ and

$$F(\epsilon^*) = 4\sqrt{m^t_{i(t)} \ln n} + 2m^t_{i(t)}$$

- If $\epsilon' > \frac{1}{2}$ ($4 \ln n > m^t_{i(t)}$), then $\epsilon^* = \frac{1}{2}$ and $F(\epsilon^*) = 4 \ln n + 3m^t_{i(t)}$

WM Algorithm: Choice of the Learning Rate (3)



$\epsilon' \leq \frac{1}{2}$
($4 \ln n \leq m^t_{i(t)}$)
$\epsilon^* = \epsilon' = \sqrt{\frac{\ln n}{m^t_{i(t)}}}$

$\epsilon' > \frac{1}{2}$
($4 \ln n > m^t_{i(t)}$)
$\epsilon^* = \frac{1}{2}$

Proof of the Weighted Majority Theorem (1)

Proof.

Let $W^t = \sum_i w_i^t$. If you make a mistake at time t, then a weighted majority of the experts must have made a wrong prediction.

- Let A^t be the total weight at time t of the experts who did NOT make a mistake
- Let B^t be the total weight at time t of the experts who DID make a mistake

Then $W^{t+1} = A^t + (1-\epsilon)B^t \leq (1-\frac{\epsilon}{2})(A^t + B^t) = (1-\frac{\epsilon}{2})W^t$ for all t, and hence

$$W^t \leq (1-\tfrac{\epsilon}{2})^{m^t} W^1 = (1-\tfrac{\epsilon}{2})^{m^t} n.$$

However, we also know that

$$(1-\epsilon)^{m_i^t} = w_i^t \leq W^t.$$

Putting these two inequalities together, we get

$$(1-\epsilon)^{m_i^t} \leq (1-\tfrac{\epsilon}{2})^{m^t} n. \qquad (\#)$$

□

Proof of the Weighted Majority Theorem (2)

Proof (Continued).

Taking logarithms on both sides of (#) gives

$$m^t \leq \frac{\ln n}{-\ln(1-\frac{\epsilon}{2})} + m_i^t \frac{\ln(1-\epsilon)}{\ln(1-\frac{\epsilon}{2})}.$$

We have obtained a bound on m^t now!

We could have stated the theorem like this, but all the logarithms hinder proper insight into the dependence of the expression on ϵ. We can simplify this though, by using (three times!) the inequality

$$-x - x^2 \leq \ln(1-x) \leq x,$$

which holds for $0 \leq x \leq \frac{1}{2}$.

□

Limitations of the Weighted Majority Setup

Four main limitations:

- At each time step t, a decision is made **deterministically** (this is analogous to playing **pure strategies**)
- There are **only two decisions** to be made at every t (this is analogous to having two pure strategies only)
- There are only two possible outcomes (this corresponds to the **column player** having just **two** pure strategies at every t)
- The **loss is binary,** depending on which of the two events occurred (this corresponds to the payoff matrix B containing only zeros and ones)

Overcoming the Limitations

We will now describe a generalized version of WM, called the **Multiplicative Weights Update Algorithm,** in which

- At each time step t, a decision is made **probabilistically** (this is analogous to playing **mixed strategies**)
- There is an **arbitrary finite number of decisions** $i^t \in S_1$ to be made at each time t
  (this is analogous to having n pure strategies)
- The number of outcomes $j^t \in S_2$ at time t is arbitrary **(even infinite!)**
  (corresponds to the **column player** having an arbitrary—even infinite—number of strategies at each t)
- The **loss** $B(i^t, j^t)$ **is arbitrary** (this corresponds to having an arbitrary payoff matrix B in a matrix game, not just a zero-one matrix)

We will assume that $-\rho \le B(i^t, j^t) \le +\rho$ for all t for some $\rho > 0$ which we will refer to as the **width** parameter.

Multiplicative Weights Update Algorithm

Problem solving was always one of Martha's greatest talents...

Multiplicative Weights Update (MWU) Algorithm: Generalized Weighted Majority

1. At time $t = 1$ assign **unit weight** to all experts: $w_i^t = 1, i = 1, \ldots, n$
2. Compute **probabilities** $p_i^t = w_i^t / \sum_j w_j^t$ for all $i = 1, \ldots, n$
3. Follow the advice of expert i **(play strategy i)** with probability p_i^t
4. Observe the outcome j^t and incur **expected loss**

$$B(p^t, j^t) = \sum_{i=1}^{n} p_i^t B(i, j^t)$$

5. **Update the weights** of all experts i in a **multiplicative** way:

$$w_i^{t+1} = \begin{cases} w_i^t (1-\epsilon)^{B(i,j^t)/\rho} & \text{if } B(i,j^t) \geq 0 \\ w_i^t (1+\epsilon)^{-B(i,j^t)/\rho} & \text{if } B(i,j^t) < 0. \end{cases}$$

6. Proceed to time $t \leftarrow t + 1$ and go to step 2

MWU: The Main Result

Theorem (Multiplicative Weights Update)

Let $0 < \epsilon \leq \frac{1}{2}$ be a learning rate. Then the output of the MWU algorithm satisfies the inequality

$$\underbrace{\sum_{t=1}^{T} B(p^t, j^t)}_{\textit{my expected loss}} \leq \frac{\rho \ln n}{\epsilon} + (1+\epsilon) \underbrace{\sum_{t:B(i,j^t)\geq 0} B(i,j^t)}_{\textit{loss of expert } i} + (1-\epsilon) \underbrace{\sum_{t:B(i,j^t)<0} B(i,j^t)}_{\textit{negative loss of expert } i}$$

for any expert i and time T.

Proof.

Similar to the proof of Weighted Majority. □

MWU: A Crucial Consequence

Corollary

Fix any $\delta > 0$ and let $\epsilon = \min\{\frac{\delta}{4\rho}, \frac{1}{2}\}$ and $T = \frac{16\rho^2 \ln n}{\delta^2}$. Then

$$\frac{1}{T}\sum_{t=1}^{T} B(p^t, j^t) \leq \delta + \frac{1}{T}\sum_{t=1}^{T} B(i, j^t) \quad \textit{for all} \quad i. \qquad (\#0)$$

This statement says how many **plays** (time steps T), following the **MWU method**, are needed for the **average expected loss** of the row player in a **repeated game-play** against a **column player** (nature picking the outcomes), to be **guaranteed to be within** a small additive constant δ of the **average loss** of his **best pure response** (i minimizing the right-hand side) to the **mixed strategy** of the column player formed from the **actual observed frequencies** of his pure strategies (outcomes).