Game Theory & Learning

Informal Notes—Not to be distributed

Bhubaneswar Mishra

Courant Institute of Mathematical Sciences

Preface

In the spring of 1998, a small group of computer science colleagues, students and I started writing notes that could be used in the context of our research in computational economy, evolving around our work on *CAFE*. The group consisted of Rohit Parikh, Ron Even, Amy Greenwald, Gideon Berger, Toto Paxia and few others. At present, these notes are intended for the consumption of only this group.

January 1 1998
251 Mercer Street, New York.

B. Mishra
mishra@nyu.edu

Contents

Chapter 1

Introduction

1.1 Stag Hunt Problem

(With Two Players)

Stag Hunt Problem

	Stag	Hare
Stag	2,2	0,1
Hare	1,0	1,1

1. If both row-player and column-player hunt stag, since a stag is worth 4 "utils", they each get 2 "utils."

2. If both row-player and column-player hunt hares, since a hare is worth 1 "util", they each get 1 "util."

3. If row-player hunts hare, while column-player hunts stag (and hence fails to hunt any thing), then the row-player gets 1 "util" and the column-player gets 0 "util."

4. The other case is symmetric.

Note that if row-player is risk aversive, he will choose to hunt hare and thus guarantee that he gets 1 "util" independent of the choice column-player makes. Thus he will maximize the minimum utility under the two possible pure strategies ("hunt stag" with a minimum utility of 0 if the opponent hunts hare vs. "hunt hare" with a minimum utility of 1 regardless of what the opponent chooses to play) and choose to hunt hare. By symmetry, it is seen that in fact both players will choose to hunt hares.

Is this the truly optimal strategy?

Quoting Rousseau (*Discourse on the origin and Basis of Equality among Men*):

> "If a group of hunters set out to take a stag, they are fully aware that they would all have to remain faithfully at their posts in order to succeed; but if a hare happens to pass near one of them, there can be no doubt that if he pursued it without qualm, and that once he had caught his prey, he cared very little whether or not he had made his companions miss theirs."

Changing the discussion slightly, suppose that column-player will play a mixed strategy by playing "hunt stag" with some probability (say, y) and by playing the other strategy ("hunt hare") with probability $(1-y)$. His best choice of these probabilities must be such that row-player is now "indifferent" to the choice of his own strategies. Thus, we have

$$2y + 0(1-y) = 1y + 1(1-y)$$

and $y = 1/2$. Thus one expects both row-player and column-player to play the strategies "hunt stag" and "hunt hare" with equal probabilities.

1.2 Why are these kinds of analysis important to us?

1. Economy

2. Evolutionary Biology
3. Large Scale Distributed Systems
4. Resource Allocation
5. Intelligent Agents

1.3 Prisoners' Dilemma

Prisoners' Dilemma

	C	D
C	0,0	-2,1
D	1,-2	-1,-1

There are two prisoners (row-player and column-player) arrested for a particular crime, but the prosecutor does not have enough evidence to convict them both. He relies one one of them testifying against the other in order to get a conviction and punish the second prisoner by sending him to jail. If both of them testify against the other (defections: "D, D") then they both go to jail for 1 year each, thus getting a "util" of -1. If, on the other hand, both maintain silence (cooperations: "C, C") then they go free with "util" of 0 each. If, on the other hand, row-player testifies (D) and column-player maintains silence (C), then row-player is rewarded with 1 util and column-player is punished with -2 util. The other case is symmetric.

The pay-offs can be made all non-negative by adding 2 utils to each and thus getting a pay-off matrix:

Prisoners' Dilemma (Modified Pay-offs)

	C	D
C	2,2	0,3
D	3,0	1,1

1. For column-player the strategy C is dominated by the strategy D independent of how row-player plays the game. Thus column player must defect.

2. Similarly, for row-player the strategy C is dominated by the strategy D independent of how column-player plays the game. Thus row player must defect.

Hence the equilibrium strategy for the players is to defect even when they could have each gotten better pay-offs by cooperating.

1.4 Second-Price Auction

1. Seller has one indivisible unit of object for sale.

2. There are I potential buyers (bidders) with valuations
$$0 \leq v_1 \leq v_2 \leq v_I.$$
(Consider the case when $I = 2$.)

3. The bidders simultaneously submit bids
$$s_i \in [0, \infty].$$

4. The highest bidder wins the object.

5. But he only pays the second bid $(\max_{j \neq i} s_j)$.

6. His utility is

$$v_i - \max_{j \neq i} s_j.$$

Consider the special case of just two players

$$v_1, v_2 = \text{valuations} \qquad s_1, s_2 = \text{bids}.$$

Pay-offs

$$\begin{aligned} u_1 &\equiv \text{ if } s_1 > s_2 \text{ then } v_1 - s_2 \text{ else } 0. \\ u_2 &\equiv \text{ if } s_2 > s_1 \text{ then } v_2 - s_1 \text{ else } 0. \end{aligned}$$

Let us look at the player 1's choices.

1. <u>Overbidding</u>
   (a) $s_1 \leq s_2$: The payoff is zero and the strategy is weakly dominated.
   (b) $s_2 \leq v_1$: The payoff is $v_1 - s_2$ and the strategy is weakly dominated with respect to bidding $s_1 = v_1$.
   (c) $v_1 < s_2 < s_1$: The payoff is $v_1 - s_2 < 0$ negative and the strategy is strongly dominated.
2. <u>Underbidding</u>
   (a) $s_2 \geq v_1$: The payoff is zero and the strategy is weakly dominated.
   (b) $s_1 \geq s_2$: The payoff is is $v_1 - s_2$ and the strategy is weakly dominated with respect to bidding $s_1 = v_1$.
   (c) $s_1 < s_2 < v_1$: The payoff is zero and the strategy is weakly dominated.

So the best strategy for player 1 is to bid exactly his own valuation ($s_1 = v_1$). And by a symmetric argument, the best strategy for player 2 is also to bid exactly his own valuation ($s_2 = v_2$).

1.5 Two Person Zero-sum Games

We define a *loss matrix* M as follows:

$M(s_i, s_j) = M(i,j) =$ Loss suffered by the row-player for the strategy profile (s_i, s_j).

Rock, Paper & Scissors

	R	P	S
R	1/2	1	0
P	0	1/2	1
S	1	0	1/2

Row-player's goal is to minimize the loss. Assume (without loss of generality) that all the losses are in the range $[0,1]$.

Row-player's expected loss

$$\begin{aligned}
&\sum_{i,j} \sigma_r(s_i)\sigma_c(s_j)M(s_i, s_j)\\
&= \sum_{i,j} \sigma_r(i)M(i,j)\sigma_c(j)\\
&= \sigma_r^{\mathrm{T}} M \sigma_c = M(\sigma_r, \sigma_c).
\end{aligned}$$

$$\begin{aligned}
\sigma_r(s_i) &= \text{Probability that the row player plays } s_i\\
\sigma_c(s_j) &= \text{Probability that the column player plays } s_j
\end{aligned}$$

Similarly,

$$M(\sigma_r, j) = \sum_{i,j} \sigma_r(i) M(i,j) \quad \text{and} \quad M(i, \sigma_c) = \sum_{i,j} \sigma_c(j) M(i,j).$$

Row-player's strategy

$$\min_{\sigma_r} \max_{\sigma_c} M(\sigma_r,\ \sigma_c).$$

A mixed strategy σ_r^* realizing this minimum is called a *minmax* strategy.

Theorem 1.5.1 The MINMAX theorem: von Neumann

$$\min_{\sigma_r} \max_{\sigma_c} M(\sigma_r,\ \sigma_c) = \max_{\sigma_c} \min_{\sigma_r} M(\sigma_r,\ \sigma_c).$$

1.6 Obstacles

1. **Imperfect Information**
   M (pay off) may be unknown.

2. **Computational complexity**
   M is so large that computing a minmax strategy using a linear program is infeasible.

3. **Irrationality**
   Opponent (column-player) may not be truly adversarial.

1.7 Repeated Play (with learning)

M unknown

1. The game is played repeatedly in a sequence of rounds.

2. On round $t = 1, \ldots, T$:

   (a) The learner (row-player) chooses mixed strategy $\sigma_{r,t}$.

   (b) The opponent (column-player) chooses mixed strategy $\sigma_{c,t}$.

   (c) Row-player observes all possible losses

   $$M(i, \sigma_{c,t}) = \sum_{i,j} \sigma_{c,t}(j) M(i, j),$$

   for each row i.

   (d) Row-player suffers loss $M(\sigma_{r,t}\ \sigma_{c,t})$.

Row-player's cumulative expected loss:

$$\sum_{t=1}^{T} M(\sigma_{r,t},\ \sigma_{c,t}).$$

The expected cumulative loss of the best strategy

$$\sum_{t=1}^{T} M(\sigma_r^*,\ \sigma_{c,t}) = \min_{\sigma_r} \sum_{t=1}^{T} M(\sigma_r,\ \sigma_{c,t}).$$

1.8 Learning Algorithm

Parameter β to be chosen. Initially,

$$\begin{aligned} W_1(i) &= 1, \qquad \forall i \\ W_{t+1}(i) &= W_t(i)\beta^{M(i,\sigma_{c,t})} \\ \sigma_{r,t}(i) &= \frac{W_t(i)}{\sum_i W_t(i)}. \end{aligned}$$

1.9 Analysis of Learning Algorithm

1.9.1 Inequality 1

$$\begin{aligned} \sum_i W_{t+1}(i) &= \sum_i W_t(i)\beta^{M(i,\,\sigma_{c,t})} \\ &= \left(\sum_i W_t(i)\right) \cdot \sum_i \sigma_{r,t}\beta^{M(i,\,\sigma_{c,t})} \\ \Rightarrow \quad \frac{\sum_i W_{t+1}(i)}{\sum_i W_t(i)} &= \sum_i \sigma_{r,t}\beta^{M(i,\,\sigma_{c,t})} \\ &\leq \sum_i \sigma_{r,t}(1-(1-\beta)M(i,\ \sigma_{c,t})) \\ &= 1-(1-\beta)M(\sigma_{r,t},\ \sigma_{c,t})). \end{aligned}$$

After telescoping, we get

$$\frac{\sum_i W_{T+1}(i)}{\sum_i W_1(i)} \leq \prod_t (1-(1-\beta)M(\sigma_{r,t},\ \sigma_{c,t}))$$

Hence,

$$\begin{aligned}
\ln\left(\frac{\sum_i W_{T+1}(i)}{n}\right) &\leq \sum_t \ln(1-(1-\beta)M(\sigma_{r,t},\ \sigma_{c,t})) \\
&\leq -(1-\beta)\sum_t M(\sigma_{r,t},\ \sigma_{c,t}).
\end{aligned}$$

1.9.2 Inequality 2

$$\begin{aligned}
\sum_i W_{T+1}(i) \geq W_{T+1}(j) &= \beta^{\sum_t M(j,\ \sigma_{c,t})} \\
&\geq \beta^{\sum_t M(\sigma_r^*,\ \sigma_{c,t})}.
\end{aligned}$$

Hence

$$\ln\left(\frac{\sum_i W_{T+1}(i)}{n}\right) \geq (\ln\beta)\sum_t M(\sigma_r^*,\ \sigma_{c,t})) - \ln n.$$

1.9.3 Final Result

Combining the two inequalities:

$$(1-\beta)\sum_t M(\sigma_{r,t},\ \sigma_{c,t}) \leq \ln n + (\ln 1/\beta)\sum_t M(\sigma_r^*,\ \sigma_{c,t})).$$

and,

$$\begin{aligned}
\sum_t M(\sigma_r^*,\ \sigma_{c,t}) &\leq \sum_t M(\sigma_{r,t},\ \sigma_{c,t}) \\
&\leq \frac{(\ln 1/\beta)}{1-\beta}\sum_t M(\sigma_r^*,\ \sigma_{c,t})) + \frac{\ln n}{1-\beta}.
\end{aligned}$$

Chapter 2

Strategic Form Games

2.1 Games

Games can be categorized in to following two forms as below. We will start here with the first category and postpone the discussion of the second category for later.

1. Strategic Form Games (also called Normal Form Games)

2. Extensive Form Games

2.2 Strategic Form Games

1. Let $\mathcal{I} = \{1, \ldots, I\}$ be a finite set of players, where $I \in \mathbb{N}$ is the number of players.

2. Let $S_i (i \in \mathcal{I})$ be the (finite) set of *pure strategies* available to player $i \in \mathcal{I}$.

3. $$S = S_1 \times S_2 \times \cdots \times S_I$$

   (Cartesian product of the pure strategies) = Set of pure strategy profiles.

Conventions

We write, $s_i \in S_i$ for a pure strategy of player i. We also write, $s = (s_1, s_2, \ldots, s_I) \in S$ for a pure strategy profile.

"$-i$" denotes the player i's "opponents" and refers to all players other than some given player i. Thus, we can write, $S_{-i} = \times_{j \in \mathcal{I}, j \neq i} S_j$

Just as before, $s_{-i} \in S_{-i}$ denotes a pure strategy profile for the opponents of i. Hence,

$$s = (s_i, s_{-i}) \in S,$$

is a pure strategy profile.

$u_i : S \to \mathbb{R}$ = Pay-off function (real-valued function on S) for player i.

$u_i(s)$ = *von Neumann-Morgenstern utility* of player i for each profile $s = (s_1, s_2, \ldots, s_I)$ of pure strategies.

Definition 2.2.1 *A strategic form game is a tuple*

$$(\mathcal{I}, \{S_1, S_2, \ldots, S_I\}, \{u_1, u_2, \ldots, u_I\})$$

consisting of a a set of players, pure strategy spaces *and* pay-off functions.

Definition 2.2.2 *A two-player zero-sum game is a strategic form game with $\mathcal{I} = \{1, 2\}$ such that*

$$\forall_{s \in S} \quad \sum_{i=1}^{2} u_i(s) = 0.$$

Definition 2.2.3 *A mixed strategy set for player i, Σ_i is the set of probability distributions over the pure strategy set S_i*

$$\Sigma_i = \left\{ \sigma_i \colon S_i \to [0,1] \,|\, \sum_i \sigma_i(s_i) = 1 \right\}.$$

The space of mixed strategy profile = $\Sigma = \times_{i \in \mathcal{I}} \Sigma_i$.

As before, we write: $\sigma_i \in \Sigma_i$, and $\sigma = \{\sigma_1, \sigma_2, \ldots, \sigma_I\} \in \Sigma$.

The *support* of a mixed strategy σ_i = The set of pure strategies to which σ_i assigns positive probability.

Player i's pay-off to profile σ is

$$\begin{aligned}
u_i(\sigma) &= \mathbb{E}_{\sigma_i} u_i(\cdot, \sigma_{-i}) \\
u_i(\sigma) &= u_i(\sigma_i, \sigma_{-i}) = \sum_{s_i \in S_i} \sigma_i(s_i) u_i(s_i, \sigma_{-i}) \\
u_i(s_i, \sigma_{-i}) &= \sum_{s_{-i} \in S_{-i}} \sigma_{-i}(s_{-i}) u_i(s_i, s_{-i}) \\
&= \sum_{s_{-i} \in S_{-i}} \left(\prod_{j \neq i} \sigma_j(s_j) \right) u_i(s_i, s_{-i}).
\end{aligned}$$

Hence,

$$\begin{aligned}
u_i(\sigma) &= \sum_{s_i \in S_i} \sum_{s_{-i} \in S_{-i}} \sigma_i(s_i) \left(\prod_{j \neq i} \sigma_j(s_j) \right) u_i(s_i, s_{-i}) \\
&= \sum_{s \in S} \left(\prod_j \sigma_j(s_j) \right) u_i(s).
\end{aligned}$$

2.3 Domination & Nash Equilibrium

Definition 2.3.1 *A pure strategy s_i is* strictly dominated *for player i if*

$$\exists_{\sigma'_i \in \Sigma_i} \forall_{s_{-i} \in S_{-i}} u_i(\sigma'_i, s_{-i}) > u_i(s_i, s_{-i}).$$

A pure strategy s_i is weakly dominated *for player i if*

$$\begin{aligned}
\exists_{\sigma'_i \in \Sigma_i} \Big(& \forall_{s_{-i} \in S_{-i}} u_i(\sigma'_i, s_{-i}) \geq u_i(s_i, s_{-i}) \\
\wedge \;\; & \exists_{s_{-i} \in S_{-i}} u_i(\sigma'_i, s_{-i}) > u_i(s_i, s_{-i}) \Big).
\end{aligned}$$

Definition 2.3.2 Best Response: *The set of best responses for player i to a pure strategy profile $s \in S$ is*

$$BR_i(s) = \left\{ s_i^* \in S_i | \forall_{s_i \in S_i} \; u_i(s_i^*, s_{-i}) \geq u_i(s_i, s_{-i}) \right\}.$$

Let the joint best response set be $BR(s) = \times_i BR_i(s)$.

Definition 2.3.3 Nash Equilibrium: *A pure strategy profile s^* is a* Nash equilibrium *if for all players i,*

$$\forall_{s_i \in S_i}\ u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*).$$

Thus a Nash equilibrium is a strategy profile s^* such that $s^* \in BR(s^*)$.

A Nash equilibrium s^* is strict *if each player has a unique best response to his rivals' strategies: $BR(s^*) = \{s^*\}$.*

$$\forall_{s_i \neq s_i^*} u_i(s_i^*, s_{-i}^*) > u_i(s_i, s_{-i}^*).$$

A mixed strategy profile σ^* is a Nash equilibrium *if for all players i,*

$$\forall_{s_i \in S_i}\ u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(s_i, \sigma_{-i}^*).$$

Remark: Since expected utilities are "linear in the probabilities," if a player uses a non-degenerate mixed strategy in a Nash equilibrium (non-singleton support), he must be indifferent between all pure strategies to which he assigns positive probability. (It suffices to check that no player has a profitable pure-strategy deviation).

2.4 Example

Example

	L	M	R
U	4,3	5,1	6,2
M	2,1	8,4	3,6
D	3,0	9,6	2,8

For column-player, M is dominated by R. Column-player can eliminate M from his strategy space. The pay-off matrix reduces to

New Pay-offs

	L	R
U	4,3	6,2
M	2,1	3,6
D	3,0	2,8

For row-player, M and D are dominated by U. Row-player can eliminate M and D. The new pay-off matrix is

New Pay-offs

	L	R
U	4,3	6,2

Next, column-player eliminates R as it is dominated by U and reduces the pay-off matrix to

New Pay-offs

	L
U	4,3

Note that

$$BR_r(U,L) = U, \quad \& \quad BR_c(U,L) = L, \quad \& \quad BR(U,L) = (U,L).$$

(U, L) is a strict Nash equilibrium.
Remark: *Mixed Strategy* (Not a Nash equilibrium.)

$$\sigma_r = (1/3, 1/3, 1/3) \quad \& \quad \sigma_c = (0, 1/2, 1/2) \quad \& \quad \sigma = (\sigma_r, \sigma_c).$$

Thus

$$\begin{aligned} u_r(\sigma_r, \sigma_c) &= \sum_s (\prod_j \sigma_j(s_j)) u_r(s) \\ &= (1/3 \times 0)4 + (1/3 \times 1/2)5 + (1/3 \times 1/2)6 \\ &\quad + (1/3 \times 0)2 + (1/3 \times 1/2)8 + (1/3 \times 1/2)3 \\ &\quad + (1/3 \times 0)3 + (1/3 \times 1/2)9 + (1/3 \times 1/2)2 \\ &= 5\frac{1}{2}, \end{aligned}$$

and

$$\begin{aligned} u_c(\sigma_r, \sigma_c) &= \sum_s (\prod_j \sigma_j(s_j)) u_c(s) \\ &= (1/3 \times 0)3 + (1/3 \times 1/2)1 + (1/3 \times 1/2)2 \\ &\quad + (1/3 \times 0)1 + (1/3 \times 1/2)4 + (1/3 \times 1/2)6 \\ &\quad + (1/3 \times 0)0 + (1/3 \times 1/2)6 + (1/3 \times 1/2)8 \\ &= 4\frac{1}{2}, \end{aligned}$$

Thus this mixed strategy leads to a much better pay-off in comparison to the pure strategy Nash equilibrium.

A pure strategy may be strictly dominated by a mixed strategy, even if it is not strictly dominated by any pure strategy.

Example

	L	R
U	2,0	-1,0
M	0,0	0,0
D	-1,0	2,0

For row-player M is not dominated by either U or D. But M is dominated by a mixed strategy $\sigma_r = (1/2, 0, 1/2)$ (payoff: $u_r(\sigma) = (1/2, 1/2)$.

Going back to the "Prisoners' Dilemma" game, note that its Nash equilibrium is in fact (D, D) [both players defect].

$$\begin{aligned}
BR_r(C,C) = BR_r(C,D) = BR_r(D,C) = BR_r(D,D) &= D,\\
BR_c(C,C) = BR_c(C,D) = BR_c(D,C) = BR_c(D,D) &= D,\\
BR(C,C) = BR(C,D) = BR(D,C) = BR(D,D) &= (D,D).
\end{aligned}$$

2.4.1 Matching Pennies

Matching Pennies

	H	T
H	1,-1	-1,1
T	-1,1	1,-1

There are two players: "Matcher" (row-player) and "Mismatcher" (column-player). Matcher and Mismatcher both have two strategies: "call head" (H) and "call tail" (T). Matcher wins 1 util if both players call the same [(H,H) or (T,T)] and mismatcher wins 1 util if the players call differently [(H,T) or (T,H)]. It is easy to see that this game has no Nash equilibrium pure strategy. However it *does have* a Nash equilibrium mixed strategy:

$$\sigma_r = (1/2, 1/2) \quad \& \quad \sigma_c = (1/2, 1/2).$$

The pay-offs are

$$\begin{aligned}
u_r(\sigma) &= (1/2 \times 1/2)1 + (1/2 \times 1/2)(-1)\\
&\quad + (1/2 \times 1/2)(-1) + (1/2 \times 1/2)1 = 0\\
u_c(\sigma) &= (1/2 \times 1/2)(-1) + (1/2 \times 1/2)1\\
&\quad + (1/2 \times 1/2)1 + (1/2 \times 1/2)(-1) = 0.
\end{aligned}$$

2.5 Key Ingredients for Nash Equilibrium

1. *Introspection* (Fictitious play)
2. *Deduction/Rationality*
3. *Knowledge of Opponents Pay-offs*
4. *Common Knowledge*

2.6 Revisiting On-line Learning

2.6.1 Convergence

Note that in the earlier discussion of the on-line learning strategy, we noted that the on-line learning algorithm is competitive [with a competitive factor of $(\ln 1/\beta)/(1-\beta) \approx 1 + (1-\beta)/2 + (1-\beta)^2/3 + \cdots$, for small $(1-\beta)$] for any sufficiently large time interval $[0, T]$. But it is also fairly easy to note that the probabilities that the row-player chooses do not necessarily converge to the best mixed strategy. Namely,

$$W_T(i) = \beta^{\sum_t M(i, \sigma_{c,t})} \quad \& \quad \sigma_{r,T}(i) = \frac{W_T(i)}{\sum_i W_T(i)}.$$

We have not explicitly shown that $\lim_{T\to\infty} \sigma_{r,T}$ converges in distribution to σ_r^*. Does the computed distribution converge to anything? In the absence of any convergence property, one may justifiably question how the algorithm can be interpreted as learning a strategy.

2.6.2 Irrationality

Let us look at the "Matching Pennies" problem again:

<u>Matching Pennies</u>

	H	T
H	1,-1	-1,1
T	-1,1	1,-1

Suppose the column-player chooses a mixed strategy at time t such that $\sigma_{c,t}(H) > 1/2$ [and $\sigma_{c,t}(T) = 1 - \sigma_{c,t}(H) < 1/2$] then for the row-player, the best response is $BR_{r,t}(\sigma_t) = H$ and is unique. By a similar reasoning, if $\sigma_{c,t}(H) < 1/2$ [and $\sigma_{c,t}(T) > 1/2$], then for the row-player, the best response is $BR_{r,t}(\sigma_t) = T$. Thus, if the rival deviates from his Nash equilibrium mixed strategy $\sigma_{c,t} = (1/2, 1/2)$, then row-player's (rational) best response is always a pure strategy H or T. Thus, if row-player had a convergent (rational) mixed strategy, then depending on $\lim_{T\to\infty}\{\sigma_{c,t}\}_0^T$, the row player must converge to one of the following three (conventional) strategies:

1. RANDOM$(1/2, 1/2)$ (the Nash equilibrium mixed strategy),
2. H^* (always H), or
3. T^* (always T).

Anything else would make the row-player irrational. Thus, a player playing the on-line learning algorithm must be almost always irrational!

2.6.3 A Meta-Theorem of Foster & Young

Definition 2.6.1 *An infinite sequence $\sigma_{c,t}$ is* almost constant, *if there exists a σ_c such that $\sigma_{c,t} = \sigma_c$ almost always (a.a.). That is*

$$\lim_{T\to\infty} \frac{|\{t \le T : \sigma_{c,t} \ne \sigma_c\}|}{T} = 0.$$

If $\sigma_{c,t}$ is not almost constant then

$$\forall_{\sigma_c = \text{ const}} \quad \sigma_{c,t} \ne \sigma_c \quad \text{infinitely often (i.o.).}$$

Consider an n-player game with a strategy space $S_1 \times S_2 \times \cdots \times S_n = S$ and with the utility functions $u_i : S \to \mathbb{R}$. All actions are publicly observed. Let Σ_i = the set of probability distributions over S_i. Let $\Sigma = \times_i \Sigma_i$ be the product set of mixture. Before every round of the game, a state can be described by a family of probability distributions

$$\{(\sigma_i, \sigma_{i,j})\}_{i \neq j}.$$

$$\begin{aligned} \sigma_i \in \Sigma_i &= \text{Player } i\text{'s mixed strategy,} \\ \sigma_{i,j} \in \Sigma_j &= \text{Player } i\text{'s \underline{belief} about player } j\text{' mixed strategy.} \end{aligned}$$

Definition 2.6.2 **Rationality**: *Each player chooses only best replies given his beliefs:*

$$\forall_{i \neq j}\ \sigma_i(s_i) > 0 \Rightarrow s_i \in BR_i(\sigma_{i,j}).$$

Definition 2.6.3 **Learning**: *Player i has its own deterministic learning process $\{f_i, f_{i,j}\}$ which it uses in determining its strategy and its beliefs. In particular, let h_t = all publicly available information up to time t. Then, player i chooses its strategy and beliefs as follows:*

$$\begin{aligned} f_i &: h_{t-1} \mapsto \sigma_{i,t} \\ f_{ij} &: h_{t-1} \mapsto \sigma_{ij,t}. \end{aligned}$$

The learning process is informationally independent *if $\sigma_{ij,t} = f_{ij}(h_{t-1})$ do not depend on any extraneous information.*

Definition 2.6.4 **Convergence**: *The beliefs are said to* converge *along a learning path $\{h_t, \sigma_{i,t}, \sigma_{ij,t}\}_0^\infty$ if*

$$\forall_{i \neq j}\ \exists_{\sigma_{ij} \in \Sigma_j}\ \lim_{t \to \infty}\ \sigma_{ij,t} = \sigma_{ij}.$$

The strategies are said to converge *along a learning path $\{h_t, \sigma_{i,t}, \sigma_{ij,t}\}_0^\infty$ if*

$$\forall_i\ \exists_{\sigma_i \in \Sigma_i}\ \lim_{t \to \infty}\ \sigma_{i,t} = \sigma_i.$$

The beliefs are said to be predictive *along a learning path if*

$$\forall_{i \neq j} \lim_{t \to \infty} \sigma_{ij,t} = \sigma_{i,t},$$

and they are strongly predictive *if in addition both the beliefs and strategies converge.*

Theorem 2.6.1 *Consider a finite 2-person game (players: row-player and column-player) with a strict (thus, unique) Nash equilibrium $\sigma^* = (\sigma_r^*, \sigma_c^*)$ which has full support on $S_r \times S_c$. Let $\{(f_r, f_{rc}), (f_c, f_{cr})\}$ be a DRIP learning process (D = Deterministic, R = Rational, I = Informationally independent and P = Predictive).*

On any learning path $(h_t, (\sigma_{r,t}, \sigma_{rc,t}), (\sigma_{c,t}, \sigma_{cr,t}))$, if the beliefs are not *almost constant with value σ^* then the beliefs do not converge.*

Proof:

Assume to the contrary: then $\sigma_{rc,t} \neq \sigma_c^*$ i.o. Then, infinitely often, $\sigma_{rc,t}$ does not have full support and

$$\exists_{s_{r,t} \in S_r} \; s_{r,t} \notin BR_r(\sigma_{rc,t}),$$

and by the finiteness of the strategies S_r:

$$\exists_{s_r \in S_r} \; s_r \notin BR_r(\sigma_{rc,t}) \;\; \text{i.o.}$$

By rationality of row-player,

$$\exists_{s_r \in S_r} \; \sigma_{r,t}(s_r) = 0 \text{ i.o.} \quad \& \quad \exists_{s_r \in S_r} \lim_{t \to \infty} \sigma_{r,t}(s_r) = 0.$$

By a similar argument,

$$\exists_{s_c \in S_c} \lim_{t \to \infty} \sigma_{c,t}(s_c) = 0.$$

Since the learning is assumed to be predictive, we get

$$\lim_{t \to \infty} \sigma_{cr,t}(s_r) = 0 \quad \& \quad \sigma_{rc,t}(s_c) = 0.$$

Thus, if the beliefs converge (say, to (σ_r, σ_c)) then the beliefs (and also, strategies—by predictivity) converge to some strategies other than the unique Nash equilibrium (as it is unique with full support). Hence one of the following two holds at the limit:

$$\exists_{t_r \in S_r \setminus \{s_r\}} \sigma_r(t_r) > 0 \quad \text{and} \quad t_r \notin BR_r(\sigma_c)$$
$$\text{or}$$
$$\exists_{t_c \in S_c \setminus \{s_c\}} \sigma_c(t_c) > 0 \quad \text{and} \quad t_c \notin BR_c(\sigma_r).$$

But, depending on which equation holds true, we shall conclude that either row-player or column-player (or both) must be irrational, a contradiction.

Chapter 3

Nash Equilibrium

3.1 Nash Equilibrium

3.1.1 Fixed Point Theorems

Definition 3.1.1 *A point $x \in K$ is a* fixed point *of an injective function $f: K \to K$, if*

$$x = f(x).$$

Definition 3.1.2 *A point $x \in K$ is a* fixed point *of a mapping $\Psi: K \to 2^K$, if*

$$x \in \Psi(x).$$

Theorem 3.1.1 Brouwer's Fixed Point Theorem: *If $f : K \to K$ is a* continuous function *from a nonempty, compact, convex subset K of a finite dimensional TVS (topological vector space) into itself, then f has a fixed point, i.e.,*

$$\exists_{x \in K}\ x = f(x).$$

Theorem 3.1.2 Kakutani's Fixed Point Theorem: *If $\Psi : K \to 2^K$ is a convex-valued, uhc (upper hemi-continuous) map from a nonempty, compact, convex subset K of a finite dimensional TVS to the nonempty subsets of K, then Ψ has a fixed point, i.e.,*

$$\exists_{x \in K}\ x \in \Psi(x).$$

Definition 3.1.3 *Topological Vector Space: L = vector space with a T_1 topology*

$$(\forall_{x \neq y \in L}\ \exists_{G_x = \text{ open set }}\ x \in G_x \wedge y \notin G_x)$$

which admits continuous vector space operations.

Example: $\mathbb{R}^n$ with standard Euclidean topology. (Only instance of a finite dimensional TVS.)

Theorem 3.1.3 Existence of a Mixed Strategy Equilibrium *(Nash 1950). Every finite strategic-form game has a mixed-strategy equilibrium.*

Proof: Player i's *reaction correspondence*, Ψ_i, maps each strategy profile σ to the set of mixed strategies that maximize player i's pay-offs when his rivals play σ_{-i}:

$$\Psi_i(\sigma) = \left\{\sigma'_i \mid \forall_{s_i \in S_i}\ u_i(\sigma'_i, \sigma_{-i}) \geq u_i(s_i, \sigma_{-i})\right\}.$$

Thus,

$$\Psi_i : \Sigma \to 2^{\Sigma_i}.$$

Define

$$\Psi : \Sigma \to 2^{\Sigma} : \sigma \mapsto \times_i \Psi_i(\sigma).$$

Thus this *correspondence map* is the Cartesian product of Ψ_i's.

A fixed point of Ψ (if exists) is a σ^* such that

$$\sigma^* \in \Psi(\sigma^*).$$

Note that

$$\forall_{s_i \in S_i}\ u_i(\sigma^*_i, \sigma^*_{-i}) \geq u_i(s_i, \sigma^*_{-i}),$$

by definition. Thus a fixed point of Ψ provides a mixed strategy equilibrium σ^*.

Claims:

1. Σ = Nonempty, compact and convex subset of a TVS.

   $\Sigma_i = \Delta_{|S_i|-1} = |S_i| - 1$ dimensional simplex, since

   $$\Sigma_i = \left\{(\sigma_{i,1}, \ldots, \sigma_{i,|S_i|}) \mid \sigma_{i,j} \geq 0, \sum_j \sigma_{i,j} = 1\right\}.$$

   Rest follows since $\Sigma = \times_i \Sigma_i$.

2. $u_i =$ Linear Function.

$$\begin{aligned} &\forall_{0<\lambda<1}\, u_i(\lambda\sigma_i' + (1-\lambda)\sigma_i'', \sigma_{-i}) \\ &\quad = \lambda u_i(\sigma_i', \sigma_{-i}) + (1-\lambda) u_i(\sigma_i'', \sigma_{-i}). \end{aligned}$$

   Hence u_i is a continuous function in his own mixed strategy. Since Σ is compact, u_i attains maxima in Σ.

$$\forall_{\sigma\in\Sigma}\, \Psi(\sigma) \neq \emptyset.$$

3.

$$\forall_{\sigma\in\Sigma}\, \Psi(\sigma) = \text{ convex.}$$

   Let σ_i', $\sigma_i'' \in \Psi(\sigma)$. By definition,

$$\begin{aligned} &\forall_{s_i\in S_i}\, (u_i(\sigma_i', \sigma_{-i}) \geq u_i(s_i, \sigma_{-i})) \\ &\quad \wedge \quad (u_i(\sigma_i'', \sigma_{-i}) \geq u_i(s_i, \sigma_{-i})). \end{aligned}$$

   Hence

$$\forall_{0<\lambda<1}\, \forall_{s_i\in S_i}\, u_i(\lambda\sigma_i' + (1-\lambda)\sigma_i'', \sigma_{-i}) \geq u_i(s_i, \sigma_{-i}),$$

   and

$$\forall_{0<\lambda<1}\, \lambda\sigma_i' + (1-\lambda)\sigma_i'' \in \Psi_i(\sigma).$$

4. $\Psi =$ uhc. Consider a sequence

$$\left\{ (\sigma^n, \hat{\sigma}^n) \mid \hat{\sigma}^n \in \Psi(\sigma^n) \right\}_n.$$

   We wish to show that

$$\text{If} \quad \lim_{n\to\infty} (\sigma^n, \hat{\sigma}^n) = (\sigma, \hat{\sigma}) \quad \text{then} \quad \hat{\sigma} \in \Psi(\sigma).$$

   Suppose Not! Then

$$\forall_n\, \hat{\sigma}^n \in \Psi(\sigma^n),$$

   but

$$\hat{\sigma} \notin \Psi(\sigma) \;\Rightarrow\; \hat{\sigma}_i \notin \Psi_i(\sigma).$$

Thus,

$$\exists_{\epsilon>0}\ \exists_{\sigma_i'\in\Sigma_i}\ u_i(\sigma_i',\sigma_{-i}) > u_i(\hat{\sigma}_i,\sigma_{-i}) + 3\epsilon.$$

Since $u_i =$ continuous, there is a sufficiently large N such that

$$\begin{aligned} u_i(\sigma_i',\sigma_{-i}^N) &> u_i(\sigma_i',\sigma_{-i}) - \epsilon \\ &> u_i(\hat{\sigma}_i,\sigma_{-i}) + 2\epsilon \\ &> u_i(\hat{\sigma}_i^N,\sigma_{-i}^N) + \epsilon. \end{aligned}$$

Thus, $\hat{\sigma}_i^N \notin \Psi(\sigma^N)$, a contradiction.

Thus we conclude that $\Psi : \Sigma \to 2^{\Sigma}$ is a convex valued, uhc map from a nonempty, compact, convex subset Σ of finite dimensional TVS to nonempty subsets of Σ. Thus by Kakutani's fixed point theorem

$$\exists_{\sigma^*\in\Sigma}\ \sigma^* \in \Psi(\sigma^*),$$

and σ^* is a mixed strategy Nash equilibrium.

Chapter 4

Beyond Nash: Domination, Rationalization and Correlation

4.1 Beyond Nash

We have seen that it is impossible to "learn" a Nash equilibrium if we insist on DRIP conditions. A resolution to this dilemma can involve one or more of the following approaches:

1. Explore simpler requirements than Nash equilibria: e.g., undominated sets, rationalizable sets and correlated equilibria. (The first two correspond to minmax and maxmin requirements. The last one requires some side information and may make the system informationally dependent.)

2. Requirement of predictivity may need to be abandoned.

3. Requirement of rationality may need to be abandoned.

4.1.1 Correlated Equilibrium

This concept extends the Nash concept by supposing that the players can build a "correlated device" that sends each of the players a private signal before they choose their strategy.

Main Ingredients: *Predictions* using only the assumption that the structure of the game (i.e., the strategy spaces and payoffs, S_i's and u_i's) and the rationality of the players are common knowledge.

4.2 Iterated Strict Dominance and Rationalizability

Definition 4.2.1 Iterated Strict Dominance: *Let*

$$S_i^0 = S_i \quad \textit{and} \quad \Sigma_i^0 = \Sigma_i$$

Let for all $n > 0$

$$S_i^n = \left\{ s_i \in S_i^{n-1} \mid \forall_{\sigma_i' \in \Sigma_i^{n-1}} \exists_{s_{-i} \in S_{-i}^{n-1}} u_i(s_i, s_{-i}) \geq u_i(\sigma_i', s_{-i}) \right\},$$

(Thus s_i dominates all the mixed strategies for some strategy profile of the rivals) and define

$$\Sigma_i^n = \left\{ \sigma_i \in \Sigma_i \mid \sigma_i(s_i) > 0 \Rightarrow s_i \in S_i^n \right\}.$$

Let

$$S_i^\infty = \bigcap_{n=0}^{\infty} S_i^n$$

be the set of player i's pure strategies that survive iterated deletion of strictly dominated strategies.

Let

$$\Sigma_i^\infty = \left\{ \sigma_i = \text{ mixed strategy } \mid \right.$$
$$\left. \forall_{\sigma_i' \in \Sigma_i} \exists_{s_{-i} \in S_{-i}^\infty} u_i(\sigma_i, s_{-i}) \geq u_i(\sigma_i', s_{-i}) \right\}$$

be the set of player i's mixed strategies that survive iterated deletion of strictly dominated strategies.

Example:

	L	R
U	1,3	-2,0
M	-2,0	1,3
D	0,1	0,1

Note that

$$S_r^0 = \{U, M, D\} \quad \& \quad \Sigma_r^0 = \{\sigma \text{ (with full support) }\}.$$

Similarly,

$$S_c^0 = \{L, R\} \quad \& \quad \Sigma_c^0 = \{\sigma \text{ (with full support) }\}.$$

Also note that

$$S_r^\infty = \cdots = S_r^2 = S_r^1 = S_r^0, \quad \& \quad S_c^\infty = \cdots = S_c^2 = S_c^1 = S_c^0.$$

Note, however, that for all values $p \in (1/3, 2/3)$ the mixed strategy $\sigma_r = (p, 1-p, 0)$ is dominated by D. Thus,

$$\Sigma_r^\infty \subset \Sigma_r^0.$$

4.2.1 Some Properties of Undominated Sets

$$S^\infty = S_1^\infty \times S_2^\infty \times \cdots \times S_I^\infty, \quad \& \quad \Sigma^\infty = \Sigma_1^\infty \times \Sigma_2^\infty \times \cdots \times \Sigma_I^\infty.$$

1. The final surviving strategy spaces are *independent of the elimination order*.

2. A strategy is strictly dominated against all pure strategies of the rivals if and only if it is dominated against all of their

strategies. Thus, the following is an equivalent definition of the undominated sets:

$$S_i^0 = S_i \quad \text{and} \quad \Sigma_i^0 = \Sigma_i$$

$$\begin{aligned}
S_i^n &= \Big\{ s_i \in S_i^{n-1} \mid \\
&\qquad \forall_{\sigma_i' \in \Sigma_i^{n-1}} \exists_{s_{-i} \in S_{-i}^{n-1}}\ u_i(s_i, s_{-i}) \geq u_i(\sigma_i', s_{-i}) \Big\}. \\
\Sigma_i^n &= \Big\{ \sigma_i \in \Sigma_i^{n-1} \mid \\
&\qquad \forall_{\sigma_i' \in \Sigma_i^{n-1}} \exists_{s_{-i} \in S_{-i}^{n-1}}\ u_i(\sigma_i, s_{-i}) \geq u_i(\sigma_i', s_{-i}) \Big\}.
\end{aligned}$$

$$S_i^\infty = \bigcap_{n=0}^{\infty} S_i^n, \quad \& \quad \Sigma_i^\infty = \bigcap_{n=0}^{\infty} \Sigma_i^n.$$

Definition 4.2.2 *A game is solvable by iterated (strict) dominance, if for each player i, S_i^∞ is a singleton, i.e., $S_i^\infty = \{s_i^*\}$. In this case, the strategy profile $(s_1^*,\ s_2^*,\ \ldots,\ s_I^*)$ is a (unique) Nash equilibrium.*

Proof: Suppose that it is not a Nash equilibrium: That is for some i

$$s_i^* \notin BR_i(s_{-i}^*)$$

Thus

$$\exists_{s_i \in S_i}\ u_i(s_i, s_{-i}^*) > u_i(s_i^*, s_{-i}^*).$$

But suppose s_i was eliminated in round n: Then

$$\exists_{s_i' \in S_i^{n-1}} \forall_{s_{-i} \in S_{-i}^{n-1}}\ u_i(s_i', s_{-i}) > u_i(s_i, s_{-i}).$$

Since $s_{-i}^* \in S_{-i}^\infty$, we have $u_i(s_i', s_{-i}^*) > u_i(s_i, s_{-i}^*)$. Repeating in this fashion we get a sequence of inequalities:

$$u_i(s_i^*, s_{-i}^*) > \cdots > u_i(s_i'', s_{-i}^*) > u_i(s_i', s_{-i}^*) > u_i(s_i, s_{-i}^*),$$

resulting in a contradiction.

4.3 Rationalizability

This notion is due to Bernheim(1984), Pearce(1984) and Aumann(1987) and provides a complementary approach to iterated strict dominance. This approach tries to answer the following question:

> "*What are all the strategies that a rational player can play?* "

Rational player will only play those strategies that are best responses to some beliefs he has about his rivals' strategies.

Definition 4.3.1 (Rationalizable Strategies) *Let*

$$\tilde{\Sigma}_i^0 = \Sigma_i.$$

For $n > 0$, let

$$\tilde{\Sigma}_i^n = \Big\{ \sigma_i \in \tilde{\Sigma}_i^{n-1} \mid \exists_{\sigma_{-i} \in \times_{j \neq i} Conv(\tilde{\Sigma}_j^{n-1})} \forall_{\sigma_i' \in \tilde{\Sigma}_i^{n-1}} \; u_i(\sigma_i, \sigma_{-i}) \geq u_i(\sigma_i', \sigma_{-i}) \Big\}.$$

The rationalizable strategies for player i are

$$R_i = \bigcap_{n=0}^{\infty} \tilde{\Sigma}_i^n$$

A strategy profile σ is rationalizable if σ_i is rationalizable for each player i. Let $\sigma^* = (\sigma_1^*, \sigma_2^*, \ldots, \sigma_I^*)$ be a Nash equilibrium. Note first, $\sigma_i^* \in \tilde{\Sigma}_i^0$, for all i. Next assume that $\sigma^* \in \times_i \tilde{\Sigma}_i^{n-1}$. Thus $\sigma_i^* \in \tilde{\Sigma}_i^{n-1}$, and $\sigma_{-i}^* \in \times_{j \neq i} \tilde{\Sigma}_j^{n-1}$. Hence,

$$\forall_{\sigma_i' \in \Sigma_i} \; u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(\sigma_i', \sigma_{-i}^*) \quad \Rightarrow \quad \sigma_i^* \in \tilde{\Sigma}_i^n.$$

Thus, $\sigma^* \in R = \times_i R_i$.

Hence,

Theorem 4.3.1 *Every Nash equilibrium is rationalizable.*

Theorem 4.3.2 *(Bernheim/Pearce (1984))*
The set of rationalizable strategies is nonempty and contains at least one pure strategy for each player. Further, each $\sigma_i \in R_i$ is (in Σ_i) a best response to an element of $\times_{j \neq i} Conv(R_j)$.

Comparing the constructions of undominated strategies with rationalizable strategies, we note that

$$\Sigma_i^0 = \Sigma_i, \quad \text{and} \quad \tilde{\Sigma}_i^0 = \Sigma_i.$$

In the nth iteration, the undominated strategies are constructed as

$$\Sigma_i^n = \Big\{ \sigma_i \in \Sigma_i^{n-1} \mid \\ \forall_{\sigma_i' \in \Sigma_i^{n-1}} \exists_{\sigma_{-i} \in \times_{j \neq i} \mathrm{Conv}(\Sigma_j^{n-1})}\, u_i(\sigma_i, \sigma_{-i}) \geq u_i(\sigma_i', \sigma_{-i}) \Big\},$$

where as rationalizable strategies are constructed as

$$\tilde{\Sigma}_i^n = \Big\{ \sigma_i \in \tilde{\Sigma}_i^{n-1} \mid \\ \exists_{\sigma_{-i} \in \times_{j \neq i} \mathrm{Conv}(\tilde{\Sigma}_j^{n-1})} \forall_{\sigma_i' \in \tilde{\Sigma}_i^{n-1}}\, u_i(\sigma_i, \sigma_{-i}) \geq u_i(\sigma_i', \sigma_{-i}) \Big\}.$$

Finally,

$$\Sigma_i^\infty = \bigcap_{n=0}^{\infty} \Sigma_i^n, \Sigma^\infty = \times_i \Sigma_i^\infty, \quad \text{and} \quad R_i = \bigcap_{n=0}^{\infty} \tilde{\Sigma}_i^n, R = \times_i R_i.$$

A direct examination of these constructions reveals that $\tilde{\Sigma}_i^n \subseteq \Sigma_i^n$ and hence, $R \subseteq \Sigma^\infty$. Also, note that the undominated strategies are computing the minmax values where as the rationalizable strategies compute maxmin values.

4.4 Correlated Equilibrium

Aumann's Example

	L	R
U	5,1	0,0
D	4,4	1,5

There are 3 Nash equilibria:

- A pure strategy: (U, L) $\mapsto$ Pay-off = 5,1,
- A pure strategy: (D, R) $\mapsto$ Pay-off = 1,5, and
- A mixed strategy: $((1/2, 1/2), (1/2, 1/2))$ $\mapsto$ Pay-off = (2.5, 2.5).

Suppose that there is a publicly observable random variable with $Pr(H) = Pr(T) = 1/2$. Let the players play (U, L) if the outcome is H, and (D, R) if the outcome is T. Then the pay-off is (3, 3).

By using publicly observable random variables, the players can obtain any pay-off vector in the convex hull of the set of Nash equilibria pay-offs.

Players can improve (without any prior contracts) if they can build a device that sends different but correlated signals to each of them.

4.4.1 Formal Definitions

- "Expanded Games" with a correlating device.
- Nash equilibrium for the expanded game.

Definition 4.4.1 *Correlating device is a triple*

$$(\Omega, \{H_i\}_{\mathcal{I}}, p)$$

- Ω = *a (finite) state space corresponding to the outcomes of the device.*

- $p =$ *probability measure on the state space* Ω
- $H_i =$ *Information Partition for player* i*.*

  Assigns an $h_i(\omega)$ *to each* $\omega \in \Omega$ *such that* $\omega \in h_i(\omega)$*.*

$$h_i \;:\; \Omega \to H_i \;:\; \omega \mapsto h_i(\omega).$$

Player i*'s posterior belief about* Ω *are given by Bayes' law:*

$$\forall_{\omega \in h_i}\; p(\omega|h_i) = \frac{p(\omega)}{p(h_i)}.$$

4.4.2 Pure Strategies for the Expanded Game

Given a correlating device $(\Omega, \{H_i\}, p)$, we can define strategies for the expanded game as follows: Consider a map

$$\tau_i \;:\; \Omega \to S_i \;:\; \omega \mapsto \tau_i(\omega),$$

such that $\tau_i(\omega) = \tau_i(\omega')$, if $\omega' \in h_i(\omega)$.

The strategies are *adapted* to the information structure.

Definition 4.4.2 DEF(1) *A* correlated equilibrium τ *relative to information structure* $(\Omega, \{H_i\}, p)$ *is a Nash equilibrium in strategies that are adapted to information structure. That is,* $(\tau_1, \tau_2, \ldots, \tau_I)$ *is a correlated equilibrium if*

$$\forall_i\, \forall_{\tilde{\tau}_i} \sum_{\omega \in \Omega} p(\omega) u_i(\tau_i(\omega), \tau_{-i}(\omega)) \geq \sum_{\omega \in \Omega} p(\omega) u_i(\tilde{\tau}_i(\omega), \tau_{-i}(\omega)).$$

Using the Bayes' rule, an equivalent condition would be:

$$\begin{aligned}
&\forall_i\, \forall_{h_i \in H_i, p(h_i) > 0}\, \forall_{s_i \in S_i} \\
&\qquad \sum_{\omega | h_i(\omega) = h_i} p(\omega|h_i) u_i(\tau_i(\omega), \tau_{-i}(\omega)) \\
&\geq \sum_{\omega | h_i(\omega) = h_i} p(\omega|h_i) u_i(s_i, \tau_{-i}(\omega)).
\end{aligned}$$

4.4.3 Correlated Equilibrium and Universal Device

"Universal Device" that signals each player how that player should play.

Definition 4.4.3 DEF(2) *A* correlated equilibrium *is any probability distribution $p(.)$ over the pure strategies $S_1 \times S_2 \times \cdots \times S_I$ such that, for every player i, and every function $d(i) : S_i \to S_i$*

$$\sum_{s \in S} p(s) u_i(s_i, s_{-i}) \;\geq\; \sum_{s \in S} p(s) u_i(d(s_i), s_{-i}).$$

Using the Bayes' rule, an equivalent condition would be:

$$\begin{aligned} &\forall_i \; \forall_{s_i \in S_i, p(s_i) > 0} \; \forall_{s_i' \in S_i} \\ &\qquad \sum_{s_{-i} \in S_{-i}} p(s_{-i}|s_i) u_i(s_i, s_{-i}) \\ &\geq \quad \sum_{s_{-i} \in S_{-i}} p(s_{-i}|s_i) u_i(s_i', s_{-i}). \end{aligned}$$

Equivalence of correlated equilibria under Def(1) and Def(2):
Claim:
Def(1) $\Leftarrow$ Def(2):
Choose $\Omega = S$. $h_i(s) = \{s' | s_i' = s_i\}$. Leave $p(s)$ unchanged.

Claim:
Def(1) $\Rightarrow$ Def(2):
Let τ be an equilibrium w.r.t. $(\Omega, \{H_i\}, \tilde{p})$. Define

$$p(s) = \sum \{\tilde{p}(\omega) | \tau_1(\omega) = s_1, \ldots, \tau_I(\omega) = s_I, \omega \in \Omega\}.$$

Let

$$J_i(s_i) = \{\omega | \tau_i(\omega) = s_i\}.$$

Thus

$\tilde{p}(J_i(s_i)) = p(s_i) =$ probability that player i is told to play s_i.

$$\sum_{\omega \in J_i(s_i)} \frac{\tilde{p}(\omega)}{\tilde{p}(J_i(s_i))} \tau_{-i}(\omega).$$

It is the mixed strategy of the rivals that player i believes he faces, conditional on being told to play s_i, and it is a convex combination of the distributions conditional on each h_i such that $\tau_i(h_i) = s_i$.

Chapter 5

Adaptive and Sophisticated Learning

5.1 Adaptive and Sophisticated Learning

The idea of best reply dynamics goes back all the way to Cournot's study of duopoly and forms the foundation of Walrasian equilibrium in economy and is created by the classical *Tatonnement* learning process.

The underlying learning processes can be categorized into successively stronger versions:

- **Best-Reply Dynamics:** However, it's also known that this dynamics lead to non-convergent, cyclic behavior. In this model, an outsider with no information about the utilities (payoffs) of the agents could eventually predict the behavior of the agents more accurately than they themselves.

- **Fictitious-Play Dynamics:** The agents choose strategies that are best reply to predictions that the probability distributions of the competitors' play at the next round is based on the empirical distribution of the past plays. Even this dynamics lead to (if there is no zero-sum restriction) cycles of exponentially increasing lengths.

- **Stationary Bayesian Learning Dynamics:** The agents choose strategies as functions from the information set (empirical distribution of the past plays) without relying on any intermediate prediction. The distribution over the strategies changes as the empirical distribution changes. (Reactive Learning: involves no model building.)

  The dynamics may converge—but to a (mixed) strategy profile that is not necessarily the perfect (Nash) equilibrium.

5.2 Set-up

Player n plays a sequence of plays: $\{x_n(t)\}$. Each $x_n(t)$ is a pure strategy and is chosen by the rules of player n's learning algorithm. We are interested in two properties that may be satisfied by $\{x_n(t)\}$: it is *approximately* best-reply dynamics, then it is consistent with *adaptive learning*; it is *approximately* fictitious-play dynamics, then it is consistent with *sophisticated learning.*

Definition 5.2.1 $\{x_n(t)\}$ **is consistent with adaptive learning.** *Player n eventually chooses only strategies that are* nearly *best replies to some probability distribution over his rivals joint strategy profiles, where near* <u>*zero*</u> *probabilities are assigned to strategies that have not been played for sufficiently long time.*

Definition 5.2.2 $\{x_n(t)\}$ **is consistent with sophisticated learning.** *Player n eventually chooses only* nearly *best replies to his probabilistic forecast of rivals' joint strategy profiles, where the support of probability may include not only past plays but also strategies that the rivals may choose if they themselves were adaptive or sophisticated learners.*

We will look at the effect of these algorithms on *finite player games, with compact strategies and continuous pay-off functions.*

Note that these assumptions are consistent with the usual model of exchange economy with infinitely divisible goods. Note that in this model, serially undominated set is a singleton and thus the Walrasian equilibrium. One of the main results that we will see is that in any process, consistent with adaptive learning, play tends towards the serially undominated set and hence, in an exchange economy, adaptive learning would lead to equilibrium.

5.3 Formulation

Definition 5.3.1 *Noncooperative game*

$$\Gamma = (N, (S_n; n \in N), \pi).$$

$$\begin{aligned} N &= \textit{Finite Player Set} \\ S_n &= \textit{Player } n\textit{'s strategy} \\ &\quad\ \textit{Compact Subset of some Normed Space} \\ \pi &= \textit{Pay-off Function} \\ &\quad\ \textit{Assumed Continuous.} \end{aligned}$$

$$S = \times_{n \in N} S_n \quad x \in S \Rightarrow x = (x_n, x_{-n}).$$

x_{-n} *is the strategy choice of* n*'s rivals.*

$$\begin{aligned} \pi &: S \to \mathbb{R}^{|N|} = \textit{Pay-off Function, Continuous} \\ \pi_n &: S \to \mathbb{R} \\ &: (x_n, x_{-n}) \mapsto \pi_n(x). \end{aligned}$$

Let T *be a set. Then* $\Delta(T) =$ *Set of all probability distributions over* T.

$\Delta(S_n)$ = *Mixed strategies on* S_n. $\Delta_{-n} = \times_{j \neq n} \Delta(T_j)$ = *Mixed strategies of* n*'s rivals.*

Definition 5.3.2 *A strategy $x_n \in S_n$ is ϵ-dominated by another strategy $\bar{x}_n \in \Delta(S_n)$ if*

$$\forall_{z_{-n} \in S_{-n}}\ \pi_n(x_n, z_{-n}) + \epsilon < \pi_n(\bar{x}_n, z_{-n}).$$

If $\forall_\epsilon$ x_n is ϵ-dominated by $\bar{x}_n$, then x_n is dominated by $\bar{x}_n$ (in the classical sense).

Let $T \subseteq S$. Define $T_n \equiv T|_{S_n}$ = projection of T onto S_n. $T_{-n} = \times_{j \neq n} T_j$.

Definition 5.3.3 *Given $T \subseteq S$. Let*

$$\begin{aligned} U_n^\epsilon(T) &= \{x_n \in S_n : \forall_{y_n \in \Delta(S_n)}\ \exists_{z_{-n} \in T_{-n}} \\ &\qquad \pi_n(x_n, z_{-n}) + \epsilon \geq \pi_n(y_n, z_{-n})\} \\ U^\epsilon(T) &= \times_{n \in N} U_n^\epsilon(T). \end{aligned}$$

$U_n^\epsilon(T)$ = Pure strategies in S_n that are not ϵ-dominated when n's rivals are limited to T_{-n}.

Fact 1
The operator U^ϵ is monotonic. Let R and T be sets of strategy profiles.

$$R \subseteq T \quad \Rightarrow \quad U^\epsilon(R) \subseteq U^\epsilon(T).$$

Fact 2

$$\exists_T\ U^\epsilon(T) \not\subseteq T.$$

In general, starting with some arbitrary set of strategy profile T one may not be able to create a monotonically descending chain of sets of strategy profiles:

$$T \supseteq U^\epsilon(T) \supseteq U^{\epsilon,2}(T) \supseteq \cdots \supseteq U^{\epsilon,k}(T) \supseteq U^{\epsilon,k+1}(T) \supseteq \cdots$$

Fact 3
However, $S \supseteq U^\epsilon(S)$. Since S is the whole nothing new can get introduced.

By the monotonicity of U^ϵ, we see that if

$$U^{\epsilon,k}(T) \supseteq U^{\epsilon,k+1}(T),$$

then

$$U^{\epsilon}(U^{\epsilon,k}(T)) \supseteq U^{\epsilon}(U^{\epsilon,k+1}(T))$$

and

$$U^{\epsilon,k+1}(T) \supseteq U^{\epsilon,k+2}(T).$$

Putting it all together, we *do* have

$$S \supseteq U^{\epsilon}(S) \supseteq U^{\epsilon,2}(S) \supseteq \cdots \supseteq U^{\epsilon,k}(S) \supseteq U^{\epsilon,k+1}(S) \supseteq \cdots$$

We then define

$$U^{\epsilon\infty}(S) = \bigcap_{k=0}^{\infty} U^{\epsilon,k}(S).$$

Hence, $U^{0\infty}(S) = \lim_{\epsilon \to 0} U^{\epsilon,\infty}(S) =$ Serially undominated strategy set. We say x is serially undominated, if $x \in U^{0\infty}(S)$.

Definition 5.3.4 *A sequence of strategies $\{x_n(t)\}$ is consistent with adaptive learning by player n if*

$$\forall_{\epsilon>0} \forall_{\hat{t}} \exists_{\bar{t}} \forall_{t \geq \bar{t}}\, x_n(t) \in U_n^{\epsilon}\Bigg(\{x(s) : \hat{t} \leq s < t\}\Bigg).$$

A sequence of strategy profiles $\{x(t)\}$ is consistent with adaptive learning if each $\{x_n(t)\}$ has this property.

5.4 Looking Forward

$$F^{\epsilon,0}(\hat{t},t) = U^{\epsilon}\Bigg(\{x(s) : \hat{t} \leq s < t\}\Bigg).$$

$\forall\, k \geq 1$:

$$F^{\epsilon,k}(\hat{t},t) = U^{\epsilon}\Bigg(F^{\epsilon,k-1}(\hat{t},t)\ \cup\ \{x(s) : \hat{t} \leq s < t\}\Bigg).$$

Lemma 5.4.1

$$F^{\epsilon,0}(\hat{t},t) \subseteq F^{\epsilon,1}(\hat{t},t) \subseteq \cdots \subseteq F^{\epsilon,k}(\hat{t},t) \subseteq F^{\epsilon,k+1}(\hat{t},t) \subseteq \cdots$$

Proof

By the monotonicity of U^ϵ,

$$F^{\epsilon,0}(\hat{t},t) \subseteq F^{\epsilon,1}(\hat{t},t).$$

Assume by inductive hypothesis,

$$F^{\epsilon,k-1}(\hat{t},t) \subseteq F^{\epsilon,k}(\hat{t},t).$$

Then

$$\begin{aligned} & F^{\epsilon,k-1}(\hat{t},t) \ \cup \ \{x(s) : \hat{t} \leq s < t\} \\ & \subseteq \ \ F^{\epsilon,k}(\hat{t},t) \ \cup \ \{x(s) : \hat{t} \leq s < t\}. \end{aligned}$$

By the monotonicity of U^ϵ,

$$\begin{aligned} & U^\epsilon\Bigg(F^{\epsilon,k-1}(\hat{t},t) \ \cup \ \{x(s) : \hat{t} \leq s < t\}\Bigg) \\ & \subseteq \ \ U^\epsilon\Bigg(F^{\epsilon,k}(\hat{t},t) \ \cup \ \{x(s) : \hat{t} \leq s < t\}\Bigg). \end{aligned}$$

Thus

$$F^{\epsilon,k}(\hat{t},t) \subseteq F^{\epsilon,k+1}(\hat{t},t).$$

Definition 5.4.1 *A sequence of strategies $\{x_n(t)\}$ is consistent with sophisticated learning by player n if*

$$\forall_{\epsilon>0}\ \forall_{\hat{t}}\ \exists_{\bar{t}}\ \forall_{t\geq\bar{t}}\ x_n(t) \in U_n^\epsilon(F^{\epsilon\infty}(\hat{t},t)).$$

A sequence of strategy profiles $\{x(t)\}$ is consistent with sophisticated learning if each $\{x_n(t)\}$ has this property.

$$\forall_{\epsilon>0}\ \forall_{\hat{t}}\ \exists_{\bar{t}}\ \forall_{t\geq\bar{t}}\ x(t) \in F^{\epsilon\infty}(\hat{t},t).$$

Chapter 6

Learning a la Milgrom and Roberts

6.1 Adaptive Learning and Undominated Sets

Example: Battle of Sexes

W\M	Ballet(B)	Football(F)
Ballet(B)	2,1	0,0
Football(F)	0,0	1,2

Let $\{x(t)\}$ be a sequence of strategy profiles. We show that $x(t) = (F, B)$ is consistent with sophisticated learning.

$$\forall_{\hat{t}}\ \{x(s)|\hat{t} \leq s < t\} = \{(F, B)\}.$$

Thus, we have

$$\begin{aligned} F_W^{\epsilon,0}(\hat{t}, t) &= U_W^{\epsilon}\Big(\{(F, B)\}\Big) = B \\ F_M^{\epsilon,0}(\hat{t}, t) &= U_M^{\epsilon}\Big(\{(F, B)\}\Big) = F \end{aligned}$$

Thus

$$F^{\epsilon,0}(\hat{t},t) = \{\ (B,F)\ \}.$$

Similarly,

$$F^{\epsilon,1}(\hat{t},t) = U^{\epsilon}\Bigg(\{(B,F),(F,B)\}\Bigg) = \{B,F\} \times \{B,F\}.$$

Continuing in this fashion, we get

$$F^{\epsilon,\infty}(\hat{t},t) = \{B,F\} \times \{B,F\}.$$

Thus

$$x(t+1) = (F,B) \in F^{\epsilon,\infty}(\hat{t},t),$$

is consistent with sophisticated learning.

6.2 Convergence

Definition 6.2.1 *A sequence of strategy profiles $\{x(t)\}$ converges omitting correlation to a correlated strategy profile*

$$G \in \Delta(S)$$

if (1) and (2) hold:

1. *G_n^t converges weakly to the marginal distribution G_n for all n.*
2. $$\forall_{\epsilon>0}\ \exists_{\bar{t}}\ \forall_{t\geq\bar{t}}\ \forall_{n\in N}\ d[x_n(t), supp(G_n)] < \epsilon,$$

   Define $d[x,T] \equiv \inf_{y\in T}\|x-y\|$.

The sequence converges to the correlated strategy $G \in \Delta(S)$ if in addition

G^t converges weakly to G.

Definition 6.2.2 *A sequence $\{x(t)\}$ converges omitting correlation to a mixed strategy Nash equilibrium if*

1. *It replicates the empirical frequency of the separate mixed strategies and*

2. *It eventually plays only pure strategies that are in or near the support of the equilibrium mixed strategies.*

Theorem 6.2.1 *If $\{x(t)\}$ converges omitting correlation to a correlated equilibrium in the game Γ, then $\{x(t)\}$ is consistent with adaptive learning.*
Proof Sketch:
G^t converges to a correlated equilibrium G.
$\Rightarrow$ G_n consists of best responses to G_{-n}
$\Rightarrow$ For sufficiently large t, $x_n(t)$ is within ϵ of G_n
$\Rightarrow$ Since S_n is compact and π is continuous

$$\forall_{y_n \in G_n}\ \exists_{z_{-n} \in G_{-n}}\ \exists_{\delta > 0}\ \pi_n(x_n(t), z_{-n}) + \delta \geq \pi_n(y_n, z_{-n})$$

$\Rightarrow$

$$x_n(t) \in U_n^\delta\Bigg(\{x(s) \mid \hat{t} \leq s < t\}\Bigg).$$

Theorem 6.2.2 *Suppose that the sequence $\{x(t)\}$ is consistent with adaptive learning and that it converges to x^*. Then x^* is a pure strategy Nash equilibrium.*
Proof Sketch*:*
Assume that x^* is not a Nash equilibrium
$\Rightarrow$ $\exists_{n \in N}\ \forall_{\epsilon > 0}\ \{x_n^*\} \neq U^\epsilon(\{x^*\})$.
$\Rightarrow$ Player n must play $x_n' \neq x_n^*$ i.o.
$\Rightarrow$ $x_n(t)$ does not converge to x_n^*
$\Rightarrow$ **Contradiction**.

Theorem 6.2.3 *Let $\{x(t)\}$ be consistent with sophisticated learning. Then for each $\epsilon > 0$ and $k \in \mathbb{N}$ there exists a time $t_{\epsilon k}$ after which (i.e., for $t \geq t_{\epsilon k}$)*

$$x(t) \in U^{\epsilon k}(S).$$

Proof Sketch:
Fix $\epsilon > 0$. Define $t_k \equiv t_{\epsilon k}$ (Change in notation).

Case $k = 0$: $t_0 = 0$. $x(t) \in U^\epsilon(S)$.

Case $k = j + 1$: By the inductive hypothesis there exists a t_j such that

$$\forall_{t \geq t_j}\ x(t) \in U^{\epsilon j}(S).$$

Hence

$$\{x(s) \mid t_j \leq s < t\} \subseteq U^{\epsilon j}(S).$$

Since $\{x(t)$ is consistent with sophisticated learning, we can choose

$$\hat{t} = t_j, \quad t_{j+1} = \max(\hat{t}, \bar{t}).$$

Then

$$\forall_{t \geq t_{j+1}}\ x(t) \in F^{\epsilon\infty}(t_j, t).$$

Claim:

$$F^{\epsilon\infty}(t_j, t) \subseteq U^{\epsilon, j+1}(S).$$

Equivalently,

$$\forall_i\ F^{\epsilon i}(t_j, t) \subseteq U^{\epsilon, j+1}(S).$$

It then follows that

$$\begin{aligned} F^{\epsilon 0}(t_j, t) &= U^\epsilon\Big(\{x(s) \mid t_j \leq s < t\}\Big) \\ &\subseteq U^\epsilon\Big(U^{\epsilon, j}(S)\Big) = U^{\epsilon, j+1}(S). \end{aligned}$$

$$\begin{aligned} F^{\epsilon, i+1}(t_j, t) &= U^\epsilon\Big(F^{\epsilon, i}(t_j, t) \cup \{x(s) \mid t_j \leq s < t\}\Big) \\ &\subseteq U^\epsilon(U^{\epsilon, j+1}(S) \cup U^{\epsilon, j}(S)) \\ &= U^\epsilon(U^{\epsilon, j}(S)) = U^{\epsilon, j+1}(S). \end{aligned}$$

$$\bigcap_k \bigcap_{\epsilon > 0} U^{\epsilon k}(S) = \bigcap_k U^{0k}(S) = U^{0\infty}(S).$$

Theorem 6.2.4 *Let $\{x(t)\}$ be consistent with sophisticated learning and S_n^∞ be the set of strategies that are played infinitely often in $\{x_n(t)\}$. Then*

$$S^\infty \;=\; \times_{n\in N} S_n^\infty \;\subseteq\; \bigcap_k \bigcap_{\epsilon>0} U^{\epsilon k}(S) \;=\; U^{0\infty}(S)$$

Corollary 6.2.5 *In particular, for any finite game Γ, all play lies eventually in the set of serially undominated strategies $U^{0\infty}(S)$.*

Theorem 6.2.6 *Suppose $U^{0\infty}(S) = \{\bar{x}\}$.*

$$\|x(t) - \bar{x}\| \to 0$$

iff $\{x(t)\}$ is consistent with adaptive learning.
Proof Sketch:
($\Rightarrow$)

$$\|x(t) - \bar{x}\| \to 0.$$

Since π is continuous,

$$\begin{aligned}
&\forall_{\epsilon>0}\, \exists_{\bar{t}}\, \forall_{t>\bar{t}}\, \forall_{n\in N}\; \pi_n(x_n(t), x_{-n}(t)) - \max\{\pi_n(y_n, x_{-n}(t)) | y_n \in S_n\} \\
&\quad < \;\; [\pi_n(\bar{x}) + \epsilon/2] - [\max\{\pi_n(y_n, \bar{x}_{-n}) | y_n \in S_n\} - \epsilon/2] \\
&\quad = \;\; \epsilon.
\end{aligned}$$

$$x_n(t) \in U_n^\epsilon(\{x(\bar{t})\}) \;\subseteq\; U_n^\epsilon\Big(\{x(s) | \bar{t} \le s < t\}\Big).$$

$\Rightarrow$ $\{x(t)\}$ is consistent with adaptive learning.

($\Leftarrow$)
Let x^* = accumulation point of $\{x(t)\}$.

$$\forall_k\, \exists_{\bar{t}}\, \forall_{t>\bar{t}}\; x(t) \in U^{\epsilon k}(S).$$

$$\begin{aligned}
x^* \;&\in\; \bigcap_{\epsilon>0} \bigcap_{k=1}^{\infty} U^{\epsilon,k}(S) \\
&=\; \bigcap_k \bigcap_{\epsilon>0} U^{\epsilon,k}(S) = U^{0\infty}(S) = \{\bar{x}\}
\end{aligned}$$

$\Rightarrow$

$$\|x(t) - \bar{x}\| \to 0.$$

Theorem 6.2.7 *Suppose* $U^{0\infty}(S) = \{\bar{x}\}$.

$$\|x(t) - \bar{x}\| \to 0$$

iff $x(t)$ *is consistent with sophisticated learning.*

6.3 Stochastic Learning Processes

We now allow the players to experiment as we will now assume that each user may not know his own pay-off function. See Freudenberg & Kreps (1988).

Game consists of alternations among

- *Exploration*: Every strategy is experimented with equi-probability.
- *Exploitation*: Good strategies —based on exploration— are played.

At each date t, player n conducts an experiment with probability ϵ_{nt} in an attempt to learn its best play.

1. **Independence:** Decision to experiment is independent of other players' decisions.
2. **Rare:** $\epsilon_{nt} \to 0$ as $t \to \infty$.
3. **Infinitely Often:** $\sum_t \epsilon_{nt} = \infty$.

$\{t(k, \omega)\}$ = Subsequence of dates at which player n conducts no experiment.

ω = Realization of the players' randomized choices.

Thus the interval $[0, t]$ consists of experiment dates $P_n(x_n, t)$ and play dates $M_n(x_n, t)$. Write $M(t)$ to denote the expected total number of experiments.

- $\Pi(x_n, t)$ = Total Pay off received with $M_n(x_n, t)$
- $\Pi(y_n, t)$ = Total Pay off received with $M_n(y_n, t)$

Claim:
Let T = set of strategy profiles.

$$\forall_{z \in T}\ \pi_n(x_n, z_{-n}) < \pi_n(y_n, z_{-n}) - 2\epsilon$$

$\Rightarrow$ For large t

$$\Pi(x_n, t) < \Pi(y_n, t) - \epsilon M(t)/|S_n|.$$

$$\begin{aligned}
&E[\Pi(x_n, \tau+1) - \Pi(y_n, \tau+1)|T] - \Pi(x_n, \tau) + \Pi(y_n, \tau)\\
&\quad = \quad \epsilon_{n,\tau+1}/|S_n|\ E[\pi_n(x_n, x_{-n}(\tau+1)) - \pi_n(y_n, x_{-n}(\tau+1))]\\
&\quad < \quad -2\epsilon \cdot \epsilon_{n,\tau+1}/|S_n|.
\end{aligned}$$

Taking expectations

$$\begin{aligned}
&E[\Pi(x_n, \tau+1) - \Pi(y_n, \tau+1)|T]\\
&\quad = \quad E[\Pi(x_n, \tau) - \Pi(y_n, \tau)|T] - 2\epsilon \cdot \epsilon_{n,\tau+1}/|S_n|.
\end{aligned}$$

and then telescoping,

$$E[\Pi(x_n, t) - \Pi(y_n, t)] \quad < \quad -2\epsilon/|S_n| \sum \epsilon_{n,t} = -2\epsilon M(t)/|S_n|.$$

Let $\Delta > 2|\pi_n|$.

$$\mathrm{Var}[\Pi(x_n, t) - \Pi(y_n, t)] \quad \leq \quad -2\Delta^2 \epsilon M(t)/|S_n|.$$

Thus $\Pi(x_n, t) - \Pi(y_n, t)] + \epsilon M(t)/|S_n|$ converges to $-\infty$ and hence represents a super-martingale.

In other words, x_n dominates y_n then the player n will discover this fact eventually by repeated experiments.

Theorem 6.3.1 *For any finite strategy game Γ, the sequence $\{x_n(t(k, \omega))\}$ constructed as described above is consistent with adaptive learning a.s.(almost surely).*

Chapter 7

Information Theory and Learning

7.1 Information Theory and Games

7.1.1 Basic Concepts

Definition 7.1.1 *Entropy is a measure of uncertainty of a random variable. Let X be a discrete random variable with alphabet $\mathcal{X}$.*

$$p(x) = Pr[X = x], \quad \textit{where } x \in \mathcal{X}.$$

The entropy $H(X)$ of the discrete random variable X is defined as

$$\begin{aligned} H(X) &= \mathbb{E}_p \lg \frac{1}{p(X)} \\ &= -\sum_{x \in \mathcal{X}} p(x) \lg p(x). \end{aligned}$$

Facts

1. $H(X) \geq 0$. Entropy is always nonnegative. $0 \leq p(x) \leq 1$; $-\lg p(x) \geq 0$. Hence, $E_p \lg(1/p(x)) \geq 0$.)

2. $H(X) \leq \lg |\mathcal{X}|$. Consider the uniform distribution $u(x)$. $\forall_{x \in \mathcal{X}}\ u(x) = 1/|\mathcal{X}|$. $H(u) = \sum_x (1/|\mathcal{X}|) \lg |\mathcal{X}| = \lg |\mathcal{X}|$.

3. $H(X)$ = Average number of bits required to encode the discrete random variable X.

7.2 Joint & Conditional Entropy

(X, Y) = A pair of discrete random variables with joint distribution $p(x, y)$.

Joint Entropy =

$$\begin{aligned} H(X,Y) &= \mathbb{E}_p \lg \frac{1}{p(X,Y)} \\ &= -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y) \lg p(x,y). \end{aligned}$$

Conditional Entropy =

$$\begin{aligned} H(Y|X) &= \mathbb{E}_p \lg \frac{1}{p(Y|X)} \\ &= -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y) \lg p(y|x) \\ &= -\sum_{x\in\mathcal{X}} p(x) \sum_{y\in\mathcal{Y}} p(y|x) \lg p(y|x) \\ &= -\sum_{x\in\mathcal{X}} p(x) H(Y|x). \end{aligned}$$

7.2.1 Chain Rule

$$\begin{aligned} p(X,Y) &= p(X)\ p(Y|X) && \text{Bayes' Rule} \\ \lg p(X,Y) &= \lg p(X) + \lg p(Y|X) \\ \lg \frac{1}{p(X,Y)} &= \lg \frac{1}{p(X)} + \lg \frac{1}{p(Y|X)} \\ \mathbb{E}_p \lg \frac{1}{p(X,Y)} &= \mathbb{E}_p \lg \frac{1}{p(X)} + \mathbb{E}_p \lg \frac{1}{p(Y|X)} && \text{Linearity of Expectation} \\ H(X,Y) &= H(X) + H(Y|X). \end{aligned}$$

Corollary 7.2.1

1. *$H(X, Y|Z) = H(X|Z) + H(Y|X, Z)$.*
2. *$H(X) + H(Y|X) = H(Y) + H(X|Y)$*
   $\Rightarrow\ H(X) - H(X|Y) = H(Y) - H(Y|X)$.
3. *Note that $H(X|Y) \neq H(Y|X)$.*

7.3 Relative Entropy & Mutual Information

Definition 7.3.1 Relative Entropy—*Also,* **Kullback-Liebler Distance** *between two probability mass functions $p(x)$ and $q(x)$.*

$$D(p\|q) = \mathbb{E}_p \lg \frac{p(x)}{q(x)} = -\sum_x p(x) \lg \frac{q(x)}{p(x)}.$$

Note that $D(p\|p) = 0$. If $u(x) = \frac{1}{|\mathcal{X}|}$, for all x. Then $D(p\|u)$ is

$$D(p\|u) = -\sum p(x) \lg \frac{1}{p(x)} + \sum p(x) \lg |\mathcal{X}| = \lg |\mathcal{X}| - H(X).$$

Definition 7.3.2 Mutual Information
Let X and Y be two discrete random variables with a joint probability mass function $p(x, y)$, and with marginal probability mass functions

$$p(x) = \sum_{y \in \mathcal{Y}} p(x, y) \quad \& \quad p(y) = \sum_{x \in \mathcal{X}} p(x, y).$$

Mutual Information,

$$\begin{aligned}
I(X;Y) &= D\Big(p(x,y) \,\|\, p(x)p(y)\Big) \\
&= \mathbb{E}_{p(x,y)} \lg \frac{p(x,y)}{p(x)p(y)} \\
&= -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x,y) \lg \frac{p(x,y)}{p(x)p(y)}
\end{aligned}$$

$$
\begin{aligned}
&= H(X) + H(Y) - H(X,Y) \\
&= \Big(H(X) + H(Y)\Big) - \Big(H(Y) + H(X|Y)\Big) \\
&= H(X) - H(X|Y) = H(Y) - H(Y|X) = I(Y;X).
\end{aligned}
$$

$$H(X) - H(X|Y) = I(X;Y) = H(Y) - H(Y|X) = I(Y;X).$$

$$I(X;X) = H(X) - H(X|X) = H(X).$$

$$I(X;Y) = I(Y;X) = H(X) + H(Y) - H(X,Y).$$

7.4 Chain Rules for Entropy, Relative Entropy and Mutual Information

$$
\begin{aligned}
&H(X_1, X_2, \ldots, X_n) \\
&\quad = H(X_1) + H(X_2|X_1) + \cdots \\
&\qquad\quad + H(X_n|X_1, \ldots, X_{n-1}) \\
&\quad = \sum_{i=1}^{n} H(X_i|X_1, \ldots, X_{i-1}).
\end{aligned}
$$

$$
\begin{aligned}
&I(X_1, X_2, \ldots, X_n; Y) \\
&\quad = H(X_1, \ldots, X_n) - H(X_1, \ldots, X_n|Y) \\
&\quad = \sum_{i=1}^{n} H(X_i|X_1, \ldots, X_i) - \sum_{i=1}^{n} H(X_i|X_1, \ldots, X_i, Y) \\
&\quad = \sum_{i=1}^{n} H(X_i|X_1, \ldots, X_i) - H(X_i|X_1, \ldots, X_i, Y) \\
&\quad = \sum_{i=1}^{n} I(X_i; Y|X_1, \ldots, X_{i-1}).
\end{aligned}
$$

$$
\begin{aligned}
&D\Big(p(x,y) \parallel q(x,y)\Big) \\
&= \sum_x \sum_y p(x,y) \lg \frac{p(x,y)}{q(x,y)} \\
&= \sum_x \sum_y p(x,y) \lg \frac{p(x)\ p(y|x)}{q(x)\ q(y|x)} \\
&= \sum_x \sum_y p(x,y) \lg \frac{p(x)}{q(x)} + \sum_x \sum_y p(x,y) \lg \frac{p(y|x)}{q(y|x)} \\
&= \sum_x p(x) \lg \frac{p(x)}{q(x)} + \sum_y p(y|x) \lg \frac{p(y|x)}{q(y|x)} \\
&= D\Big(p(x) \parallel q(x)\Big) + D\Big(p(y|x) \parallel q(y|x)\Big).
\end{aligned}
$$

7.5 Information Inequality

$$
\begin{aligned}
-D(p\|q) &= \sum_x p(x) \lg \frac{q(x)}{p(x)} \quad \text{lg is a concave function} \\
&\le \lg \sum_x p(x) \frac{q(x)}{p(x)} \le \lg \sum_x q(x) = \lg 1 = 0.
\end{aligned}
$$

Theorem 7.5.1 $D(p\|q) \ge 0$ *(with equality iff $p(x) = q(x)$ for all x.)*

Corollary 7.5.2

$$
I(X;Y) = D\Big(p(x,y) \parallel p(x)p(y)\Big) \ \ge\ 0,
$$

(with equality iff X and Y are independent, i.e., $p(x,y) = p(x)p(y)$ for all x and y.)

Let $u(x) = \frac{1}{|\mathcal{X}|}$.

$$
D(p \parallel u) = \lg |\mathcal{X}| - H(X) \ge 0.
$$

Hence,

$$H(X) \leq \lg |\mathcal{X}|,$$

(with equality iff X has a uniform distribution over $\mathcal{X}$.)

$$I(X;Y) = H(X) - H(X|Y) \ \geq \ 0.$$

Theorem 7.5.3

$$H(X|Y) \ \leq \ H(X).$$

Conditioning reduces entropy.

$$\begin{aligned} H(X_1, \ldots, X_n) &= \sum_{i=1}^{n} H(X_i | X_1, \ldots, X_{i-1}) \\ &\leq \sum_{i=1}^{n} H(X_i) \end{aligned}$$

Corollary 7.5.4

$$H(X_1, \ldots, X_n) \ \leq \ \sum_{i=1}^{n} H(X_i),$$

with equality iff X_i's are independent.

7.6 Stationary Markov Process

- *Markovian*

$$Pr[X_n | X_1, \ldots, X_i] = Pr[X_n | X_i], \quad i \leq n.$$

- *Stationary*

$$Pr[X_n | X_1, \ldots, X_i] = Pr[X_{n+1} | X_2, \ldots, X_{i+1}].$$

$$
\begin{aligned}
H(X_n|X_1) &\geq H(X_n|X_1, X_2) \quad \text{conditioning reduces entropy} \\
&= H(X_n|X_2) \quad \text{Markov} \\
&= H(X_{n-1}|X_1) \quad \text{Stationary .}
\end{aligned}
$$

2nd Law of Thermodynamics

Theorem 7.6.1 *Conditional entropy $H(X_n|X_1)$ increases with time n for a stationary Markov process.*

Relative entropy $D(\pi_n \| \pi'_n)$ decreases with time n.
Let π_n and π'_n be two postulated probability distributions on the state space of a Markov Process. At time $n+1$, the distribution changes to π_{n+1} and π'_{n+1}, governed by the transition probabilities $r(x_n, x_{n+1})$.

Thus

$$
\begin{aligned}
p(x_n, x_{n+1}) &= p(x_n) r(x_n, x_{n+1}) \\
&= p(x_n) p(x_{n+1}|x_n)
\end{aligned}
$$

similarly,

$$
\begin{aligned}
q(x_n, x_{n+1}) &= q(x_n) r(x_n, x_{n+1}) \\
&= q(x_n) q(x_{n+1}|x_n)
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
& D\Big(p(x_n, x_{n+1}) \parallel q(x_n, x_{n+1})\Big) \\
&= D\Big(p(x_n) \parallel q(x_n)\Big) + D\Big(p(x_{n+1}|x_n) \parallel q(x_{n+1}|x_n)\Big) \\
&= D\Big(p(x_n) \parallel q(x_n)\Big).
\end{aligned}
$$

And

$$
D\Big(p(x_n, x_{n+1}) \parallel q(x_n, x_{n+1})\Big)
$$

$$
\begin{aligned}
&= D\Big(p(x_{n+1}) \parallel q(x_{n+1})\Big) + D\Big(p(x_n|x_{n+1}) \parallel q(x_n|x_{n+1})\Big) \\
&\geq D\Big(p(x_{n+1}) \parallel q(x_{n+1})\Big).
\end{aligned}
$$

We conclude that

$$D\Big(p(x_n) \parallel q(x_n)\Big) \geq D\Big(p(x_{n+1}) \parallel q(x_{n+1})\Big).$$

Thus the relative entropy for this system must decrease:

$$
\begin{aligned}
&D(\pi_1\|\pi_1') \;\geq\; D(\pi_2\|\pi_2') \;\geq\; \cdots \\
&\quad\geq\; D(\pi_n\|\pi_n') \;\geq\; D(\pi_{n+1}\|\pi_{n+1}') \;\geq\; \cdots\cdots \to 0.
\end{aligned}
$$

7.7 Gambling and Entropy

Horse Race

$$\#\text{ horses } = m, \quad \{H_1, H_2, \ldots, H_m\}.$$

$$p_i = Pr[H_i \text{ wins }]$$

$$u_i = \text{ pay-off if } H_i \text{ wins }.$$

If b_i = bet on the ith horse then the payoff =

$$
\begin{cases}
b_i u_i, & \text{if } H_i \text{ wins with probability } p_i; \\
0, & \text{if } H_i \text{ loses with probability } (1 - p_i).
\end{cases}
$$

Assume that the gambler has 1 dollar. Let b_i = fraction of his wealth invested in H_i. Thus

$$0 \leq b_i \leq 1. \qquad \sum_{i=1}^{m} b_i = 1.$$

Note that the gambler's pay-off is $b_i u_i$ if H_i wins (with probability p_i.)

$$S(X) = b(X)u(X)$$

= factor by which the gambler increases his wealth if X wins.

Repeated game with reinvestment.

$$\begin{aligned} S_0 &= 1, \\ S_n &= S_{n-1} S(X_n), \quad \text{if } X_n \text{ wins in the } n\text{th game.} \end{aligned}$$

Thus

$$S_n = \prod_{i=1}^{n} S(X_i) = 2^{\sum \lg S(X_i)}.$$

Let

$$\mathbb{E}_p[\lg S(X)] = \sum p_k \lg(b_k u_k) = W(b, p) = \text{ Doubling Rate},$$

where $b =$ the betting strategy. Then

$$\frac{1}{n} \lg S_n \to \mathbb{E}_p[\lg S(X)] \quad \text{in probability},$$

by "Law of Large Number." Hence

$$S_n \to 2^{nW(b,p)}.$$

Definition 7.7.1 Doubling Rate

$$W(b, p) = \sum_{k=1}^{m} p_k \lg(b_k u_k).$$

Theorem 7.7.1 *Let the race outcomes $X_1, \ldots, X_n$ be i.i.d.$\sim p(x)$. Then the wealth of the gambler using betting strategy b grows exponentially at rate $W(b, p)$, i.e.*

$$S_n \equiv 2^{nW(b,p)}.$$

$$
\begin{aligned}
W(b,p) &= \sum p_k \lg(b_k u_k) \\
&= \sum p_k \left[\lg \frac{b_k}{p_k} - \lg \frac{1}{p_k} + \lg u_k\right] \\
&= \sum p_k \lg u_k - H(p) - D(p\|b) \\
&\leq \sum p_k \lg u_k - H(p),
\end{aligned}
$$

with equality iff $p = b$.

The optimal doubling rate

$$W^*(p) = \max_b W(b,p) = W(p,p) = \sum p_k \lg u_k - H(p).$$

Theorem 7.7.2 Proportional gambling is log-optimal.
The optimum doubling rate is given by

$$W^*(p) = W(b^*, p) = \sum p_k \lg u_k - H(p),$$

and is achieved by the proportional gambling scheme, $b^* = p$.

Define $r_k = \frac{1}{u_k}$ = Bookie's estimate of the win "probabilities." Thus

$$\sum_k r_k = \sum \frac{1}{u_k} = 1.$$

Odds are fair and there is no track take.

$$
\begin{aligned}
W(b,p) &= \sum p_k \lg \frac{b_k}{r_k} \\
&= \sum p_k \left[\lg \frac{b_k}{p_k} - \lg \frac{r_k}{p_k}\right] \\
&= D(p\|r) - D(p\|b).
\end{aligned}
$$

Doubling Rate = Difference between the distance of the bookie's estimate from the true distribution and the distance of the gambler's estimate from the true distribution.

Special Case: Odds are m-for-1 on each horse:

$$\forall_k \;\; r_k = \frac{1}{m}.$$

Thus,

$$W^*(p) \;=\; D(p\|u) - D(p\|b^*) \;=\; \lg m - H(p).$$

Theorem 7.7.3 Conservation Theorem

$$W^*(p) + H(p) = \lg m$$

for uniform odds.

Low-Entropy Races are Most Profitable.
Case of a not fully invested gambler.

$$\begin{aligned} b_0 &= \text{wealth held out as cash} \\ b_i &= \text{proportional bet on } H_i. \end{aligned}$$

$$b_0 \geq 0, \quad b_i \geq 0, \quad \sum_{i=0}^{m} b_i = 1.$$

Thus

$$S(X) = b_0 + b(X)u(X).$$

- **Fair Odds:** $\sum \frac{1}{u_i} = 1$.
  If there is a non-fully-invested strategy with b_0, b_1, $\ldots$, b_m, then there is also a full investment as follows

$$\begin{aligned} b'_0 &= 0 \\ b'_i &= b_i + \frac{b_0}{u_i}, \quad 1 \leq i \leq m \\ \sum_{i=0}^{m} b'_i &= \sum_{i=1}^{m} b_i + b_0 \sum_{i=1}^{m} \frac{1}{u_i} = 1. \end{aligned}$$

  Thus

$$\begin{aligned} S(X) &= b'(X)u(X) = \frac{b_0}{u(X)} u(X) + b(X)u(X) \\ &= b_0 + b(X)u(X). \end{aligned}$$

  Thus in this case there is a risk-neutral investment.

- **Super-Fair Odds:** $\sum \frac{1}{u_i} < 1$.
  "Dutch Book" betting strategy.

  $$b_0 = 1 - \sum \frac{1}{u_i}, \quad b_i = \frac{1}{u_i}, 1 \le i \le m.$$

  Thus

  $$S(X) = \left(1 - \sum \frac{1}{u_i}\right) + \frac{1}{u(X)} u(X) = 2 - \sum \frac{1}{u_i} > 1$$

  with no risk!

  This, however, implies a strong arbitrage opportunity.

- **Sub-Fair Odds:** $\sum \frac{1}{u_i} > 1$.
  In this case, proportional gambling is no longer log-optimal and this case represents a risky undertaking for the gambler.

7.8 Side Information

Some external information about the performance of the horses may be available—for instance, previous games.
$X = \{1, 2, \ldots, m\}$, represent the horses.
$Y =$ Some other arbitrary discrete random variable
(*Side Information*).

$$p(x, y) = \text{ joint probability mass function for } (X, Y) \; .$$

$b(x|y) =$ conditional betting strategy depending on Y
$=$ proportion of wealth bet on horse x given that $y \in Y$ is observed.
$b(x) =$ unconditional betting strategy.

$$\begin{array}{ll} b(x) \ge 0, & \sum_x b(x) = 1. \\ b(x|y) \ge 0, & \sum_x b(x|y) = 1. \end{array}$$

$$
\begin{aligned}
W^*(X) &= \max_{b(x)} \sum_x p(x) \lg(b(x)u(x)) \\
&= \sum_x p(x) \lg u(x) - H(X).
\end{aligned}
$$

$$
\begin{aligned}
W^*(X|Y) &= \max_{b(x|y)} \sum_x p(x) \lg(b(x|y)u(x)) \\
&= \sum_x p(x) \lg u(x) - H(X|Y).
\end{aligned}
$$

$$
\begin{aligned}
\Delta W &= W^*(X|Y) - W^*(X) \\
&= \sum_x p(x) \lg u(x) - H(X|Y) - \sum_x p(x) \lg u(x) + H(X) \\
&= H(X) - H(X|Y) = I(X;Y) \geq 0.
\end{aligned}
$$

Increase in Doubling Rate =
Mutual information between the horse race and side information.

7.9 Learning

$\{X_k\}$ = Sequence of horse race outcomes from a stochastic process.

$$
\begin{aligned}
&W^*(X_k|X_{k-1}, X_{k-2}, \ldots, X_1) \\
&\quad= \mathbb{E}\left[\max_{b(.|x_{k-1},\ldots,x_1)} E[\lg S(X_k)|X_{k-1}, X_{k-2}, \ldots, X_1]\right] \\
&\quad= \lg m - H(X_k|X_{k-1}, X_{k-2}, \ldots, X_1),
\end{aligned}
$$

and is maximized for

$$
b^*(x_k|x_{k-1}, \ldots, x_1) = p(x_k|x_{k-1}, \ldots, x_1).
$$

Note that since

$$
S_n = \prod_{i=1}^n S(X_i),
$$

we have

$$
\begin{aligned}
\frac{1}{n} E \lg S_n &= \frac{1}{n} \sum E \lg S(X_i) \\
&= \frac{1}{n} \sum (\lg m - H(X_i | X_1, \ldots, X_{i-1})) \\
&= \lg m - \frac{H(X_1, \ldots, X_n)}{n} \\
&= \lg m - H(\mathcal{X}).
\end{aligned}
$$

$H(\mathcal{X})$ is simply the entropy rate.

Chapter 8

Universal Portfolio

8.1 Universal Portfolio

1. Sequential Portfolio Selection Procedure. An adapted process.

2. No statistical assumption about the behavior of the market.

3. Robust procedure with respect to arbitrary market sequences occurring in the real world.

We shall consider growth of wealth for arbitrary market sequences. For example, our goal may be to outperform the best buy-and-hold strategy—i.e., we wish to be competitive against a competing investor who can predict n future days. A different goal may be to outperform all constant rebalanced portfolio strategies.

$$\begin{aligned} m &= \#\text{ stocks traded in a market} \\ x_i &= \text{price relative for the } i\text{th stock} \\ &= \frac{\text{stock price at close}}{\text{stock price at open}} = \frac{P_i(c)}{P_i(o)} \\ &= 1 + \frac{\Delta P_i}{P_i}. \end{aligned}$$

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \text{ stock market vector .}$$

8.1.1 Portfolio

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = \text{ portfolio ,} \quad \left\{ \begin{array}{l} b_i \geq 0 \\ \sum_i b_i = 1. \end{array} \right.$$

Portfolio is simply the *proportion of the current wealth invested in each of the stocks.*

$$S = b \cdot x = b^T x = \sum_i b_i x_i,$$

= Factor by which the wealth increases in one period.

$$x(1), x(2), \ldots, x(n)$$

= stock market vectors for n consecutive days.

$$b = \text{ Fixed (constant) portfolio}$$

We shall follow a constant rebalanced portfolio strategy.

$$S_n(b) = \prod_{i=1}^{n} b^T x(i), \quad \left\{ \begin{array}{rcl} S_0(b) & = & 1 \\ S_n(b) & = & S_{n-1}(b)\, b^T x(n). \end{array} \right.$$

$$S_n^* = \max_b S_n(b) = S_n(b^*).$$

This is the maximum wealth achievable on the given stock sequence maximized over all constant rebalanced portfolios.

8.2 Universal Portfolio Strategy

$$\hat{b}(k)$$

depends only the past price relatives: $x(1)$, $x(2)$, …, $x(k-1)$.

It performs as well as the best constant rebalanced portfolio based on a clairvoyant knowledge of the sequence of price relatives.

8.2.1 Questions

Since we wish to compete against a clairvoyant investor (who knows the future) and universal portfolios only depend on the past (past has no *causal* or *correlated* relation with the future), how is it possible that universal portfolio can be competitive?

Malicious/adversarial nature is free to structure the future so as to help the competing investor.

$$\hat{b}(1) = \begin{pmatrix} 1/m \\ 1/m \\ \vdots \\ 1/m \end{pmatrix}.$$

$$S_k(b) = \prod_{i=1}^{k} b^T x(i), \quad B = \left\{ b \in \mathbb{R}_+^m \mid b_i \geq 0, \sum b_i = 1 \right\}.$$

$$\hat{b}(k+1) = \frac{\int_B b S_k(b) db}{\int_B S_k(b) db}$$

Note that

$$\hat{b}(k+1)^T x(k+1) = \frac{\int_B b^T x(k+1) S_k(b) db}{\int_B S_k(b) db} = \frac{\int_B S_{k+1}(b) db}{\int_B S_k(b) db}$$

The "learned" portfolio is the performance weighted average of all portfolios $b \in B$.

Thus

$$\hat{S}_n = \prod_{k=1}^{n} \hat{b}(k)^T x(k) = \frac{\int_B S_n(b) db}{\int_B db} = (m-1)! \int_B S_n(b) db.$$

We will show that

$$\hat{S}_n \approx S_n^* \frac{(m-1)!(\sqrt{2\pi/n})^{m-1}}{\sqrt{|J_n|},}$$

where J_n = a positive semidefinite $(m-1) \times (m-1)$ sensitivity matrix.

8.3 Properties & Analysis

Let F be some arbitrary probability distribution for price relatives over $\mathbb{R}_+^m$. Let F_n be the empirical distribution associated with $x(1)$, $x(2)$, ..., $x(n)$. $Pr[X = x(i)] = 1/n$. $Pr[X \neq x(i), \forall_i] = 0$.

$$\lim_{n \to infty} F_n \ \to \ F.$$

8.3.1 Doubling Ratio

$$\begin{aligned} W(b, F) &= \int \lg(b^T x) dF(x) \\ W(b, F_n) &= \sum_{i=1}^{n} \frac{1}{n} \lg(b^T x(i)) \\ W^*(F) &= \max_b W(b, F) \\ W^*(F_n) &= \max_b W(b, F_n) \end{aligned}$$

$$S_n^* = \max_b S_n(b) = \max_b \prod_{i=1}^{n} b^T x(i) = 2^{nW^*(F_n)}.$$

Let e_j be the vector

$$e_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad 1 \text{ in } j\text{th position only.}$$

$$\begin{aligned} S_n(e_j) &= \prod_{k=1}^{n} e_j^T x(k) = \prod_{k=1}^{n} x_j(k) \\ &= \text{Wealth due to buy-and-hold strategy} \\ &\qquad \text{associated wit the } j\text{th stock.} \end{aligned}$$

Since S_n^* is a maximization of $S_n(b)$ over the entire simplex,

$$\forall_j \;\; S_n^* \geq S_n(e_j).$$

Corollary 8.3.1

1. **Target Exceeds Best Stock.**

$$S_n^* \geq \max_j S_n(e_j).$$

2. **Target Exceeds Value Line.**

$$S_n^* \geq \left(\prod_j S_n(e_j) \right)^{1/m}$$

3. **Target Exceeds Arithmetic Mean.**

$$S_n^* \geq \sum_j \alpha_j S_n(e_j), \qquad \alpha_j \geq 0, \sum_j \alpha_j = 1.$$

4. $S_n^*(x(1), x(2), \ldots, x(n))$ is invariant under permutations of the sequence $x(1), x(2), \ldots, x(n)$.

Lemma 8.3.2

$$\hat{S}_n = \prod_{k=1}^{n} \hat{b}(k)^T x(k) = \frac{\int_B S_n(b)db}{\int_B db}$$

where

$$S_n(b) = \prod_{i=1}^{n} b^T x(i).$$

$\hat{S}_n$ *= Wealth from universal portfolio is the average of* $S_n(b)$ *over the simplex.*

Proof*:*

Recall that

$$\hat{b}(k+1)^T x(k+1) = \frac{\int S_{k+1}(b)db}{\int S_k(b)db}.$$

Telescoping the products

$$\begin{aligned}
\hat{S}_n &= \prod_{k=1}^{n} \hat{b}(k)^T x(k) \\
&= \frac{\int S_n(b)db}{\int S_{n-1}(b)db} \times \cdots \times \frac{\int S_1(b)db}{\int db} \\
&= \frac{\int S_n(b)db}{\int db} \\
&= \frac{\int_B \prod_{i=1}^{n} b^T x(i)db}{\int_B db} \\
&= \mathbb{E}_b S_n(b) = \mathbb{E}_b 2^{nW(b,F_n)}.
\end{aligned}$$

Corollary 8.3.3 $\hat{S}_n(x(1),\ x(2),\ \ldots,\ x(n))$ *is invariant under permutations of the sequence* $x(1),\ x(2),\ \ldots,\ x(n)$*.*

Claim

$$\mathbb{E}_b W(b, F_n) \geq \frac{1}{m} \sum_j W(e_j, F_n).$$

$$\mathbb{E}_b W(b, F_n) \quad = \quad \mathbb{E}_b \int \lg(b^T x)\, dF_n(x)$$

$$\begin{aligned}
&= \mathbb{E}_b \int \lg \sum b_j (e_j^T x)\ dF_n(x) \\
&\geq \mathbb{E}_b \sum b_j \int \lg(e_j^T x)\ dF_n(x) \\
&= \frac{1}{m} \sum \int \lg(e_j^T x) dF_n(x) \\
&= \frac{1}{m} \sum_j W(e_j, F_n).
\end{aligned}$$

By Jensen's inequality

$$\begin{aligned}
E_b 2^{nW(b,F_n)} &\geq 2^{nE_b W(b,F_n)} \\
&\geq 2^{1/m \sum nW(e_j,F_n)} \\
&\geq \left(\prod 2^{nW(e_j,F_n)}\right)^{1/m}.
\end{aligned}$$

Thus

$$\begin{aligned}
\hat{S}_n &= \mathbb{E}_b S_n(b) = \mathbb{E}_b 2^{nW(b,F_n)} \\
&\geq \left(\prod 2^{nW(e_j,F_n)}\right)^{1/m} \geq \left(\prod_{j=1}^{m} S_n(e_j)\right)^{1/m}.
\end{aligned}$$

Corollary 8.3.4 *Universal portfolio exceeds Value Line index.*

$$\hat{S}_n \geq \left(\prod_{j=1}^{m} S_n(e_j)\right)^{1/m}.$$

8.4 Competitiveness

$F_n(x)$ = Empirical probability mass function. Mass on each $x(i) \in \mathbb{R}_+^m$ is $\frac{1}{n}$.

$$\begin{aligned}
S_n(b) &= \prod_{i=1}^{n} b^T x(i) = 2^{nW(b,F_n)} = e^{nV(b,F_n)}, \\
b^*(F_n) = b^* &= \arg\max S_n(b) = \arg\max V(b, F_n) \in \mathbb{R}_+^m.
\end{aligned}$$

$$S_n^* = \max_{b \in B} S_n(b) = e^{nV^*(F_n)}.$$

Definition 8.4.1 *All stocks are active at time n, if*

$$\exists_{b^*; S_n(b^*)=S_n^*} \forall_{i\in[1..m]} \ (b^*(F_n))_i > 0.$$

All stocks are strictly active at time n, if

$$\forall_{b^*; S_n(b^*)=S_n^*} \forall_{i\in[1..m]} \ (b^*(F_n))_i > 0.$$

If

$$\text{Lin } (x(1), x(2), \ldots, x(n)) = \mathbb{R}^m,$$

then we say that the price relatives $x(1)$, $x(2)$, ..., $x(n)$ are of full rank.

$$J(b) = \ (m-1) \times (m-1) \text{ matrix .}$$

$J(b)$ = Sensitivity Matrix Function of a market with respect to distribution $F(x)$, $x \in \mathbb{R}_+^m$.

$$J_{ij}(b) = \int \frac{(x(i)-x(m))(x(j)-x(m))}{(b^T x)^2} dF(x).$$

$J^* = J(b^*)$ = Sensitivity Matrix.

$$J_{ij}^* = -\frac{\partial^2 V((b_1^*, \ldots, b_{m-1}^*, 1-\sum_{i=1}^{m-1} b_i^*), F)}{\partial b_i \, \partial b_j}$$

= Positive Semidefinite Matrix.

It is positive definite if all stocks are strictly active.

Let

$$C = \left\{ (c_1, c_2, \ldots, c_{m-1}) \mid c_i \geq 0, \sum c_i \leq 1 \right\}.$$

Define

$$b(c) = \left(c_1, \ldots, c_{m-1}, 1 - \sum_{i=1}^{m-1} c_i \right).$$

Thus

$$V_n(c) = \frac{1}{n} \sum_{i=1}^{n} \ln\left(b(c)^T x(i) \right) = \int \ln(b^T x) \, dF_n(x) \equiv \mathbb{E}_{F_n} \ln(b^T x).$$

Using Taylor series expansion:

$$\begin{aligned} V_n(c) &= V_n(c^*) + (c-c^*)^T \nabla V_n(c^*) \\ &\quad - \frac{1}{2}(c-c^*)^T J_n^*(c-c^*) \\ &\quad + \frac{1}{6}\sum_{ijk}(c_i - c_i^*)(c_j - c_j^*)(c_k - c_k^*) \\ &\quad \times \mathbb{E}_{F_n} \frac{(x(i)-x(m))(x(j)-x(m))(x(k)-x(m))}{S^3(\tilde{c})} \end{aligned}$$

where

$$\begin{aligned} \tilde{c} &= \lambda c^* + (1-\lambda)c, \quad \lambda \geq 0, \\ S(\tilde{c}) &= \sum_{i=1}^{m-1} b(\tilde{c})_i X(i). \end{aligned}$$

Assume that all stocks are strictly active:

$$J^* = -\left[\frac{\partial^2 V}{\partial c_i \partial c_j}\right] = \text{ positive definite.}$$

Hence its determinant is strictly positive:

$$|J^*| > 0.$$

Let $u = \sqrt{n}(c - c^*)$. Then since the second term is 0 in the Taylor series formulation, we have

$$\begin{aligned} nV_n(c) &= nV^*(F_n) - \frac{1}{2}u^T J_n^* u \\ &\quad + \frac{1}{6\sqrt{n}} \sum u_i u_j u_k \\ &\quad \times \mathbb{E}_{F_n} \frac{(x(i)-x(m))(x(j)-x(m))(x(k)-x(m))}{S^3(\tilde{c})}. \end{aligned}$$

Next assume that $0 < a \leq x(i) \leq c < \infty$.

$$S(\tilde{c}) \geq a, \quad x(i) - x(m) \leq 2c.$$

Thus the last term in the preceding expression can be bounded by

$$-\frac{1}{6\sqrt{n}}\|u\|^3 m^{3/2}\frac{(2c)^3}{a^3}.$$

Hence, we have

$$nV_n(c) \quad = \quad nV^*(F_n) - \frac{1}{2}u^T J_n^* u - \frac{4m^{3/2}c^3}{3\sqrt{n}a^3}\|u\|^3.$$

We thus conclude that

$$\begin{aligned} S_n(c) = 2^{nW_n(c)} &\geq e^{(nV_n^*)-(u^T J_n^* u/2)-(4m^{3/2}c^3\|u\|^3/3\sqrt{n}a^3)} \\ &= S_n^*\, e^{-(u^T J_n^* u/2)-(4m^{3/2}c^3\|u\|^3/3\sqrt{n}a^3)}. \end{aligned}$$

Since $\hat{S}_n = \int S_n(b)db/\int db$, and since $\int db = 1/(m-1)!$, we have

$$\hat{S}_n \geq S_n^*(m-1)! \int_{u\in U} e^{-(u^T J_n^* u/2)-(4m^{3/2}c^3\|u\|^3/3\sqrt{n}a^3)} \left(\frac{1}{\sqrt{n}}\right)^{m-1} du.$$

Thus,

$$\hat{S}_n \approx S_n^* \frac{(m-1)!(2\pi/n)^{(m-1)/2}}{|J_n|^{1/2}}.$$

In other words,

$$\frac{1}{n}\lg\frac{S_n^*}{\hat{S}_n} = \frac{1}{n}\lg\frac{|J_n|^{1/2}}{(m-1)!(2\pi/n)^{(m-1)/2}} \;\to\; 0, \qquad \text{as } n \to \infty\ .$$

Summarizing, we have

$$\begin{aligned} \frac{1}{n}\lg S_n^* &\sim \frac{1}{n}\lg \hat{S}_n \\ V_n^* &\sim \hat{V}_n. \end{aligned}$$

Chapter 9

Portfolios and Markets

9.1 Portfolio Theory

9.1.1 Itô Calculus

X = asset price at time t. In a continuous time model, one can study the return on the asset dX/X over a small period of time dt.

$$\frac{dX}{X} = \mu \; dt + \sigma \; dZ.$$

This is a so-called Itô process.

$$\begin{aligned} \mu &= \text{average rate of growth: } \textbf{DRIFT} \\ \sigma &= \text{volatility: } \textbf{DIFFUSION} \end{aligned}$$

9.1.2 Market Model

Assume that there are m stocks, represented by m Itô processes:

$$X_1(t), X_2(2), \ldots, X_m(t).$$

Furthermore,

$$\frac{dX_i}{X_i} = \mu_i \; dt + \sum_{j=1}^{m} \sigma_{ij} \; dZ_j,$$

Here, Z_j's are independent Brownian motions.

$$\mu = \textbf{Drift Vector} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{pmatrix}.$$

$$\sigma = \textbf{Diffusion Matrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1} & \sigma_{m2} & \cdots & \sigma_{mm} \end{pmatrix}$$

$$\Sigma = \textbf{Instantaneous Covariance Matrix} = \sigma\, \sigma^T.$$

In general, the term dZ corresponds to a *Wiener Process*.

- dZ = Normal Random Variable.
- $dZ \sim \mathcal{N}(0, \sqrt{dt})$. Mean of dZ is zero and variance of dZ is dt.

$$dZ = \phi\sqrt{dt}, \qquad E[\phi] = 0, \qquad E[\phi^2] = 1.$$

This holds in continuous time in the limit as $dt \to 0$.

Lemma 9.1.1 Itô's Lemma [Analogous to Taylor's theorem in case of functions of random variables. The key ideas is based on the observation that with probability 1, $dZ^2 \to dt$ as $dt \to 0$.]

Suppose $f(X)$ is a function of X (where X is possibly stochastic).

$$\begin{aligned}
df &= \frac{\partial f}{\partial X} dX + \frac{1}{2}\frac{\partial^2 f}{\partial X^2} dX^2 + \text{ smaller order terms} \\
dX^2 &= (\mu X dt + \sigma X dZ)^2 \\
&= \sigma^2 X^2 dZ^2 + 2\sigma\mu X^2 dZ dt + \mu^2 X^2 dt^2 \\
&\to \sigma^2 X^2 dt \qquad \text{as } dt \to 0 \\
df &= \frac{\partial f}{\partial X}(\mu X dt + \sigma X dZ) + \frac{1}{2}\sigma^2 X^2 \frac{\partial^2 f}{\partial X^2} dt \\
&= \left(\mu X \frac{\partial f}{\partial X} + \frac{1}{2}\sigma^2 X^2 \frac{\partial^2 f}{\partial X^2}\right) dt + \sigma X \frac{\partial f}{\partial X} dZ.
\end{aligned}$$

Example

$$\frac{dX}{X} = \mu\ dt + \sigma\ dZ$$

Let $f(X) = \ln X$. Then

$$\frac{\partial f}{\partial X} = \frac{1}{X}, \quad \& \quad \frac{\partial^2 f}{\partial X^2} = -\frac{1}{X^2}$$

$$\begin{aligned}
df &= \frac{\partial f}{\partial X}dX + \frac{1}{2}\frac{\partial^2 f}{\partial X^2}dX^2 \\
&= \frac{dX}{X} - \frac{1}{2X^2}\sigma^2 X^2 dt \\
&= \frac{dX}{X} - \frac{\sigma^2}{2}dt \\
d(\ln X) &= \frac{dX}{X} - \frac{\sigma^2}{2}dt \\
\frac{dX}{X} &= d(\ln X) + \frac{\sigma^2}{2}dt \\
\int_0^t \frac{dX}{X} &= \int_0^t d(\ln X) + \frac{1}{2}\int_0^t \sigma^2 dt \\
&= \ln X(t) - \ln X(0) + \frac{1}{2}\int_0^t \sigma^2 dt \\
\exp\left\{\int_0^t \frac{dX}{X}\right\} &= \frac{X(t)}{X(0)} \exp\left\{\frac{1}{2}\int_0^t \sigma^2 dt\right\}
\end{aligned}$$

9.2 Rebalanced Portfolio

Market Model with m stocks:

$$\begin{aligned}
\frac{dX_i(t)}{X_i(t)} &= \mu_i(t)dt + \sum_{j=1}^{m} \sigma_{ij}(t)dZ_j(t) \\
\Sigma(t) &= \sigma(t)\sigma(t)^T.
\end{aligned}$$

A portfolio of long stocks at time t is identified by its weighted vector process $b(t) \in B$.

$$B = \left\{ b \in \mathbb{R}^m \mid b_i \geq 0, \sum_{i=1}^{m} b_i = 1 \right\}.$$

Rebalanced Portfolio
(A self-financing portfolio without dividends).

$$\begin{aligned} \frac{dS(t)}{S(t)} &= \sum_{i=1}^{m} b_i(t) \frac{dX_i(t)}{X_i(t)} \\ &= \left(\sum_{i=1}^{m} b_i(t)\mu_i(t) \right) dt + \left(\sum_{i=1}^{m} \sum_{j=1}^{m} b_i(t)\sigma_{ij}(t) dZ_j \right). \end{aligned}$$

Let $g(S) = \ln S$ and $f(X) = \sum b_i \ln X_i = \ln \prod X_i^{b_i}$.

$$\begin{aligned} dg &= \frac{dS}{S} - \frac{1}{2S^2}(b^T \Sigma b) S^2 dt \\ \frac{dS}{S} &= d(\ln S) + \frac{1}{2}(b^T \Sigma b) dt \\ df &= \sum b_i \frac{dX_i}{X_i} - \sum \frac{1}{2X_i^2}(b_i \Sigma_{ii}) X_i^2 dt \\ \sum b_i \frac{dX_i}{X_i} &= d(\sum b_i \ln X_i) + \frac{1}{2} \sum b_i \Sigma_{ii} dt \end{aligned}$$

Hence

$$\begin{aligned} d(\ln S) &= d(\sum b_i \ln X_i) - \frac{1}{2} b^T \Sigma b dt + \frac{1}{2} \sum b_i \Sigma_{ii} dt \\ \ln \frac{S(t;b)}{S(0)} &= \sum b_i \ln \frac{X_i(t)}{X_i(0)} - \frac{1}{2} b^T \Lambda b + \frac{1}{2} \sum b_i \Lambda_{ii}, \end{aligned}$$

where $\Lambda \equiv \int_0^t \Sigma(s) ds$.

$$S(t;b) = S(0) \prod_{i=1}^{m} \left(\frac{X_i(t)}{X_i(0)} \right)^{b_i} \exp\left\{ -\frac{1}{2} b^T \Lambda b + \frac{1}{2} \sum \Lambda_{ii} b_i \right\}.$$

Maximizing the above expression we have

$$S^*(t) = \max_{b \in B} S(t; b) = S(t; b^*(t))$$

Note that $b^*(t)$ = optimal solution of the following quadratic programming problem:

$$\max_{b \in B} -\frac{1}{2} b^T \Lambda b + \sum_{i=1}^{m} \left(\ln \frac{X_i(t)}{X_i(0)} + \frac{1}{2} \Lambda_{ii} \right) b_i.$$

Define the matrix V, an $(m-1) \times (m-1)$ symmetric positive semidefinite matrix

$$V = (V_{ij}) \quad V_{ij} = \Lambda_{ij} - \Lambda_{im} - \Lambda_{jm} + \Lambda_{mm}, \ \ 1 \leq i, j \leq m.$$

Lemma 9.2.1 *If V = positive definite then the portfolio problem has a unique optimal solution.*

Definition 9.2.1 *A stochastic process $X(t)$ is* weakly regular *if*

$$\forall_t \ \ |E[X(t)]| < \infty.$$

$$\lim \frac{E[X(t)]}{t} = \theta \qquad \textit{exists}$$

$$\frac{X(t)}{t} \to \theta \qquad \textit{in probability as } t \to \infty.$$

The stock market model is weakly regular (easily satisfied if the market is stationary)

$$\forall_t \ \ |E[\Lambda(t)]| < \infty, \quad \& \quad |E[\ln X(t)]| < \infty,$$

$$\lim \frac{E[\Lambda(t)]}{t} = \Sigma^\infty, \quad \& \quad \lim \frac{E[\ln X(t)]}{t} = \eta^\infty \qquad \text{exist}$$

$$\frac{\Lambda(t)}{t} \to \Sigma^\infty, \quad \& \quad \frac{\ln X(t)}{t} \to \eta^\infty \qquad \text{in probability as } t \to \infty.$$

Note that

$$
\begin{aligned}
\frac{dX_i}{X_i} &= \mu_i dt + \sum \sigma_{ij} dZ_j \\
d(\ln X_i) &= \frac{dX_i}{X_i} - \frac{dX_i^2}{2X_i^2} \\
&= \left(\mu_i - \frac{\Sigma_{ii}}{2}\right) dt + \sum \sigma_{ij} dZ_j \\
\eta_i^\infty &= \mu_i^\infty - \frac{\Sigma_{ii}^\infty}{2}.
\end{aligned}
$$

Thus

$$\mu_i^\infty = \eta_i^\infty + \frac{\Sigma_{ii}^\infty}{2}.$$

Similarly,

$$
\begin{aligned}
\frac{dS}{S} &= \sum b_i \mu_i dt + \sum \sum b_i \sigma_{ij} dZ_j \\
d(\ln S) &= \left(b^T \mu - \frac{1}{2} b^T \Sigma b\right) dt + \sum \sum b_i \sigma_{ij} dZ_j \\
r(b) &= \lim \frac{E[\ln S(t;b)]}{t} = -\frac{1}{2} b^T \Sigma^\infty b + b^T \mu^\infty.
\end{aligned}
$$

Asymptotically optimal constant weight $b^\infty \in B$.

$$r(b^\infty) = \max_{b \in B} r(b) = \max_{b \in B} -\frac{1}{2} b^T \Sigma^\infty b + b^T \mu^\infty.$$

9.2.1 Optimal Portfolio

Recall

$$S(t;b) = S(0) \prod_{i=1}^m \left(\frac{X_i(t)}{X_i(0)}\right)^{b_i} \exp\left\{-\frac{1}{2} b^T \Lambda b + \frac{1}{2} \sum \Lambda_{ii} b_i\right\}.$$

$$V_{ij}(t) = \Lambda_{ij} - \Lambda_{im} - \Lambda_{jm} + \Lambda_{mm}.$$

Define

$$\lambda_i = \ln\left(\frac{X_m(t)}{X_m(0)}\right) - \ln\left(\frac{X_i(t)}{X_i(0)}\right) - \frac{V_{ii}(t)}{2}.$$

Notation:

$$b = (b', b_m) \quad b'_1 + \cdots + b'_{m-1} + b_m = 1, \quad b'_i \geq 0, b_m > 0.$$

Rewriting the previous equation, we have

$$S(t;b) = S(0)\frac{X_m(t)}{X_m(0)} \exp\left\{-\frac{1}{2}b'^T V b' - \lambda^T b'\right\}.$$

The above value $S(t;b)$ is maximized at $b' = \beta^*$

$$\begin{aligned} V(t)\beta^*(t) &= -\lambda(t) \\ \beta^*(t) &= -V^{-1}(t)\lambda(t) \end{aligned}$$

$$S^*(t) = S(0)\frac{X_m(t)}{X_m(0)} e^{\beta^{*T} V \beta^*/2},$$

and

$$S(t;b) = S^*(t) \ \exp\left\{-\frac{1}{2}(b' - \beta^*)^T V (b' - \beta^*)\right\}.$$

$$\frac{S(t;b)}{S^*(t)} = \exp\left\{-\frac{1}{2}(b' - \beta^*)^T V (b' - \beta^*)\right\}.$$

9.2.2 Long Term Effects

$$\left.\begin{aligned} V_{ij} &= \Lambda_{ij} - \Lambda_{im} - \Lambda_{jm} + \Lambda_{mm} \\ J^\infty_{ij} &= \Sigma^\infty_{ij} - \Sigma^\infty_{im} - \Sigma^\infty_{jm} + \Sigma^\infty_{mm} \end{aligned}\right\} \quad \lim \frac{V(t)}{t} = J^\infty$$

$$\left.\begin{aligned} \lambda_i &= \ln\left(\tfrac{X_m(t)}{X_m(0)}\right) - \ln\left(\tfrac{X_i(t)}{X_i(0)}\right) - \tfrac{V_{ii}(t)}{2} \\ \gamma^\infty_i &= \eta^\infty_m - \eta^\infty_i - \tfrac{J^\infty_{ii}}{2} \\ &= \mu^\infty_m - \tfrac{\Sigma^\infty_{mm}}{2} - \mu^\infty_i - \tfrac{\Sigma^\infty_{ii}}{2} \\ &\quad - \tfrac{\Sigma^\infty_{ii}}{2} + \Sigma^\infty_{im} - \tfrac{\Sigma^\infty_{mm}}{2} \\ &= \mu^\infty_m - \mu^\infty_i + \Sigma^\infty_{im}. \end{aligned}\right\} \quad \lim \frac{\lambda_i(t)}{t} = \gamma^\infty_i$$

Since

$$\begin{aligned} r(b) &= -\frac{1}{2}b^T \Sigma^\infty b + b^T \mu^\infty \\ &= -\frac{1}{2}b'^T J^\infty b' - b'^T \gamma^\infty, \end{aligned}$$

it is maximized at

$$\beta^\infty = -(J^\infty)^{-1}\gamma^\infty.$$

Note, however, that

$$\beta^*(t) = -\left(\frac{V(t)}{t}\right)^{-1}\left(\frac{\lambda(t)}{t}\right) \to -(J^\infty)^{-1}\gamma^\infty = \beta^\infty.$$

Problem: Construction of b^∞ requires the long-term average of future instantaneous expected returns and covariances. This however is impossible.
Remedy: *Universal Portfolio*

9.3 Universal Portfolio

Rebalanced portfolio with weights:

$$\hat{b}_i(t) = \frac{\int_B b_i S(t;b)db}{\int_B S(t;b)db}.$$

Let

$$\bar{S} = \frac{\int_B S(t;b)db}{\int_B db}$$

Note that

$$\bar{S}(0) = \hat{S}(0).$$

Furthermore,

$$\begin{aligned} \frac{d\bar{S}}{\bar{S}} &= \frac{\int_B dS(t;b)db}{\int_B S(t;b)db} = \frac{\int_B \sum_i S(t;b)b_i(dX_i/X_i)db}{\int_B S(t;b)db} \\ &= \sum_i \hat{b}_i(t)\frac{dX_i}{X_i} \\ &= \frac{d\hat{S}}{\hat{S}} \end{aligned}$$

Hence,

$$\forall_t \; \hat{S}(t) = \bar{S}(t).$$

Lemma 9.3.1 *The wealth accumulated by a universal portfolio is given by*

$$\hat{S}(t) = \frac{\int_B S(t;b)db}{\int_B db}.$$

This is the average wealth accumulated by all possible portfolios.

9.3.1 Competitiveness

$$S(t;b) = S^*(t)e^{-\frac{1}{2}(b'-\beta^*)^T V(b'-\beta^*)}.$$

Let $x = V^{1/2}(t)(b' - b'^*)$. Thus

$$\Delta(t) = V^{1/2}(t)(B' - b'^*),$$

where

$$B' = \left\{ b' \in \mathbb{R}^{m-1} \mid b'_i \geq 0, \sum b'_i < 1 \right\}.$$

Note that

$$\text{Vol}\,(B') = \frac{1}{(m-1)!}.$$

We have

$$\hat{S}(t) = \frac{S^*(t) \int_{\Delta(t)} e^{-|x|^2/2}\, dx}{|V(t)|^{1/2}(1/(m-1)!)}.$$

$$\begin{aligned} \frac{\hat{S}(t)}{S^*(t)} &= \frac{(m-1)! \int_\Delta e^{-|x|^2/2} dx}{\left(\left|\frac{V(t)}{t}\right|\right)^{1/2} t^{m-1/2}} \\ &= \frac{(m-1)! (\sqrt{2\pi})^{m-1}}{|J\infty|^{1/2} t^{m-1/2}} \\ &= \frac{(m-1)!}{|J\infty|^{1/2}} \left(\frac{2\pi}{t}\right)^{m-1/2}. \end{aligned}$$

Thus,

$$\frac{1}{t} \ln \frac{\hat{S}(t)}{S^*(t)} = \frac{C(m) - C'(m) \ln t}{t} \to 0$$

and

$$\frac{\ln \hat{S}(t)}{t} \to \frac{\ln S^*(t)}{t} \to \frac{\ln S(t; b^\infty)}{t}.$$

Bibliography

Text Books

[1] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*, John Wiley & Sons, 1991. `ISBN 0-471-06259-6.`

[2] Darrel Duffie. *Dynamic Asset Pricing Theory*, Princeton, 1997. `ISBN 0-691-04302-7.`

[3] Drew Fudenberg and Jean Tirole. *Game Theory*, MIT, 1995. `ISBN 0-262-06141-4.`

[4] Alan Kirman and Mark Salmon. *Learning and Rationality in Economics*, Basil Blackwell, 1995. `ISBN 0-631-18488-0.`

[5] H.M. Markovitz. *Means-Variance Analysis in Portfolio Choice and Capital Markets*, Blackwell, 1991. `ISBN 0-631-17854-6.`

Popular Books

[6] William Poundstone. *Prisoner's Dilemma* Doubleday, 1992.

[7] Anatol Rapoport. *Two-Person Game Theory: The Essential Ideas*, Ann Arbor Science Paperbacks, University of Michigan, 1966.

[8] Karl Sigmund. *Games of Life: Explorations in Ecology, Evolution and Behaviour* Oxford University Press, 1993.