1. (5pts) Let $f(x, y) = 4x + 2y - x^2 + xy - y^2$ on $(x, y) \in \mathbb{R}^2_+$.

   (a) Is f concave? Prove or disprove.

   (b) Does f have a global maximum? If yes, prove this and find the global maximum. Otherwise, disprove this.

   (c) Does f have a global min? If yes, prove this and find the global minimum. Otherwise, disprove this.

(a) show $D^2 f(x,y)$ is NSD by checking principal minors

(b) Unconstrained optimization w/ $f: E \subset \mathbb{R}^n \to \mathbb{R}$, f C^2 & E nonempty open & convex

f concave?

Yes

$\nabla f(x^*) = 0$

$\Rightarrow x^*$ is global max

(prop 5, module 6)

No

If a global max x^* exists, then it must satisfy:

① $\nabla f(x^*) = 0$ (FONC)

② $D^2 f(x^*)$ is NSD (SONC)

If no pts satisfy ① & ②, then $\exists$ no global max.

If we replace NSD w/ ND in ②, then we have the sufficient conds for a pt to be a local max.

(Prop 59, module 5)

(c) For global min ...

f ~~concave~~ convex ?

Yes

$\nabla f(x^*) = 0$

$\Rightarrow x^*$ is global ~~max~~ min

(prop 5, module 6)

No

If a global ~~max~~ min x^* exists, then it must satisfy:

1. $\nabla f(x^*) = 0$ (FONC)
2. $D^2 f(x^*)$ is ~~NSD~~ PSD (SONC)

If no pts satisfy 1 & 2, then $\exists$ no global ~~max~~ min.

If we replace ~~NSD~~ PD w/ ~~ND~~ PSD in 2, then we have the <u>sufficient conds</u> for a pt to be a <u>local</u> ~~max~~ min

(Prop 59, module 5)

SONC requires $D^2 f(x^*)$ PSD $\Rightarrow$ (x)

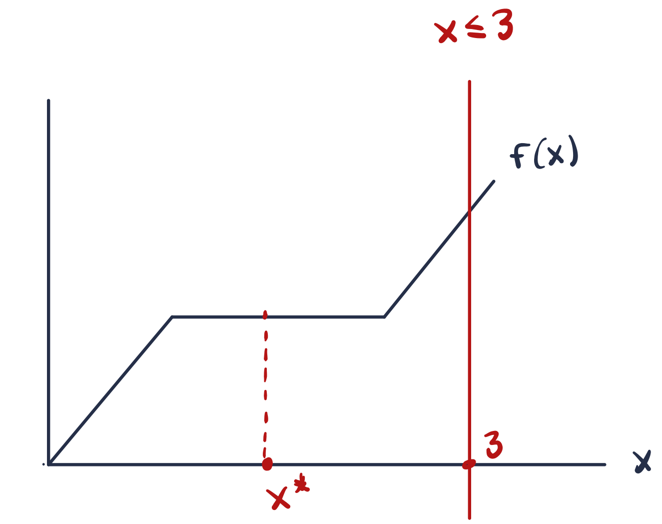
2. (5pts) Let $f : \mathbb{R}_+ \to \mathbb{R}$ be quasiconcave, and let $g : \mathbb{R}_+ \to \mathbb{R}$ be quasiconvex. Consider the *quasiconcave program*,

$$\max_{\{x \in \mathbb{R}_+ : g(x) \leq 0\}} f(x).$$

(a) Prove or disprove: if x^* is a local (constrained) maximum, it is also a global (constrained) maximum.

(b) Prove or disprove: if f is strictly quasiconcave, then the constrained global maximum is unique.

(a)



$f(\alpha x_1 + (1-\alpha)x_2) \geq \min\{f(x_1), f(x_2)\}$

Takeaway: quasiconcavity doesn't guarantee global solution. It's weaker than concavity.

But if you have a unique local max, then quasi-concavity tells you that's also the unique global max.

Proof: Define $C \equiv \{x \in \mathbb{R}_+ : g(x) \leq 0\}$ set of feasible x

If x^* is the unique constrained local max. Then $\exists\, \delta > 0$ s.t. $\forall\, x \in B_\delta(x^*) \cap C$, $f(x) < f(x^*)$.

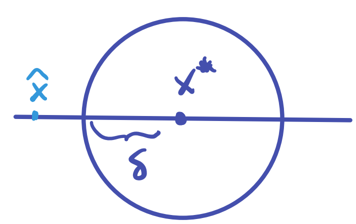
If x^* is not a global max. Then $\exists\, \hat{x} \in C$ s.t. $f(\hat{x}) > f(x^*)$.

Then $\forall\, \alpha \in [0,1]$, $f(\lambda \hat{x} + (1-\lambda)x^*) \geq \min\{f(\hat{x}), f(x^*)\} = f(x^*)$

Since g is quasiconvex, C is convex (Module 3, Ex. 8)

$\Rightarrow \lambda \hat{x} + (1-\lambda)x^* \in C$. To make it $\in B_\delta(x^*)$, shrink λ arbitrarily close to 0: $\lambda = \dfrac{\delta}{2\|\hat{x} - \bar{x}\|}$ will do.

Then $\lambda \hat{x} + (1-\lambda)x^* \in B_\delta(x^*) \cap C$ but $f(\lambda \hat{x} + (1-\lambda)x^*) \geq f(x^*)$. Contradiction.



(b) Strict quasiconcavity $\Rightarrow$ global max x^*, if exists, is unique.

If not. Then $\exists\ \hat{x} \in C$ s.t. $f(\hat{x}) = f(x^*)$

f is strictly quasiconcave $\Rightarrow f(\alpha\hat{x} + (1-\alpha)x^*) > \min\{f(\hat{x}), f(x^*)\} = f(x^*)$

$\alpha\hat{x} + (1-\alpha)x^* \in C$, which contradicts x^* is a constrained global max

Note that strict quasiconcavity also gives local max $\Rightarrow$ global max

proof is similar to (a).

3. (5pts) Loosely speaking, the implicit function theorem gives conditions under which the set of solutions to a (potentially nonlinear) equation (or system of equations) locally takes on a particular form, i.e., some of the variables can be expressed as a differentiable function of the others.

   (a) Let $f : U \subset \mathbb{R}^2_+ \to \mathbb{R}$, where U is open, and consider the equation $f(x, y) = 0$. Suppose $(x_0, y_0) \in U$ solves this equation. Precisely state the implicit function theorem for this case, including all the necessary assumptions and the conclusions.(+2)
   Given what's already assumed, if we further impose

   i. f is continuously differentiable;
   ii. $f_y(x_0, y_0) = \frac{\partial f}{\partial y}\Big|_{(x_0,y_0)} \neq 0$

   Then $\exists \epsilon > 0$ and a continuously differentiable function $h : B_\epsilon(x_0) \to \mathbb{R}_+$ such that

   i. $f(x, h(x)) = 0 \forall x \in B_\epsilon(x_0)$ and $B_\epsilon(x_0, y_0) \cap \{(x, y) \in U : f(x, y) = 0\} = graph(h)$
   ii. $h'(x) = -\frac{f_x(x_0,y_0)}{f_y(x_0,y_0)}$

   (b) Give an example of a smooth function f and a point (x_0, y_0) where the key necessary condition of the implicit function theorem fails, but the conclusion that the set of (x, y) that solve $f(x, y) = 0$ is locally the graph of a (not necessarily differentiable) function of x still holds. (+3)
   Hint: Draw a picture if you cannot immediately think of a numerical example.
   Consider $f(x, y) = y^3 - x$. The point $(0, 0)$ solves $f(x, y) = 0$. The function $y(x) = x^{\frac{1}{3}}$ is a local solution (in fact, a global solution) to this equation. But $\frac{\partial f(x,y)}{\partial y}\Big|_{(0,0)} = 3y^2|_{(0,0)} = 0$ so IFT doesn't apply.

Use Prop 4 in module 6 : f concave iff $D^2f(x)$ is NSD
$\Rightarrow$ show diagonal matrix $D^2f(x)$ is NSD iff its diagonal elements are all non positive.

4. (5pts) Let $f : \mathbb{R}^n \to \mathbb{R}$ be given twice continuously differentiable. Suppose moreover that the hessian, $D^2 f(x)$, is a diagonal matrix for every $x \in \mathbb{R}^n$, i.e., the off diagonal entries of $D^2 f(x)$ are 0.

   Prove or disprove: f is concave if and only if all the diagonal entries of $D^2 f(x)$ are non-positive for all $x \in \mathbb{R}^n$.

   Remark: If we interpret f to be a production function, we would say there is no complementarity between the inputs if all the cross partials are zero.

   Suppose all the diagonal entries of $D^2 f(x)$ are non-positive for all $x \in \mathbb{R}^n$. Take any $v \in \mathbb{R}^n$. Then

$$\begin{aligned}
v^T D^2 f(x) v &= \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \begin{bmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \\
&= \begin{bmatrix} v_1 d_1 & \cdots & v_n d_n \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \\
&= \sum_{i=1}^{n} v_i^2 d_i \\
&\leq 0
\end{aligned}$$

   where $d_1, ..., d_n$ are the diagonal elements of $D^2 f(x)$. Hence $D^2 f(x)$ is NSD for all $x \in \mathbb{R}^n \implies f$ is concave by Proposition 4 in Module 6.

   Now suppose f is concave and $D^2 f(x)$ is a diagonal matrix. Then again by Proposition 4 in Module 6, $D^2 f(x)$ is NSD for all $x \in \mathbb{R}^n$. Take the ith standard basis vector $e^i \in \mathbb{R}^n$, i.e., e^i has all 0s except for the ith component, which is 1. Then $\forall i \in \{1, ..., n\}$ and $x \in \mathbb{R}^n$,

$$e^{i^T} D^2 f(x) e^i = d_i \leq 0$$

   Therefore all the diagonal entries of $D^2 f(x)$ are non-positive for all $x \in \mathbb{R}^n$.

5. (Extra Credit: 2 pts)

(a) Let $f : U \subset \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^n$, where U is open. Precisely state the implicit function theorem for f, including the key necessary assumptions and conclusions. (1 point)
Module 6, Proposition 22 (Implicit Function Theorem): Let $F : U \subset \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^n$ be continuously differentiable, where U is open. Consider the set $S \equiv \{(x, p) \in U : F(x, p) = y\}$. Let $(a, b) \in S$. Let $D_x F(a, b)$ denote the $n \times n$ matrix of partials of F with respect to only the first n variables, and suppose $D_x F(a, b)$ is invertible. There exists an $\epsilon > 0$ and a continuously differentiable function $h : B_\epsilon(p) \subset \mathbb{R}^k \to \mathbb{R}^n$ such that $F(h(p), p) = y$ for all $p \in B_\epsilon(p)$, and moreover, $B_\epsilon(a, b) \cap S = \text{graph}(h)$. Finally $Dh(b) = -D_x F(a, b)^{-1} D_p F(a, b)$.

(b) State a general parameterized constrained optimization program with (1) an objective function that has n choice variables and k parameters, and (2) m constraint functions of the same dimensionality. Precisely state the envelope theorem for this constrained maximization program. (1 point)
Let $f : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$, and $g_i : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$ for all $i \in \{1, \ldots, m\}$; f and all g_i are continuously differentiable. Let $y_1, \ldots y_m \in \mathbb{R}$. We let x denote a vector in $\mathbb{R}^n$ and p denote a vector in $\mathbb{R}^k$. Consider the following problem:

$$\text{Maximize } f(x, p) \text{ with respect to } x\text{, subject to } g_i(x, p) \leq y_i \text{ for all } i \in \{1, \ldots, m\}. \quad (1)$$

Define $V(p) \equiv \max_{x : g(x,p) \leq y} f(x, p)$.
Module 6, Proposition 25 (Envelope Theorem): Suppose that for some open subset $A \times B \subset \mathbb{R}^n \times \mathbb{R}^k$, the solution to Problem (1) is the graph of some continuously differentiable function $h : B \to A$. Let $I(h(p))$ denote the set of binding constraints at the solution $h(p)$, and suppose the KKT conditions hold at $h(p)$ for any $p = (p_1, \ldots, p_k) \in B$ with λ_j denoting the associated multipliers for each $j \in I(h(p))$. Then $\frac{\partial V(p)}{\partial p_i} = \frac{\partial f(h(p),p)}{\partial p_i} - \sum_{j \in I(h(p))} \lambda_j \frac{\partial g_j(h(p),p)}{\partial p_i}$

Dynamic Programming (Pls read Krep's Micro Foundations, A.6 and Chloe & Helen's notes. Also watch the awesome vids in Canvas announcement!)

- Timing : a discrete series of dates where the agent takes an action
- Actions a_t : The action the agent chooses at t
  $a_t \in A_t$, where A_t is set of possible actions at t.
  Note both a_t & A_t can depend on h_t.
- Histories h_t : Everything known to the agent at t before making a decision. Usually $h_t = \{a_0, \cdots, a_{t-1}\}$
  $h_t \in H_t$, where H_t is the set of all possible histories up to t. H (w/o subscript t) denotes the set of all possible histories of the game.
- Strategy σ_t : $\sigma_t : H_t \to A_t^{H_t}$; given h_t, specifies what action a_t to take at each time period t ; σ w/o the subscript t specifies what action to take at h_t $\forall t$. Define Σ_t as the set of all possible strategies at t, and Σ the set of all strategies (cross product of $\{\Sigma_t\}$ $\forall t$)

- utility func : $u : H \to \mathbb{R}$

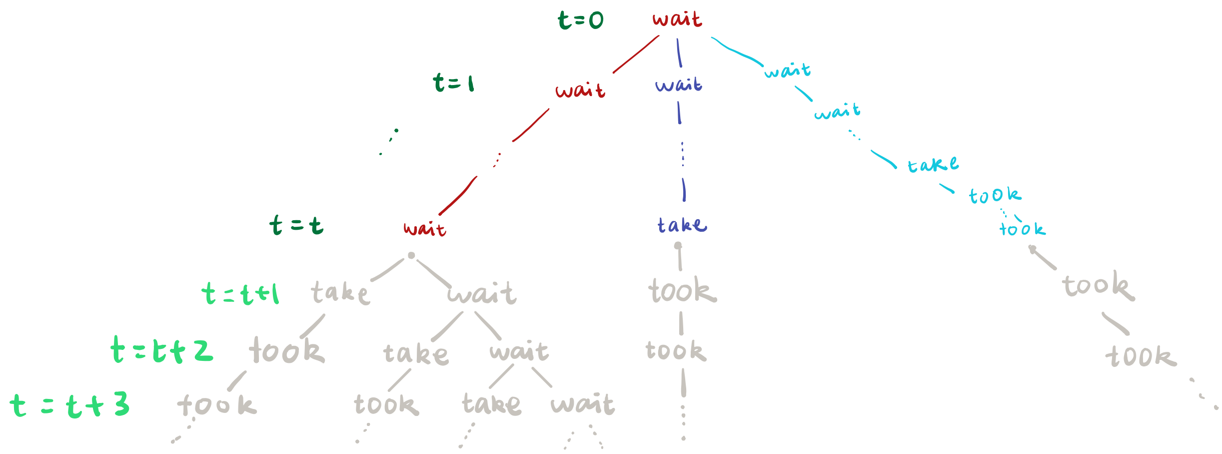
Given history h_t, the payoff of continuing w/ strategy σ starting at t is $f_t^\sigma(h_t) := E^\sigma[u(h) \mid h_t]$

By law of iterated conditional expectation,

$$\begin{aligned} f_t^\sigma(h_t) &:= E^\sigma[u(h) \mid h_t] \\ &= E^\sigma\{\underbrace{E^\sigma[u(h) \mid h_{t+1}]} \mid h_t\} \quad \text{since } h_t \subseteq h_{t+1} \\ &= E^\sigma\{f_{t+1}^\sigma(h_{t+1}) \mid h_t\} \end{aligned}$$

$\boxed{\text{Ex}}$ You've waiting at $t=0$. At each $t \geq 1$, you choose an action.

1. If you've been waiting up until time t, you can choose to either continue to wait, or choose to take
2. If you've taken before t, then you stay took afterwards and can't switch back to wait



• Timing: $t = 1, 2, \cdots$ (infinite)

At t, there are 3 possible h_t: (wait, $\cdots$, wait)
(wait, $\cdots$, wait, take)
(wait, $\cdots$, take, took, $\cdots$, took)

• Action space: $A_t(h_t) = \{\text{take}, \text{wait}\}$
$A_t(h_t) = \{\text{took}\} = A(h_t)$

• One strategy can be: wait if $h_t = h_t$
took if $h_t = h_t$ & h_t

Note a strategy should spit out a <u>feasible</u> action $a_t \in A_t(h_t)$
<u>$\forall$ possible h_t</u>.

• Let's assume no uncertainty in payoff,
i.e., $f_t^\sigma(h_t)$ = a deterministic number following:
 • If you take at t, you get $\frac{t+1}{t+2}$
 • If you never take, you get 0.

<u>Quiz</u>: If $\sigma(h_t)$ is $\begin{cases} \text{wait if } h_t = h_t \\ \text{took if } h_t = h_t \text{ \& } h_t \end{cases}$, what's $f_t^\sigma(h_t)$?

Define the optimal value function by $f_t^*(h_t) = \sup_{\sigma \in \Sigma} f_t^\sigma(h_t)$

What's $f_t^*(h_t)$?

$\boxed{\text{Def}}$ A strategy $\hat{\sigma}$ is conserving if $f_t^*(h_t) = E^{\hat{\sigma}}[f_{t+1}^*(h_{t+1}) \mid h_t] \;\forall\, t, h_t$

Note $f_t^*(h_t) = \sup_{\sigma \in \Sigma} E^{\sigma}[f_{t+1}^*(h_{t+1}) \mid h_t]$ is the Bellman eqn. This says given any h_t, playing $\hat{\sigma}$ gives you the optimal expected payoff in subsequent time periods. (i.e., it "conserves" the optimal value)

What $\hat{\sigma}$ is conserving for this game? $\hat{\sigma}$ as specified above.

$$f_t^*(\underbrace{\text{wait}, \cdots, \text{wait}}_{h_t}) = E^{\hat{\sigma}}\big[f_{t+1}^*(\underbrace{\text{wait}, \cdots, \text{wait}, \text{wait}}_{h_{t+1} \text{ according to } \hat{\sigma}}) \mid h_t\big] = 1$$

$\boxed{\text{Def}}$ A Strategy is unimprovable if $\forall\; t \;\&\; h_t$,

$$f_t^{\hat{\sigma}}(h_t) = \sup_{\sigma \in \Sigma} E^{\sigma}[f_{t+1}^{\hat{\sigma}}(h_{t+1}) \mid h_t]$$

This says if at $t+1$ you use $\hat{\sigma}$, then using $\hat{\sigma}$ at t is optimal, i.e., if you know you'll use $\hat{\sigma}$ next period, deviating to any other σ gives you a lower expected payoff this period.

Is σ specified above unimprovable?

$f_t^{\hat{\sigma}}(\text{wait}, \cdots, \text{wait}) = 0$ since you never take.

Consider deviating to take at t (call this strat σ')

Then $f_t^{\sigma'}(\text{wait}, \cdots, \text{wait}) = \frac{t+1}{t+2}$ and $h_{t+1} = (\text{wait}, \cdots, \text{wait}, \text{take})$

Now, $E^{\sigma'}[f_{t+1}^{\hat{\sigma}}(h_{t+1}) \mid h_t] = \frac{t+1}{t+2} > f_t^{\hat{\sigma}}(h_t) = 0$

$\boxed{\text{Def}}$ A strategy σ is everywhere optimal if $f_t^{\sigma}(h_t) = f_t^*(h_t) \;\forall\, t, h_t$

What σ is optimal for this game?