Online Optimization and Learning in Games

Instructor: Chen-Yu Wei

This is a tentative syllabus. The current texts are drafted by AI, but I will revise them before the course starts, mostly by cutting or compressing some topics. At least 70%-75% of the topics should stay. The order and emphasis may change as the course develops.

Course overview

This course studies online learning through one recurring idea: *no-deviation guarantees*. Imagine that a learner makes decisions over time. Afterward, an auditor asks: "On the rounds when you did X, would you have done better by doing Y instead?" A learning guarantee says that no auditor in a specified family can find a profitable deviation.
Different choices of the auditor's test family lead to different concepts: external regret, internal and swap regret, equilibria in games, calibrated forecasting, multi-group guarantees, omniprediction, and obedience in strategic settings. The course develops this common language from the simplest deviation tests to richer tests involving groups, forecasts, and strategic agents.

Prerequisites

Probability at the level of concentration inequalities, linear algebra, and mathematical maturity. No prior background in online learning, convex optimization, or game theory is assumed. The small amount of convex analysis needed for the course will be introduced when it is first used.

Key ideas

Regret. How much better the learner could have done under some simple change to its past behavior. No-regret means this improvement is negligible over time.

Deviation. A proposed change to behavior, such as "whenever you played action a, play action b instead."

Test class. The family of deviations or checks that the learner is required to pass. Larger test classes give stronger guarantees.

Equilibrium. A way of playing a game in which no player can benefit from the allowed deviations.

Calibration. A forecast is calibrated if, among the rounds where it predicts probability p, the event happens about a p fraction of the time.

Part I: Regret and games (Weeks 1–4)

Week 1. The experts problem. Prediction with expert advice; external regret; Hedge / multiplicative weights; potential-function regret proof; lower bound sketch. *Thread:* external regret is the first no-deviation guarantee: no fixed expert would have been much better in hindsight.

Week 2. Zero-sum games and minimax. Two-player zero-sum games; the minimax theorem via no-regret dynamics; approximate Nash equilibria; no-regret vs. no-regret and no-regret vs. best-response dynamics; convergence of average play. Brief remarks on GANs, with the limitation of the analogy stated explicitly. *Thread:* even the simplest no-regret guarantee already proves a central theorem in game theory.

Week 3. Internal and swap regret. The deviation-family viewpoint; external vs. internal vs. swap regret; separations; reduction from swap regret to external regret; fixed-point / stationary-distribution computation. *Thread:* we move from fixed deviations to conditional deviations such as "when I played a, I should have played b."

Week 4. General-sum games. Definitions and separations among Nash equilibrium, correlated equilibrium (CE), and coarse correlated equilibrium (CCE); external regret $\Rightarrow$ CCE; internal/swap regret $\Rightarrow$ CE; why no-regret dynamics do not generally converge to Nash equilibrium. *Thread:* the equilibrium hierarchy is the deviation hierarchy in game form.

Part II: From actions to forecasts (Weeks 5–7)

Week 5. Online gradient descent. Convex sets and functions; subgradients; projections; online gradient descent and its regret proof. OMD/FTRL will be provided as an optional reading handout rather than treated as a main lecture. *Thread:* vector-valued guarantees require a little geometry, so we introduce only the tools needed for the next week.

Week 6. Blackwell approachability. Vector-valued online learning; approachable target sets; halfspace and separation-oracle characterizations; reduction to scalar regret; the auditing-oracle and audit-and-fix viewpoint. *Thread:* approachability is the general machine for driving many violations to zero at once.

Week 7. Calibration and proper scoring. Calibration of probabilistic forecasts; calibration via approachability; calibration via internal regret; proper scoring rules as incentives for truthful belief reporting; what calibration does and does not certify. *Thread:* we now audit forecasts rather than actions. The checks are conditioned on the forecaster's own predictions.

Part III: Richer auditors (Weeks 8–9)

Week 8. Multiaccuracy and multicalibration. Residual errors; multiaccuracy as small residual correlation with every group in a class; multicalibration as calibration within many groups and prediction levels; subgroup and subsequence guarantees; audit-and-fix algorithms. *Thread:* calibration checks the forecaster's own level sets; multicalibration allows outside auditors to choose the groups.

Week 9. Omniprediction. Omniprediction as one forecast serving many downstream users; how calibration plus multiaccuracy can imply good performance for many convex losses; proof sketch for a representative setting; scope and limitations. *Thread:* passing audits can make a forecast safe to use for many downstream decision problems.

Part IV: Strategic auditors (Weeks 10–12)

Week 10. Incentives, obedience, and manipulation. [core] Mechanism design as rule design: payments, losses, recommendations, and information structures that make desired behavior incentive-compatible. Learning algorithms as adaptive response under these rules. CE as obedient recommendations; no-swap-regret as empirical obedience; strategic manipulation of learning algorithms and what no-swap-regret does and does not protect against. *Thread:* the CE condition from Week 4 becomes an incentive condition: no player wants to disobey the recommendation.

Week 11. Non-collusion and calibrated best response. [core] Algorithmic pricing and the collusion question; calibrated best response; non-collusion audits as behavioral tests run by a regulator; how failure of such tests provides evidence against plausible non-collusion. *Thread:* the auditor from approachability, the conditional checks from calibration, and the incentive viewpoint from obedience meet in one applied question.

Week 12. Commitment and persuasion. [survey] Stackelberg commitment; security games; Bayesian persuasion; information structures; Bayes plausibility; receiver obedience. Pointers to the literature and open questions at the boundary of learning and information design. *Thread:* now the designer moves first and shapes the information or rules under which agents respond.