**CS 405/6001: Game Theory and Algorithmic Mechanism Design** **Jul–Nov 2026**

# Lecture 4: Security, Elimination, and Matrix Games

*Lecturer: Swaprava Nath* *Scribe(s): Claude Scribe*

**Disclaimer**: *These notes aggregate content from several texts and have not been subjected to the usual scrutiny deserved by formal publications. If you find errors, please bring to the notice of the Instructor.*

## 4.1 Recap

Last time we saw that trusting the opponent to play their part of a Nash equilibrium is a strong assumption. Today we make the alternative, *worst-case*, notion of rationality precise: the **maxmin** value. We then study how iterated elimination interacts with maxmin values and with PSNE, and finish with **two-player zero-sum (matrix) games**.

## 4.2 Maxmin Strategies

**Example 4.1 (Risk aversion)** *Consider the two-player game*

| | | *Player 2* $L$ | $R$ |
|---|---|---|---|
| | $T$ | $2, 1$ | $1, -20$ |
| *Player 1* | $M$ | $3, 0$ | $-10, 1$ |
| | $B$ | $-100, 2$ | $3, 3$ |

*What if Player 2 does not play a Nash-equilibrium action? Which row should a risk-averse Player 1 choose?*

**Definition 4.2 (Maxmin value and strategy, MSZ §4.10)** *Player $i$'s **maxmin value** is $\underline{v}_i := \max_{s_i \in S_i} \min_{s_{-i} \in S_{-i}} u_i(s_i, s_{-i})$. Any $s_i^{\max\min} \in \arg\max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i})$ is a **maxmin strategy**: it guarantees $u_i \geqslant \underline{v}_i$ no matter what the others do.*

In Example 4.1, Player 1's worst case for each row is min over columns: $T$ gives $\min(2, 1) = 1$, $M$ gives $\min(3, -10) = -10$, $B$ gives $\min(-100, 3) = -100$. Player 1's maxmin strategy is $T$, guaranteeing at least 1 – far less than the 3 available at $M$ if Player 2 happens to play $L$, but it is what can be *guaranteed* with no assumption about Player 2's behaviour at all. Picking $T$ is simply the less risky choice.

**Remark 4.3 (Connection to robust optimization)** *This worst-case logic is exactly what underlies robust optimization formulations of adversarial training in modern machine learning, where a model is trained to maximize performance under a worst-case perturbation of its input rather than an expected one [MMS18]. The maxmin value is precisely the guaranteed performance of such a model against any perturbation within the assumed budget.*

**Theorem 4.4** *If $s_i^*$ is a dominant strategy for player $i$, it is also a maxmin strategy for $i$.*

**Proof:** Since $s_i^*$ is dominant, $u_i(s_i^*, s_{-i}) \geqslant u_i(s_i', s_{-i})$ for every $s_{-i} \in S_{-i}$ and every $s_i' \in S_i$. Fix any $s_i'$ and let $s_{-i}^{\min}(s_i') \in \arg\min_{s_{-i}} u_i(s_i', s_{-i})$ be the worst reply to $s_i'$. Plugging in $s_{-i} = s_{-i}^{\min}(s_i')$:

$$u_i(s_i^*, s_{-i}^{\min}(s_i')) \geqslant u_i(s_i', s_{-i}^{\min}(s_i')) = \min_{s_{-i}} u_i(s_i', s_{-i}), \quad \forall s_i' \in S_i.$$

So $s_i^*$'s worst case (the left side, minimized over $s_{-i}$) is at least every other strategy's worst case. Hence $s_i^* \in \arg\max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i})$. ■

**Theorem 4.5 (MSZ Thm. 4.29)** *Every PSNE $s^*$ of $G$ satisfies $u_i(s^*) \geqslant \underline{v}_i$ for every $i \in N$.*

**Proof:** For any $s_i \in S_i$, $u_i(s_i, s_{-i}^*) \geqslant \min_{s_{-i}} u_i(s_i, s_{-i})$ by definition of min. Since $s^*$ is a PSNE, $u_i(s_i^*, s_{-i}^*) = \max_{s_i} u_i(s_i, s_{-i}^*) \geqslant \max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i}) = \underline{v}_i$. ■

So a Nash equilibrium payoff can never fall below what a player could have guaranteed alone – reassuring, but Example 4.1-style reasoning shows the *guarantee itself* is often what matters when you cannot be sure the opponent will cooperate with an equilibrium.

## 4.3 Iterated Elimination: Effects on Maxmin and on PSNE

The story so far: dominance cannot explain all outcomes, and games may not have dominant strategies; PSNE captures *stability* (no unilateral deviation); maxmin captures *security* (a worst-case guarantee). What happens to stability and to security when some strategies are eliminated from a game?

**Example 4.6 (Elimination order, revisited)** *Recall the $3 \times 3$ game from Lecture 3:*

| | | *Player 2* | | |
|---|---|---|---|---|
| | | $L$ | $C$ | $R$ |
| | $T$ | $1,2$ | $2,3$ | $0,3$ |
| *Player 1* | $M$ | $2,2$ | $2,1$ | $3,2$ |
| | $B$ | $2,1$ | $0,0$ | $1,0$ |

*eliminating in the order $T, R, B, C$ reaches $(M, L)$, while eliminating in the order $B, L, C, T$ reaches $(M, R)$. Does eliminating a dominated strategy change a player's maxmin value?*

To answer this concretely, tweak the game slightly: change Player 2's payoff at $(B, L)$ from 1 to 0, leaving everything else the same:

| | | Player 2 | | |
|---|---|---|---|---|
| | | $L$ | $C$ | $R$ |
| | $T$ | $1,2$ | $2,3$ | $0,3$ |
| Player 1 | $M$ | $2,2$ | $2,1$ | $3,2$ |
| | $B$ | $2,0$ | $0,0$ | $1,0$ |

Here $M$ weakly dominates $B$ for Player 1 (tied at column $L$, strictly better at $C$ and $R$). Player 1's maxmin value: before elimination, row minimums are $T:0$, $M:2$, $B:0$, so $\underline{v}_1 = 2$; after removing $B$, still $\underline{v}_1 = 2$ – unchanged. Player 2's maxmin value: before elimination, column minimums are $L:0$, $C:0$, $R:0$ (all dragged down by row $B$'s zeros), so $\underline{v}_2 = 0$; after removing $B$, column minimums become $L:2$, $C:1$, $R:2$, so $\underline{v}_2 = 2$ – Player 2's own security *increased* once Player 1's dominated row was gone, simply because Player 2 no longer has to guard against it.

**Theorem 4.7 (MSZ Thm. 4.30)** *Let $s'_j \in S_j$ be a dominated strategy of player $j$, and let $G'$ be the game after removing $s'_j$. Then player $j$'s maxmin value is unchanged: $\underline{v}_j$ in $G$ equals $\underline{v}'_j$ in $G'$.*

**Proof:** Let $t_j$ dominate $s'_j$, so $u_j(t_j, s_{-j}) \geqslant u_j(s'_j, s_{-j})$ for all $s_{-j}$. Then

$$\min_{s_{-j}} u_j(t_j, s_{-j}) \geqslant \min_{s_{-j}} u_j(s'_j, s_{-j}),$$

so removing $s'_j$ cannot lower the best worst-case available to $j$ among $S_j \setminus \{s'_j\}$, since $t_j$ was already there and already does at least as well against every $s_{-j}$. Hence

$$\underline{v}_j = \max\left\{ \max_{s_j \neq s'_j} \min_{s_{-j}} u_j(s_j, s_{-j}),\ \min_{s_{-j}} u_j(s'_j, s_{-j}) \right\} = \max_{s_j \neq s'_j} \min_{s_{-j}} u_j(s_j, s_{-j}) = \underline{v}'_j.$$

■

**Remark 4.8** *Removing player $j$'s* own *dominated strategy never hurts $j$'s security level (Theorem 4.7). But as the worked example above shows, removing $j$'s dominated strategy can* help *another player $k$: fewer strategies for $j$ means fewer bad cases $k$ must guard against, so $k$'s maxmin value is monotonically non-decreasing under elimination of* anyone's *dominated strategies. This is why, when computing player $j$'s maxmin value specifically, we should eliminate only $j$'s own dominated strategies – eliminating others' can inflate the number we get.*

### 4.3.1 Preservation of PSNE

**Theorem 4.9 (MSZ Thm. 4.31)** *Let $G$ and $\hat{G}$ be games before and after elimination of a strategy (not necessarily dominated). If $s^*$ is a PSNE in $G$ and survives in $\hat{G}$, then $s^*$ is a PSNE in $\hat{G}$ too.*

The intuition: a PSNE is a maximizer of $i$'s utility among $i$'s *own* strategies, holding others fixed; removing strategies (of anyone) cannot break a maximality that already held over a superset. We omit the short formal proof (an exercise).

**Theorem 4.10 (MSZ Thm. 4.32)** *Let $\hat{s}_j$ be a weakly dominated strategy of player $j$ in $G$, and let $\hat{G}$ be $G$ with $\hat{s}_j$ removed. Then every PSNE of $\hat{G}$ is already a PSNE of $G$: elimination of a dominated strategy creates no new equilibria.*

**Proof:** Let $\hat{S}_j = S_j \setminus \{\hat{s}_j\}$ and $\hat{S}_i = S_i$ for $i \neq j$. Let $s^* = (s^*_j, s^*_{-j})$ be a PSNE of $\hat{G}$, so

$$u_i(s^*) \geqslant u_i(s_i, s^*_{-i}),\ \forall i \neq j,\ \forall s_i \in S_i, \qquad u_j(s^*) \geqslant u_j(s_j, s^*_{-j}),\ \forall s_j \in \hat{S}_j.$$

We must show $s^*$ has no profitable deviation in $G$ either. For $i \neq j$ this is immediate, since no strategy of theirs was removed. For player $j$, we already know there is no profitable deviation to any $s_j \in \hat{S}_j$; it remains to check $\hat{s}_j$ itself. Since $\hat{s}_j$ is dominated, there is some $t_j \in \hat{S}_j$ with $u_j(t_j, s_{-j}) \geqslant u_j(\hat{s}_j, s_{-j})$ for every $s_{-j}$; in particular at $s_{-j} = s^*_{-j}$,

$$u_j(s^*_j, s^*_{-j}) \geqslant u_j(t_j, s^*_{-j}) \geqslant u_j(\hat{s}_j, s^*_{-j}),$$

where the first inequality is because $s^*$ is a PSNE of $\hat{G}$ and $t_j \in \hat{S}_j$. So switching to $\hat{s}_j$ is not profitable either, and $s^*$ is a PSNE of $G$. ■

**Remark 4.11** *Combining Theorems 4.9 and 4.10: elimination of a dominated strategy never* creates *a new PSNE, but it can still* destroy *one that does not survive the elimination. Example 4.6 already shows this concretely: eliminating in the order $T, R, B, C$ reaches $(M, L)$, but if the elimination order instead removes the strategies needed to reach $(M, L)$ before Player 1 or 2 can settle there, that equilibrium is gone from the process even though it was a perfectly valid PSNE of the original game – only $(M, R)$ survives that particular order. Which equilibrium you find can depend on which strategies you happen to eliminate first, and in what order.*

- Elimination of a *strictly* dominated strategy has no effect on the set of PSNE.
- Elimination of a *weakly* dominated strategy may shrink the set of PSNE, but never adds a new one.
- The maxmin value of the player whose own (strictly or weakly) dominated strategy is removed is unaffected; other players' maxmin values can only rise.

## 4.4 Two-Player Zero-Sum Games and Matrix Games

**Definition 4.12 (Zero-sum game, MSZ Def. 4.39)** *A two-player game is* ***zero-sum*** *if $u_1(s)+u_2(s) = 0$ for every strategy profile $s$. Writing $u := u_1 = -u_2$, the entire game is captured by one matrix – hence* ***matrix game****: Player 2's payoff is simply the negative of the single matrix $u$ that records Player 1's payoffs.*

**Example 4.13 (Penalty shoot-out)** *Two players, Shooter and Goalkeeper, each independently choose $L$ or $R$. If they choose the same side the shot is saved (Shooter loses, Goalkeeper wins); if they choose opposite sides the shot scores (Shooter wins). Writing the payoff as $+1/-1$ to the winner/loser gives a zero-sum game, summarized by the single matrix $u$ (Shooter's payoff):*

| | | *Goalkeeper* $L$ | $R$ |
|---|---|---|---|
| *Shooter* | $L$ | $-1$ | $1$ |
| | $R$ | $1$ | $-1$ |

$$u = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

*Row minimums are $-1$ and $-1$, so the Shooter's maxmin value is $\underline{v} = \max\{-1, -1\} = -1$. Column maximums are $1$ and $1$, so the Goalkeeper's minmax value is $\overline{v} = \min\{1, 1\} = 1$. Since $\underline{v} = -1 \neq 1 = \overline{v}$, verify for yourself (by checking all four cells for a profitable deviation) that* ***this game has no PSNE****.*

**Definition 4.14 (Saddle point, MSZ Def. 4.47)** *$(s_1^*, s_2^*)$ is a* ***saddle point*** *of $u$ if $u(s_1^*, s_2^*) \geqslant u(s_1, s_2^*)\ \forall s_1$ and $u(s_1^*, s_2^*) \leqslant u(s_1^*, s_2)\ \forall s_2$: simultaneously the max in its column and the min in its row.*

**Theorem 4.15** *$(s_1^*, s_2^*)$ is a saddle point of a matrix game $u$ if and only if it is a PSNE.*

**Proof:** Since $u \equiv u_1 \equiv -u_2$, the saddle-point inequalities $u(s_1^*, s_2^*) \geqslant u(s_1, s_2^*)\ \forall s_1$ and $u(s_1^*, s_2^*) \leqslant u(s_1^*, s_2)\ \forall s_2$ translate exactly into $u_1(s_1^*, s_2^*) \geqslant u_1(s_1, s_2^*)\ \forall s_1$ and $u_2(s_1^*, s_2^*) \geqslant u_2(s_1^*, s_2)\ \forall s_2$ (flipping sign reverses the second inequality's direction) – which is precisely the PSNE condition for both players. ■

**Corollary 4.16** *A matrix game has a PSNE if and only if $\underline{v} = \overline{v}$, i.e., maxmin equals minmax.*

Example 4.13 is therefore a game with provably *no* pure-strategy equilibrium at all – not a failure of our analysis, but a real feature of these games: a shooter who always shot the same side would be trivially exploited, and vice versa for the goalkeeper. That unpredictability-by-necessity is the intuitive content of $\underline{v} \neq \overline{v}$. Resolving this gap is the job of **mixed strategies**, which is where we pick up next week with von Neumann's minimax theorem [vN28] and Nash's existence theorem [Nas51].

## 4.5 Summary and Preview

We defined the maxmin (security) value and maxmin strategies, related them to dominant strategies (Theorem 4.4) and to PSNE (Theorem 4.5); showed iterated elimination of a player's own dominated strategy preserves that player's maxmin value while it can only help other players' (Theorem 4.7); showed elimination of a dominated strategy never creates new PSNE but can destroy old ones (Theorems 4.9–4.10); and specialized to two-player zero-sum (matrix) games, where PSNE, saddle points, and maxmin=minmax coincide (Theorem 4.15), with the penalty shoot-out as a game with none of these. Next week: mixed strategies, the characterization of mixed-strategy Nash equilibrium, and Nash's 1951 existence theorem, which guarantees every finite game has *some* equilibrium once randomization is allowed.

## References

[MMS18] A. Madry, A. Makelov, L. Schmidt, D. Tsipras and A. Vladu, "Towards deep learning models resistant to adversarial attacks," *ICLR*, 2018.

[MSZ13] M. Maschler, E. Solan and S. Zamir, *Game Theory*, Chapter 4, Cambridge University Press, 2013.

[Nas51] J. Nash, "Non-cooperative games," *Annals of Mathematics*, 54(2), 1951, pp. 286–295.

[vN28] J. von Neumann, "Zur Theorie der Gesellschaftsspiele," *Mathematische Annalen*, 100(1), 1928, pp. 295–320.