Inverse Decision Making

- So far, the focus has been on planning:
  - Making good decisions
  - Planning ahead
  - Creating strong policies

- Let’s focus now on the agent’s themselves. In particular we are interested in the following questions:
  - What do the agents want? What are they trying to accomplish in the environment?
  - Do they even have preferences?
  - What are preferences?
  - Can we “learn” their preferences? How?

- These are the question we will try to answer on the remainder portion of our lectures!

Preferences and Utility Functions

- Let's first tackle the notion of preferences.

- Abstractly, agents (humans and machines) have preferences:
  - Humans naturally prefer certain things over others
  - Machines have preferences over outcomes, due to their programming

- Formally, a preference is a **partial order.**
  - like positive semi-definiteness for matrices (it is a partial order)

- Given two elements x_1, x_2 belonging to some set X, we say that if an agent (weakly) prefers x_1 over x_2 then we write:

$$x_1 \succeq x_2$$

- And if we are indifferent between x_1 and x_2 we write: $x_1 \sim x_2$

Preferences and Utility Functions

- In order for the partial order to be useful we need to “quantify” the ordering somehow.

- That is the role of the Utility Function.

- Consider again the two elements x_1 and x_2. We will say that an agent (weakly) prefers x_1 over x_2 if and only if there is an Utility function $\mathrm{U}(x)$ such that:

$$U(x_1) \geq U(x_2) \Leftrightarrow x_1 \succeq x_2$$

- Hence the Utility function enforces an **ordering** among elements:
  - It does not really matter that the utility is 1 or 10
  - All it matters is the order (whether it is bigger or smaller than something else)

Preferences and Utility Functions

- One good question to make is whether it even makes sense to represent an agent's preferences by an Utility Function.
- It turns out that if the preferences (i.e.: the partial order) satisfies some conditions we can, in fact, represent the agent's preferences by an Utility Function.
- This the **Von Neumann–Morgenstern Utility Theorem.**
- In particular, the following conditions are essential:

$$x_1 \succ x_2, \quad x_1 \prec x_2, \text{or} \quad x_1 \sim x_2 \quad, \forall x_1, x_2 \in X$$

Complete Preferences

$$\text{If } x_1 \succeq x_2 \text{ and } x_2 \succeq x_3, \text{ then } \quad x_1 \succeq x_3 \quad, \forall x_1, x_2, x_3 \in X$$

Transitive Preferences

Preferences and Utility Functions

- Preferences that are both complete and transitive are called **Rational.**

- Hence, in this definition, a **rational agent** is the one that have complete and transitive preferences.

- And any rational agent can have their preferences represented by an Utility Function.

- We will work with Utility Functions, so we will assume rationality for any agent we consider.

Preferences and Choices

- Preferences cannot be observed:
  - They are "inside" a person's mind

- What we can observe are the person's choices (or actions).

- We can then use the information from those choices/actions to infer the agent's preferences.

- Of course, we may think: "What if we just ask the agent directly what are their preferences?"
  - (1) The agents can lie (they may not tell the truth)
  - (2) The agents may not be able to answer due to uncertainty in the environment

Example: Contract Selection

- One of the earliest methods of interfering preferences is by leveraging the concept of “self-selection”.
- We provide the agent a set of possible choices and ask the agent to select one of them.
- If we craft the choice set carefully, we can induce the agent to reveal to us their preference, even if they are not willing.
- This is the “self-selection” phenomenon: the agent’s select an option that reveals their underlying preferences.

Example: Contract Selection

- Suppose we own a company and we would like to hire new employees.

- Our goal is to maximize our profit and our profit function is given as follows:

$$P(q) - t, \quad P(q) = 10\sqrt{(}q)$$

- Where:
  - q is the amount of "work" done by the employees.
  - t is their salary

- Now, assume, there are two types of workers:
  - (1) a worker who is a perfect fit for the job (called the "efficient" worker).
  - (2) a worker who is not (called the "inefficient" worker).

Example: Contract Selection

- Suppose the workers have preferences over pairs of "work" and salaries. So if the worker (weakly) prefers the work-salary pair (q_1, t_1) over (q_2, t_2) we say:

$$(q_1, t_1) \succeq (q_2, t_2)$$

- Under our Rationality assumption, suppose that the workers utility functions are given by:

$$U_i(q,t) = t - \theta_i q - F \quad \forall i \in \{E, I\}$$

- Where θ_i is a scalar that represents the agent's **private information (type)**
  - We let θ_E denote the type of the efficient worker and θ_I the type of the inefficient worker.
  - And $\theta_I > \theta_E$
  - F is the utility of staying at home(!)

Example: Contract Selection

- Now the problem is to design a hiring offer where we are able to hire the efficient worker.

- Suppose $F = 20$ and $\theta_E = 0.25$ and $\theta_I = 0.30$

- Let's suppose that our hiring process is to just "ask" the worker what is their type. And then we offer a contract of work-salary (q, t) to them.

- First suppose they tell the truth (or we have perfect information about their types)

- Then, in order to hire the workers we need to make sure we make then an offer such that:

$$U_I(q_I, t_I) = t_I - 0.30q_I \geq 20$$
$$U_E(q_E, t_E) = t_E - 0.25q_E \geq 20$$

Participation Constraints

Example: Contract Selection

- As the company, it is enough if we offer a contract such that each agent's is indifferent between coming to work or staying at home.

- That is we would offer a contract (q, t) such that:

$$t_i - \theta_i q_i = 20 \quad , \forall i \in \{E, I\}$$

- Then we wish to maximize:

$$P(q_i) - t_i = 10\sqrt{q_i} - t_i = 10\sqrt{q_i} - \theta_i q_i - 20$$

- Taking the derivative and setting it to zero, gives the following set of contracts:

$$(q_E, t_E) = (400, 120) \qquad (q_I, t_I) = (277, 103.1)$$

Example: Contract Selection

- Now, let's see what happens if we **do not know** the workers type and we offer both choices of contracts:

$$(q_E, t_E) = (400, 120) \qquad\qquad (q_I, t_I) = (277, 103.1)$$

- Note that the inefficient worker will still select the contract intended to them (self-select):

$$t_E - \theta_I q_E - F = -20 \qquad\qquad t_I - \theta_I q_I - F = 0$$

- But the efficient worker will **not** prefer the contract intended to them:

$$t_I - \theta_E q_I - F = 13.85 \qquad\qquad t_E - \theta_E q_E - F = 0$$

- So our set of choices, is not good: we are not able to infer the workers type by their contract choices

Example: Contract Selection

- Let's design a new set of contracts.
- Since the types are unknown, suppose there is $v = 50\%$ chance of a worker that we offer a contract to be efficient.
- In order to **induce** self-selection (or in order words, **truth-telling**) we need make sure:

$$t_I - \theta_I q_I - F \geq t_E - \theta_I q_E - F$$
$$t_E - \theta_E q_E - F \geq t_I - \theta_E q_I - F$$

Incentive Compatibility Constraints

- Those enforce self-selection when offering the set of contract choices.
- Let's put everything together into an optimization problem.

Example: Contract Selection

- The company finds the set of contract choices by solving the following optimization problem.

$$\max_{(q_I,t_I),(q_E,t_E)} \quad v(P(q_E)-t_E)+(1-v)(P(q_I)-t_I)$$

$$\text{s.t.:}$$

Incentive Compatibility Constraints

$$\begin{cases} t_I-\theta_I q_I - F \geq t_E - \theta_I q_E - F \\ t_E - \theta_E q_E - F \geq t_I - \theta_E q_I - F \end{cases}$$

Participation Constraints

$$\begin{cases} t_I-\theta_I q_I - F \geq 0 \\ t_E - \theta_E q_E - F \geq 0 \end{cases}$$

- The solution can be derived from the optimality conditions (KKT conditions)

Example: Contract Selection

- But the problem is nice enough that we can argue which constraints are binding or not.

- First, observe that we can "set" the payoff of the inefficient agent to zero.
  - This is often the case: the lowest type obtain zero utility in the final setting
  - This relates to the notion of utility functions reflecting partial orders

- Then the inefficient worker participation constraints will be binding:

$$t_I - \theta_I q_I - F = 0 \implies t_I = \theta_I q_I$$

- Next, as we saw, the efficient will attempt to lie, if we do not provide them with an incentive (that is a positive utility) to behave truthfully. Hence it must follow that:

$$t_E - \theta_E q_E - F > 0$$

Example: Contract Selection

- Now let's focus on the incentive compatibility constraints:
  - For the inefficient worker, we do need to be concerned, as we saw, they do not select the contract meant for the efficient agent
  - For the efficient worker we do: we have to ensure the incentive compatibility constraint holds.

- Hence let's suppose we make the efficient worker's IC constraints binding:

$$t_E - \theta_E q_E - F = t_I - \theta_E q_I - F \implies t_E = \theta_E q_E + (\theta_I - \theta_E) q_I + F$$

- And the inefficient worker's IC constraints are not:

$$0 > t_E - \theta_I q_E - F$$

Example: Contract Selection

- With this (informal) argument, we write an optimization problem with just equalities:

$$\max_{(q_I, t_I), (q_E, t_E)} v(P(q_E) - t_E) + (1 - v)(P(q_I) - t_I)$$

$$\text{s.t.:}$$

$$t_I = \theta_I q_I + F$$

$$t_E = \theta_E q_E + (\theta^I - \theta^E) q_I + F$$

- Substituting t_I and t_E in the objective function, makes the entire problem unconstrained:

$$\max_{q_I, q_E} v(P(q_E) - \theta_E q_E - (\theta^I - \theta^E) q_I - F) + (1 - v)(P(q_I) - \theta_I q_I - F)$$

- Now we can take the derivative w.r.t. q_I and q_E and set them to zero.

Example: Contract Selection

- That leads us to:

$$P'(q_E) = \theta_E \qquad\qquad P'(q_I) = \theta_I + \frac{v}{1-v}(\theta_I - \theta_E)$$

- If we select q_E^* and q_I^* that solve the above equations, and let:

$$t_I = \theta_I q_I^* + F$$
$$t_E = \theta_E q_E^* + (\theta^I - \theta^E) q_I^* + F$$

- Then we have the solution to the original problem

- It is a good exercise to show that this point satisfy the KKT conditions.

Example: Contract Selection

- Now back to the numerical example. Recall that $\mathrm{v} = 0.5$, $(\theta_I, \theta_E) = (0.30, 0.25)$ and we have:

$$P(q) = 10\sqrt(q)$$

- So:

$$P'(q_E) = \theta_E \implies \frac{5}{\sqrt{q_E}} = 0.25 \implies q_E^* = 400$$

$$P'(q_I) = \theta_I + \frac{v}{1-v}(\theta_I - \theta_E) \implies \frac{5}{\sqrt{q_I}} = 0.35 \implies q_I^* = 204.08$$

- Then, since $F = 20$:

$$t_I = \theta_I q_I^* + F = 81.22$$

$$t_E = \theta_E q_E^* + (\theta^I - \theta^E) q_I^* + F = 130.20$$

Inference with Rational Agents

- As the previous example highlights, when inferring private information from agents, we need to design a scheme that ensures truthful behavior.

- In order to describe the problem formally, suppose the rational agent seeks to maximize their utility function:

$$u_i^* = \arg \max_{u \in U(x)} \{U(x_i, u; \theta)\}$$

- Where:
  - x_i is the state of the "environment"
  - u_i^* is the action/control taken by the agent
  - θ is the type (private information) that we wish to infer

- Note that the utility function can be seen as the "cost-to-go" (value function) of the agent.

Inference with Rational Agents

- Suppose that we get to observe to observe the agent interacting with the environment. And we collect observations of state-action pairs:

$$((x_1, u_1^*), (x_2, u_2^*), ..., (x_n, u_n^*))$$

- Let's assume that the agent behaves optimality. Hence we can pose our inference as the following optimization problem:

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}} \quad 0$$

$$\text{s.t.:}$$

$$u_i^* = \arg \max_{u \in U(x_i)} \{U(x_i, u, \theta)\}, \forall i \in \{1, ..., n\}$$

- Where Θ is some bounded set of possible type values.

Inference with Rational Agents

- This problem is called an **Inverse Optimization Problem.**

- It is a feasibility problem, where the feasible region is given by solutions of optimization problems.

- These in general are very hard to solve, because a computer cannot handle constraints that themselves are optimization problems.

- We will study a simplified version of the problem:
  - The set $U(x)$ will be a polyhedron
  - The utility function will be quadratic and strictly concave.
  - This means that agent is solving some convex Quadratic Optimization with a unique maximizer for every x_i
  - For example: they may be solving a linear MPC problem (or LQR if unconstrained)

Inference with Rational Agents

- For a given state x_i, the agent's problem will be given by:

$$\max_u \{u^\top R u + x_i^\top F^\top u + c^\top u\}$$

$$\text{s.t.: } Ax_i + Bu \leq d$$

- Where $R \prec 0$. So $\theta = (R, F, c)$.

- In addition, the problem is a strictly concave problem and it has a unique optimal solution. (note that x_i is fixed here)

- Let's suppose the agent solve the above problem and let λ_j be the Lagrange multiplies associated with constraint j:

$$a_j^\top x_i + b_j^\top u \leq d_j \qquad (\lambda_j)$$

Inference with Rational Agents

- Let $I^*(x_i)$ be the set of constraints that are active. Then we characterize the optimal solution as:

$$2Ru_i^* + Fx_i + c = \sum_{j \in I^*(x_i)} \lambda_j b_j$$

- Note that the active set of constraints depend on the state x_i.

- For a given state-control pair (x_i, u_i) it is easy to verify which constraints are active in order to find the set $I^*(x_i)$:

$$j \in I^*(x_i) \text{ if and only if } a_j^\top x_i + b_j^\top u_i^* = d_j$$

- Again, note that x_i and u_i^* are given here (they are “data”).

Inference with Rational Agents

- Hence the Inverse Optimization Problem can be written as:

$$\hat{\theta} = (\hat{R}, \hat{F}, \hat{c}) = \underset{(R,F,c,\lambda)}{\operatorname{argmin}} \quad 0$$

$$\text{s.t.:}$$

$$2Ru_i^* + Fx_i + c = \sum_{j \in I^*(x_i)} \lambda_j^i b_j, \ \forall i \in \{1, ..., n\}$$

$$\lambda_j^i \geq 0, \quad \forall j \in I^*(x_i), \forall i \in \{1, ..., n\}$$

$$R \preceq 0$$

- Note that the Langrange multipliers λ are *decision variables*, even though our goal is to just estimate (R, F, c)

Inference with Rational Agents

- There is an issue with previous formulation.

- Namely there can be multiple optimal solutions: For example the trivial solution $R = 0, F = 0, c = 0$ solves the Inverse Optimization Problem above.

- This relates back to the notion of the Utility Function capturing only the ordinal relations between preferences (the ranking).

- So we need to set a baseline, or normalize the utility function in order to obtain meaningful results.

- One possible normalization is to change the semidefinite constraint to:

$$R \preceq \mathbb{I}$$

Inference with Rational Agents

- So the inference problem becomes:

$$\hat{\theta} = (\hat{R}, \hat{F}, \hat{c}) = \underset{(R,F,c,\lambda)}{\operatorname{argmin}} \quad 0$$

$$\text{s.t.:}$$

$$2Ru_i^* + Fx_i + c = \sum_{j \in I^*(x_i)} \lambda_j^i b_j, \; \forall i \in \{1, ..., n\}$$

$$\lambda_j^i \geq 0, \quad \forall j \in I^*(x_i), \, \forall i \in \{1, ..., n\}$$

$$R \preceq \mathbb{I}$$

- This problem is tractable to solve by a computer (it is convex).

Inference with Rational Agents

- Often, we do not know the parametric form of the agent's utility.
- We can instead suppose there are a set of performance criterions that the agent will try to optimize over and we wish to find out the "ranking" of those criterions
- That is how the criterions are ranked in order of importance.
- Suppose the utility function of the agent is given by the following:

$$U(x,u) = \sum_{j=1}^{K} k_j f_j(x,u)$$

Weighted sums of the criterions $f_j(\cdot,\cdot)$, concave in u

$\theta = (k_1, k_2, \dots, k_K)$ is the private information of the agent

Inference with Rational Agents

- So the inference problem becomes:

$$\hat{\theta} = (\hat{k_1}, ..., \hat{k_K}) = \underset{(k_1, ..., k_K)}{\operatorname{argmin}} \quad 0$$

$$\text{s.t.:}$$

$$\sum_{j=1}^{K} k_j \nabla_u f_j(x_i, u_i^*) = \sum_{j \in I^*(x_i)} \lambda_j^i b_j, \, \forall i \in \{1, ..., n\}$$

$$\lambda_j^i \geq 0, \quad \forall j \in I^*(x_i), \, \forall i \in \{1, ..., n\}$$

$$k_1 \geq 1$$

- Where $k_1 \geq 1$ is our normalization choice. This problem is still tractable (note that the equality constraint is linear on $(k_1, \dots, k_K)$

Inference with Rational Agents

- Note that we made a very strong assumption:

- Namely that the agent is able to solve their optimization problem exactly.

- As we saw throughout the course, agent’s often cannot solve their planning problems to optimality, often having to make due with suboptimal solutions.

- Alternatively, they may solve the problem exactly, however we collect the data with some i.i.d. noise.

- Let’s see how can we adapt our inference formulation to treat this case.

Inference with Rational Agents

- Typically, sub-optimality can be characterized by violations of the KKT conditions (if the problem is convex).

- In our scenario, that means the following:
  - (1) the agent always compute a feasible control/action u_i
  - (2) the control/action choice may be suboptimal
  - (3) given the active set of constraints $I(x_i)$, we may have violations of the gradient equation.

- The idea is then to add a loss-term in the objective (instead of just 0) that minimized the degree of sub-optimality in the agent's problem.

Inference with Rational Agents

- Hence the problem becomes:

$$\hat{\theta} = (\hat{k_1}, ..., \hat{k_K}) = \underset{(k_1, ..., k_K)}{\operatorname{argmin}} \quad \sum_{i=1}^{n} ||r_i||_2^2$$

$$\text{s.t.:}$$

$$\sum_{j=1}^{K} k_j \nabla_u f_j(x_i, u_i) - \sum_{j \in I(x_i)} \lambda_j^i b_j = r_i, \forall i \in \{1, ..., n\}$$

$$\lambda_j^i \geq 0, \quad \forall j \in I(x_i), \forall i \in \{1, ..., n\}$$

$$k_1 \geq 1$$

- So $(r_1, \ldots, r_n)$ are the residuals of sub-optimality in the agent's planning for the data points $((x_1, u_1), \ldots, (x_n, u_n))$

Example: Hypothesis Testing

- One interesting question to ask is given observed state-action pairs data points:

$$((x_1, u_1), ..., (x_n, u_n)), \forall i \in \{1, ..., n\}$$

- Is to ask if the agent is behaving optimality or not in their environment

- Let's say that if they are the residuals $(r_1, \dots, r_n)$ should be i.i.d. white noise with some variance Σ.

- Then if we solve the inverse optimization problem as before, and compute the estimated residuals:

$$\hat{r}_i = \sum_{j=1}^{K} \hat{k}_j \nabla_u f_j(x_i, u_i) - \sum_{j \in I(x_i)} \lambda_j^i b_j, \forall i \in \{1, ..., n\}$$

Example: Hypothesis Testing

- We can then formulate a hypothesis testing to test whether:

$$(\hat{r}_1, ...\hat{r}_n) \sim \mathcal{N}(0, \Sigma),\ \forall i \in \{1, ..., n\}$$

- Or not.

- The correctness and consistency of such estimation and hypothesis testing is beyond our scope.
    - It is work based on "Inverse Optimization with Noisy Data, Aswani et al, 2017."

- It turns out that on the tractable scenario we considered (strictly concave utilities, polyhedral feasible regions, etc), such hypothesis testing is indeed correct and our residual estimates are consistent.