Lecture 1: Sets and Numbers

Problem 1. Show that if $A \cup B = A$ and $A \cap B = A$, then $A = B$.

Problem 2. Show that $(A \setminus B) \cap C = (A \cap C) \setminus (B \cap C)$.

Problem 3. Let $b - a > 2$. Write down a simpler expression for the following sets:

a. $\bigcup_{n=1}^{\infty} [a + \frac{1}{n}, b - \frac{1}{n}]$.

b. $\bigcap_{n=1}^{\infty} (a - \frac{1}{n}, b + \frac{1}{n})$.

Problem 4. Use induction to solve the following problems.

a. Let $S_n = 1+3+5+\cdots+2n-1$. Find an expression for S_n and prove its correctness.

b. Prove that every finite set of integers has a smallest and a largest element.

Problem 5. Prove that, given a finite set X with n elements, the set $P(X)$ of all subsets of X has 2^n elements.

Hint. *If you feel like you need to practice proofs by induction, go ahead and do it by induction. There is also a combinatoric way to do it, using indicator functions.*

Note. *The set of all subsets of a set X is often denoted 2^X.*

Problem 6. Suppose the computer language you use does not know how to compute $n!$. In that case, you need to code a function $factorial : \mathbb{N} \to \mathbb{N}$ where $factorial(n) = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1$. How would you code it?

Hint. *Recursion.*

Problem 7. Consider the following statement: "If a polygon x is in collection A, then x is green." Write below the negation, the inverse, the converse and the contrapositive of that statement. Which of these is equivalent to the original statement?

Problem 8. A function $f : X \to \mathbb{R}$ is *bounded above* when its image $f(X) = \{f(x) \in \mathbb{R} : x \in X\}$ is a bounded set. In that case, we write $\sup f = \sup f(X)$. Prove that if $f, g : X \to \mathbb{R}$ are bounded, then their sum $f + g$ is also bounded and $\sup(f + g) \leqslant \sup f + \sup g$. State and prove an analogous result for the infimum.

Problem 9. Let X be a finite set. We say a binary relation $R \subset X \times X$ is *acyclic* when for any finite 'chain' $x_1 R x_2 R \cdots R x_n$ of elements of X, it is not the case that $x_n R x_1$.

Prove that if $R \subseteq X \times X$ is acyclic, then there exists $x \in X$ such that there is no $y \in X$ for which yRx.

Note. *The definition of an acyclic binary relation is not always the same. Pay attention to the precise definition in each context. The definition of R-unbeaten is sometimes defined as R-maximal. Again, pay attention to the context.*

Problem 10. Let $X_1, X_2, \ldots, X_n, \ldots$ be a family of non-empty sets. Prove or find a counterexample for the statements below.

a. If $X_1 \supseteq X_2 \supseteq \cdots \supseteq X_n$, with $n \in \mathbb{N}$, then $\bigcap_{i=1}^{n} X_i \neq \emptyset$.

b. If $X_1 \supseteq X_2 \supseteq \cdots \supseteq X_n \supseteq \cdots$, then $\bigcap_{i=1}^{\infty} X_i \neq \emptyset$.

Problem 11. Consider two arbitrary sets X and Y and a function $f : X \to Y$. In the items below, mark *True* or *False*, and justify your answer in no more than one line.

a. If X is countable and f is one-to-one (injective), then Y is countable.

b. If X is countable and f is onto (surjective), then Y is countable.

c. If Y is countable and f is one-to-one (injective), then X is countable.

d. If Y is countable and f is onto, then X is countable.

Problem 12. Prove that $2^{\mathbb{N}}$ is not countable. Can you generalize this for sets other than $\mathbb{N}$?

Hint. *Use Cantor's diagonal argument.*

Lecture 2: Topology of $\mathbb{R}^n$

Problem 13. Prove the following statements about open sets.

a. Let I be an arbitrary index set, $(A_i)_{i \in I}$ a family of sets where each A_i is open in $\mathbb{R}^n$. Then $\bigcup_{i \in I} A_i$ is an open set.

b. Let I be a finite index set, $(A_i)_{i \in I}$ a family of sets where each A_i is open in $\mathbb{R}^n$. Then $\bigcap_{i \in I} A_i$ is an open set.

c. $\mathbb{R}^n$ and $\emptyset$ are open in $\mathbb{R}^n$.

Problem 14. Can you write down in plain English the three statements in the previous problem?

Hint. *There are better ways to write that in English than simply translating to English all those symbols.*

Problem 15. Write formally the negation of the following statements:

a. $\forall \varepsilon > 0\ \exists \delta > 0$ such that $\|x - a\| < \delta \Rightarrow \|f(x) - f(a)\| < \varepsilon$.

b. $\forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}$ such that $\|x_n - a\| < \varepsilon\ \forall n > n_0$.

Problem 16. Prove that every convergent sequence is bounded.

Problem 17. Let $K \subset \mathbb{R}^n$, and $f : K \to \mathbb{R}^m$ be a continuous function.

a. If K is closed, but unbounded, show that $f(K)$ is not necessarily closed.

b. If K is bounded, but not closed, show that $f(K)$ is not necessarily bounded.

c. Show that if K is compact, then $f(K)$ is compact. Use that to prove Weirstrass' theorem: a real-valued function f defined on a compact set K has a maximizer and a minimizer, that is, there are $m, M \in K$ such that $f(M) \geqslant f(x) \geqslant f(m)$ for all $x \in K$.

Note. *Compact sets are what we call* topological invariants*: they are preserved under continuous transformations. Another example of topological invariant is connectedness.*

Problem 18.

a. Show that for $\lambda > 0$, $\sum_{n=0}^{\infty} \lambda^n = \frac{1}{1-\lambda}$.

b. Show that for $\lambda > 0$, $\sum_{n=1}^{\infty} \lambda^n = \frac{\lambda}{1-\lambda}$.

c. It is often handy to know a series of positive numbers that adds up to one. Using the formulas above, can you guess one such series?

d. Let's see an interesting example where having a series of positive numbers that adds up to one is useful. Fix $\varepsilon > 0$. Using the power series from the previous item, create a set that includes all the rational numbers between 0 and 1, but that has total "size" smaller than ε. Note that you didn't use any special property of the rational numbers besides the fact that they are a countable set. If you cannot solve this problem, instead the rationals, pick any large finite set of points in $[0, 1]$; if you managed that, go back to the case with the rationals.

e. Based on the previous item, answer the following question: if we draw a real number from the interval $[0, 1]$ "at random" (that is, using a uniform distribution), what is the probability that this number is a rational number?

Note. *A precise definition of what we mean by "size" would take us too far; formally, by size we mean* measure*, as defined in measure theory. Intuitively, for this exercise, all you need to know is that the measure of an interval is its length, and that the the measure of a union of sets A_i is no larger than the sum of the measures of each set A_i.*

Problem 19. Let $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ be two sequences. Show that if (x_n) is bounded and $y_n \to 0$, then $x_n y_n \to 0$.

Note. *A sequence $x_1, x_2, x_3, \ldots$ is often denoted simply x_n; if you do this, make sure there is no ambiguity between the sequence x_n and the n-th element of that sequence. It is safer to put parentheses around it. If there is ambiguity about the indexing set, write $(x_n)_{n\in\mathbb{N}}$. Do not write $\{x_n\}$ as curly brackets indicate sets: if $x_n = 1$ for all n, then $(x_n)_{n\in\mathbb{N}} = (1, 1, 1, \ldots)$, while $\{x_n\}_{n\in\mathbb{N}} = \{1\}$.*

Problem 20. Show that if $f : \mathbb{R} \to \mathbb{R}$ is monotonic, then the set of points where f is discontinuous is countable.

Problem 21. Let (x_n) be a sequence of positive real numbers. Show that if $\sum_{n=1}^{\infty} x_n < \infty$, then $x_n \to 0$

Problem 22. Let $X \subseteq \mathbb{R}^L_+$ be a set of alternatives and $u : X \to \mathbb{R}$ be a continuous utility function, and let R be the preference relation induced by R: xRy if and only if $u(x) \geqslant u(y)$. For every alternative $x \in X$, define

$$U(x) = \{y \in X : yRx\},$$

and

$$L(x) = \{y \in X : xRy\}.$$

Show that $U(x)$ and $L(x)$ are closed for all $x \in X$.

Problem 23. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a continuous function and $K \subset \mathbb{R}^n$ a connected set. Show that $f(K)$ is a connected set.

Problem 24. Let $f : [a, b] \to \mathbb{R}$ be a continuous function such that $f(a) < 0$ and $f(b) > 0$. Show that there exists $c \in (a, b)$ such that $f(c) = 0$.

Note. *This result is known as the* intermediate value theorem*.*

Hint. *Use the previous problem. If you did not solve the previous problem, you cannot use it.*

Problem 25. For every positive integer k and every positive real number x define, recursively, $h(x, 0) = 1$ and $h(x, k) = x^{h(x,k-1)}$. Let

$$f(x) = \lim_{k\to\infty} h(x, k)$$

Is there a solution to the equation $f(x) = 2$? Prove it. What about equations of the form $f(x) = s$ for some $s \in \mathbb{R}$?

Lecture 3: Linearity and Convexity

Problem 26. Show that the inverse of a linear operator is unique.

Problem 27. Let A, B be linear operators on $\mathbb{R}^n$.

a. Show that AB is a linear operator.

b. Show that the inverse of AB is $B^{-1}A^{-1}$.

Problem 28. Let A be a $m \times n$ matrix.

a. What is the domain of the linear transformation induced by A. What is the image?

b. Show that the columns of A span the image of A in $\mathbb{R}^m$. In other words, the image of the linear transformation defined by A is the *column-space* of A.

c. Based on the previous item, interpret geometrically the system of linear equations given by $Ax = b$.

d. Characterize the number of solutions of the system $Ax = b$ in terms of m, n, b, and $rank(A)$.

e. Show that the rows of A are the orthogonal complement of the kernel of A in $\mathbb{R}^n$. In other words, the kernel of the linear transformation defined by A is the orthogonal complement of the *row-space* of A.

f. Based on the previous item, interpret geometrically the system of linear inequalities given by $Ax >= 0$.

Problem 29. Define $f : \mathbb{R}^n \to \mathbb{R}$ by

$$f(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j,$$

where each a_{ij} is a real number.

a. Write down $f(x)$ in matrix notation (no vectors, no inner products, no sums, just products of matrices).

b. Verify that as long as we only care about $f(x)$, the matrix you used in the previous item can be assumed to be a symmetric matrix without any loss of generality.

c. Write down $f(x)$ in vector notation (no sums, no products of matrices — Ax is allowed, but see at as a linear transformation acting on a vector).

d. Give an example where $f(x) > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$. Plot the graph of f. You can use a computer.

e. Give an example where $f(x) < 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$. Plot the graph of f. You can use a computer.

f. Give an example where there are $x, y \in \mathbb{R}^n$ such that $f(x) > 0$ and $f(y) < 0$. Plot the graph of f. You can use a computer.

Hint. *Keep the examples simple.*

Note. *If the expression* quadratic form *did not come to your mind when you saw this exercise, you might want to grab a linear algebra book and study that topic a little bit.*

It is going to be very useful when we study optimization, as second order conditions are essentially determining if quadratic forms (like f here) are positive or negative.

Problem 30. Consider a set of monetary prizes $X = \{\$1, \$2, \$3\}$ and let P be the set of all probability distributions over X, that is, all triples $p = (p_1, p_2, p_3)$ such that $p_1 + p_2 + p_3 = 1$, where p_i is the probability of receiving i dollars. A reasonable way to decide between two lotteries is the following: given $p, q \in P$, if

$$\begin{aligned} p_3 &\geqslant q_3, \\ p_3 + p_2 &\geqslant q_3 + q_2, \\ p_3 + p_2 + p_1 &\geqslant q_3 + q_2 + q_1, \end{aligned}$$

then p is better than q. If the lines above are true with the roles of p and q reversed, then q is better than p.

Define the relation R over X given by pRq if and only if p is better than q.

a. Is this relation complete? Prove your answer is correct.

b. Is this relation convex? Prove your answer is correct.

c. Find a matrix A such that pRq if and only if $Ap \geqslant Aq$.

Note. *The relation R we defined above is called* stochastic dominance. *In general, we say that a random variable X stochastically dominates a random variable Y if and only if $Prob(X > z) \geqslant Prob(Y > z)$ for every $z \in \mathbb{R}$. It can be shown that this is equivalent to saying that $Ef(X) \geqslant Ef(Y)$ for every non-decreasing function $f : \mathbb{R} \to \mathbb{R}$. That is a very important concept when studying attitudes towards risk: if a lottery X (a probability distribution over prizes) stochastically dominates a lottery Y, then X is better than Y independently of risk-behavior (think of f as a utility function): all that is required is that larger prizes be preferred to smaller prizes.*

Problem 31. Let A, P be $n \times n$ matrices. Show that $\det(PAP^{-1}) = \det(A)$.

Note. *This shows that the determinant does not depend on the particular matrix used to represent the underlying linear transformation. If you don't know what I'm talking about, grab a linear algebra book and study change of basis and similarity between matrices.*

Problem 32. Show that if the column of a square matrix is zero, then the determinant of the matrix is also zero.

Hint. *If you cannot prove this, let's cheat a little bit and use the following theorem, which we quote without proof: if $f : \mathbb{R}^n \to \mathbb{R}^n$ is a linear operator and $A \subset \mathbb{R}^n$ is a set for which the volume is well defined, then the volume of $f(A)$ is given by $|\det f|\, volume(A)$. Now, on* **problem 28** *we showed that the image of a linear transformation is its column space. Using these two facts you should be able to solve this problem. Yes, there are shorter algebraic proofs, but this one has geometric intuition!*

Problem 33. Let X be some euclidean space, and define the sum of convex sets A, B in X as $A + B = \{a + b \in X : a \in A, b \in B\}$.

a. Show that the sum of two convex sets is a convex set.

b. Prove by induction that the sum of any finite number of convex sets is a convex set.

Problem 34. Show that the interior of a convex set is a convex set.

Problem 35. Let $f : D \to \mathbb{R}$ be a convex function, $D \subseteq \mathbb{R}^n$.

a. Show that f is continuous in the interior of D.

b. From now on, suppose $D = \mathbb{R}^n$. Show that if f is also concave, then it is in fact affine.

c. Show that if $g : D \to \mathbb{R}$ is also convex, then $f + g$ is convex.

d. Show that if $g : \mathbb{R} \to \mathbb{R}$ is convex and increasing, then $g \circ f$ is convex.

Problem 36. Give an example of a convex function that has a point of discontinuity. A graphic is enough.

Problem 37. Let $A \subseteq \mathbb{R}^n$ be a finite set and (f_α) a family of affine functions $f : \mathbb{R}^n \to \mathbb{R}$ indexed by $\alpha \in A$. Show that the function $g : \mathbb{R}^n \to \mathbb{R}$ defined by $g(x) = \max\{f_\alpha(x) : \alpha \in A\}$ is a convex function. Can you generalize this in any way?

Problem 38. Show that the determinant of a square matrix is the product of its eigenvalues.

Problem 39. Use the fact that the determinant, as a functional defined on the space $\mathbb{R}^n \times \mathbb{R}^n$ of $n \times n$ matrices is a polynomial to show that the set of invertible matrices is open and dense.

Note. *The above result says that we can approximate any singular matrix A by a nonsingular matrix that is arbitrarily close to A. It also says that we cannot approximate a nonsingular matrix by a singular matrix, so this is* not *like the case of the rational and the irrational numbers, where an element of a set can be approximated by the other, and vice-versa.*

Problem 40. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

a. Show that f maps bounded sets to bounded sets.

b. Show that if f is continuous at the origin, then f is continuous.

c. Show that if $rank(f) = n$, then f maps subspaces of dimension $k < n$ to subspaces of dimension k.

d. Show that, in case $m = n$, f is an open mapping if and only if f is full-rank.

Note. *We call an* open mapping *a transformation that maps open sets to open sets.*

Problem 41. Consider the set X of all differentiable functions $f : (0, 1) \to \mathbb{R}$. If f and g are in X, define $f + g$ as the function $x \mapsto f(x) + g(x)$. If $\alpha \in \mathbb{R}$, define αf as the function $x \mapsto \alpha f(x)$.

a. Show that X is a vector space over $\mathbb{R}$.

b. Show that X is not finite dimensional.

c. Define the operator $D : X \to X$ by setting $Df(x)$ as the derivative of f at $x \in (0, 1)$. Using what you know about the rules of differentiation, can you say the operator D is linear?

d. Define the fuctional $I : X \to \mathbb{R}$ by setting If as the integral of f over $(0, 1)$. Using what you know about the rules of integration, can you say the functional I is linear?

Note. *A linear transformation of a vector space on itself is usually called an* operator*. A function from a space of functions (like X in this example) to the real numbers is usually called a* functional*.*

Lecture 4: Separation and Tangency

Problem 42. Show that a closed set $A \subset \mathbb{R}^n$ is convex if and only if it is the intersection of a collection of half-spaces.

Hint. *Use the separating hyperplane theorem.*

Problem 43. We are watching a horse race between horses A and B, and there are people selling lotteries $(x_A, x_B) \in \mathbb{R}^2_+$ that pay x_A to the holder of the lottery ticket in case horse A wins and x_B in case horse B wins. A man with a moustache approaches you and says: "these lotteries are pretty good:$(4, 1)$ and $(2, 3)$. They are definitely better than the lotteries $(3, 2), (1.5, 5)$". Is it possible that the man with the moustache is basing his judgement on the expected value of the lotteries? What if he had not mentioned the lottery $(3, 2)$?

Hint. *Plot the lotteries on $\mathbb{R}^2$ and relate the comparisons of expected value with separating hyperplanes.*

Problem 44. Let $x \in \mathbb{R}^n$. Show that if $\langle x, y\rangle \geqslant 0$ for all $y \in \mathbb{R}^n$, then $x = 0$.

Problem 45. The following result is a famous result known as *Farkas' lemma*, or *Farkas' alternative*: for any $m \times n$ matrix A and any $b \in \mathbb{R}^n$, one, and only one of the following statements is true:

i. $\exists x \in \mathbb{R}^n$ such that $Ax \leqslant 0$ and $\langle b, x\rangle > 0$;

ii. $\exists y \in \mathbb{R}^m_+$ such that $A^T y = b$.

Prove it.

Hint. *Think of the rows of A as collection of vectors in $\mathbb{R}^n$. What does the first item say, geometrically? What does the second item say, geometrically? Can you see how that relates to the separating hyperplane theorem? The two conditions or not relate directly to the possibility or impossibility to separate the vector b from a certain closed convex set $C(A)$ that contains the rows of A. After you figure that out, you should be able to construct a convex and compact set $C(b)$ containing b such that conditions (i) and (ii) above relate to the possibility of impossibility to separate these two sets.*

Note. *Farkas' lemma is often used in microeconomic theory and optimization. It is one of the many known* theorems of the alternative. *The name comes from the fact that the statement of these theorems always follow the format that one, and only one of two statements can be true. Kim Border has notes on many theorems of the alternative.*

Problem 46. Below there is a list of functions $f : \mathbb{R}^2 \to \mathbb{R}$. For each one of them, compute the partial derivatives, write down the gradient, plot the level curve that passes through $(1, 1)$ and the gradient of f at $(1, 1)$, translated to $(1, 1)$. You should get at least the shape of the level curves right without the aid of a computer!

a. $f(x, y) = x^\alpha y^\beta$, where α and β are strictly positive real numbers. For plotting, pick $\alpha = 2$ and $\beta = 1$.

b. $f(x, y) = x + y$.

c. $f(x, y) = \min\{x, y\}$.

d. $f(x, y) = \log x + \log y$.

e. $f(x, y) = x + \sqrt{y}$.

f. $f(x, y) = x$.

Problem 47. Let $\phi : \mathbb{R} \to \mathbb{R}$ be defined by $\phi(t) = f(x[u(t), v(t)], y(t))$, where $f : \mathbb{R}^2 \to \mathbb{R}$, $x : \mathbb{R}^2 \to \mathbb{R}$ and u, v and y map $\mathbb{R}$ to $\mathbb{R}$. What is the expression for $\phi'(t)$?

Problem 48. Find all the partial derivatives of the following functions $f : \mathbb{R}^2 \to \mathbb{R}$.

a. $f(x, y) = \frac{\sqrt{x}+\sqrt{y}-xy}{x^2+y^2}$.

b. $f(x, y) = \exp(xy)\sin(\frac{x}{y})$.

Problem 49. Show that if a function $f : \mathbb{R}^n \to \mathbb{R}^m$ has a derivative at some point $x \in \mathbb{R}^n$, then f is continuous at x.

Problem 50. For each $f : \mathbb{R}^n \to \mathbb{R}$ below, find the derivative Df.

a. $f(x) = \langle p, x\rangle$, where $p \in \mathbb{R}^n$.

b. $f(x) = \langle x, Qx\rangle + \langle p, x\rangle$, where $p \in \mathbb{R}^n$ and Q is a $n \times n$ matrix.

Now write down all the functions above, and their derivatives, in matrix notation (no inner products of vectors, just products of matrices).

Problem 51. Let $f : \mathbb{R} \to \mathbb{R}$ be a concave function. Show that the set where f is not differentiable has measure zero.

Hint. *Prove convex functions always have directional derivatives and use* **problem 20**. *Remember that countable sets have measure zero (see* **problem 18**.*)*

Note. *This result is also true when the domain of f is $\mathbb{R}^n$. However, the proof is more involved.*

Lecture 5: Optimization I

Problem 52. Find all the local minimizers and maximizers of the function

$$f : \mathbb{R}^2 \to \mathbb{R}, \qquad f(x) = x_1^3 + x_2^3 - 3x_1x_2$$

Problem 53. Let Q be a positive definite $n \times n$ matrix and $p, c \in \mathbb{R}^n$. Show that

$$f : \mathbb{R}^n \to \mathbb{R}, \qquad f(x) = \langle Qx, x\rangle + \langle p, x\rangle + c$$

has a unique local maximizer, which is also the global maximizer.

Problem 54. Let $D = \{x \in \mathbb{R}^2 : x_1 + x_2 > 0\}$ and $f : D \to \mathbb{R}$ defined by $f(x) = x_1^2 + x_2 + \frac{1}{x_1+x_2}$. Show that this function has a minimizer.

Hint. *The domain D is open and unbounded, so you cannot use Weirstrass theorem directly. Note, however, that for purposes of minimization, you do not have to stick to D: you can look at the restriction of f to a compact lower contour set. Try to find that lower contour set.*

Problem 55. Show that every upper semi-continuous function defined on a compact set has a maximizer.

Problem 56. Let $f : D \to \mathbb{R}$ with $D \in \mathbb{R}^n$ be a differentiable function. We define a direction of *increase* of f at x as a vector $d \in \mathbb{R}^n$ for which there exists $\varepsilon > 0$ such that $f(x + td) > f(x)$ for all $t \in (0, \varepsilon]$. We define $v \in \mathbb{R}^n$ as a *feasible* (with respect to D) direction at $x \in D$ when there exists $\varepsilon > 0$ such that $x + tv \in D$ for all $t \in [0, \varepsilon]$ (take a moment to picture these concepts; actually drawing an example might help). We also define a *cone* K as a subset of $\mathbb{R}^n$ characterized by the following rule: if $x \in K$, then $tx \in K$ for all $t \in [0, +\infty)$.

a. Show that the union of $\{0\}$ and the set of all directions of increase at x is a cone. Show that the set of all feasible directions at x is a cone.

b. Show that for every direction d of increase at x, we have $\langle \nabla f(x), d\rangle \geqslant 0$.

c. Show that if $d \in \mathbb{R}^n$ satisfies $\langle \nabla f(x), d\rangle > 0$, then d is a direction of increase at x.

d. Show that if x is a local maximizer of f, then $\langle \nabla f(x), v \rangle \leqslant 0$ for all feasible directions. More than that, show that at a local maximizer, no feasible direction can be a direction of increase.

e. Show that if x is a local maximizer in the interior of D, then the set of all feasible directions at x is all of $\mathbb{R}^n$.

f. Conclude that if x is a local maximizer in the interior of D, then $\nabla f(x) = 0$.

Note. *The generalization of this analysis would be the following: if x is a local maximizer of f, then $\langle \nabla f(x), d \rangle \geqslant 0$ for all d in the* contingent/Bouligand *cone at y; if $\langle \nabla f(y), d \rangle > 0$ for all d in the contingent/Bouligand cone, then y is a local maximizer. The contingent/Bouligand cone is akin to the tangent cone, but not quite the same.*

Lecture 6: Projections

Problem 57. Let A be a $n \times k$ matrix, and B a $k \times n$ matrix. What are the number of rows and columns of AB? Show that $(AB)^T = B^T A^T$.

Problem 58. Let X be a $n \times k$ matrix. Show that $X^T X$ is symmetric.

Problem 59. Let X be $n \times k$ matrix. Show that the matrix $X^T X$ is invertible if and only if the columns of X are linearly independent. If this is too hard, try the case where $k = n$; this case has a very simple proof.

Problem 60. Let X be a $n \times k$ matrix, y be a $n \times 1$ matrix and $\hat{\beta}$ be a $k \times 1$ matrix. Interpret the equation

$$X^T (y - X\hat{\beta}) = 0$$

geometrically.

Hint. *Orthogonality.*

Problem 61. Let A be a full-rank $n \times n$ matrix. Show that $(A^T)^{-1} = (A^{-1})^T$.

Problem 62. Let $x_1, x_2, \ldots, x_k$ be $n \times 1$ matrices, and consider the matrix X that has x_i as it's i-th column, $i \in \{1, \ldots, k\}$.

a. Show that the matrix $P_X = X(X^T X)^{-1} X^T$ is an idempotent matrix, that is, $P^2 = P$. This means that it is the matrix of a projection operator. What subspace does it project on? You should not try an argument based on orthogonality.

b. Let M_X be the matrix defined by $P_X + M_X = I$. Show that M_X is also an idempotent matrix, and therefore, a projection. What subspace does it project on? (*Hint*: think of the kernel and the image of P_X). You should not try an argument based on orthogonality.

c. Show that $\mathbb{R}^n$ can be written as the direct sum of the kernel and the image of P_X.

d. Finally, let us use the concept of orthogonality. Show that the matrix P_X is symmetric. Show that this implies that P_X is the matrix of a orthogonal projection operator.

e. Show that the kernel of P_X is the orthogonal complement of the image of P_X.

f. Show that the quadratic form $z \mapsto \langle P_X z, z\rangle$ is positive semidefinite.

Lecture 7: Least-Squares

Problem 63. Consider two functions $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$. Prove the following statements.

a. If g is nondecreasing, then every maximizer of f is a maximizer of $g \circ f$.

b. If g is *strictly* increasing, then every maximizer of $g \circ f$ is a maximizer of f.

Problem 64. Fix $y \in \mathbb{R}^n$ and let $x_1, x_2, \dots, x_k \in \mathbb{R}^n$ be a set of linearly independent vectors. Find the orthogonal projection of y on the linear space spanned by $x_1, \dots, x_k$, that is, find the solution to the problem:

$$\text{minimize}_{\beta\in\mathbb{R}^k} \left\| y - X\beta \right\|,$$

where X is the matrix that has x_i as the i-th column vector, $i \in \{1, \dots, k\}$. Why do we know that there exists a solution to this problem? Why do we know the solution is unique?

Problem 65.

a. Consider L^2 as the space of random variables with finite variance. Show that the covariance $cov : L^2 \times L^2 \to \mathbb{R}$, is a positive, symmetric bilinear functional on L^2. What is the norm with respect to this inner product? What about the square of the norm?

b. The *Cauchy-Schwarz* inequality holds in any vector space with inner product. In particular, it holds in L^2. Write down the Cauchy-Schwarz inequality using the covariance as the inner product of L^2 and conclude that the correlation coefficient is always between -1 and 1, and is only equal to 1 if one of the random variables is a multiple of the other.

Lecture 8: Optimization II

Problem 66. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of class C^1, and let $M = f^{-1}(c)$ where $\nabla f(x) \neq 0$ for all $x \in M$. We define T_pM as the plane *tangent* to M at p as the collection of all "velocity vectors" $v = \lambda'(0)$ of all differentiable functions $\lambda : (-\varepsilon, \varepsilon) \to M$ such that $\lambda(0) = p$ (take a moment to picture that). In the following, let $x \in M$.

a. Show that $\nabla f(x)$ is a direction of increase at x.

b. Show that $\nabla f(x)$ is the direction of steepest/largest increase at x: that is, $\left\langle \nabla f(x), \frac{\nabla f(x)}{\|\nabla f(x)\|} \right\rangle \geqslant \left\langle \nabla f(x), \frac{d}{\|d\|} \right\rangle$ for all $d \in \mathbb{R}^n$.

c. Show that the gradient $\nabla f(x)$ is orthogonal to the plane tangent to M at x.

d. Define the plane tangent to M at x in terms of the gradient of f.

Hint. *Item (b): Cauchy-Schwarz inequality. Item (c): chain rule. Item (d): implicit function theorem.*

Problem 67. Let $0 < \sigma_1 < \sigma_2$ be fixed real numbers. Find all the minimizers and maximizers of the function

$$f : \mathbb{R}^2 \to \mathbb{R}, \qquad f(x) = \frac{\sigma_1 x_1^2 + \sigma_2 x_1^2}{2}$$

subject to the constraint $h(x) = x_1^3 + x_2^3 = 1$. Write down the local optimizers, classying them as maxima or minima, the locally optimum values, and the lagrange multipliers when $\sigma_1 = 1$, $\sigma_2 = 2$.

Problem 68. Let p_1, p_2 and w be striclty positive real numbers, and $\alpha \in (0, 1)$. Find all the minimizers and maximizers of the following *Cobb-Douglas* utility function

$$f : \mathbb{R}^2 \to \mathbb{R}, \qquad u(x) = x_1^{\alpha} x_2^{1-\alpha}$$

subject to the constraint $h_1(x) = p_1 x_1 + p_2 x_2 = w$. Write down the maximizer (or maximizers, in case there is more than one), the maximum value of the function and the Lagrange multipliers for the particular case where $p_1 = 0.5, p_2 = 1, w = 4$, and $\alpha = 1/3$. What if $p_1 = p_2 = 1$?

Problem 69. State and prove the envelope theorem for unconstrained maximization problems. Draw a picture to illustrate the theorem, making explicit the reason of the name “envelope”.

Hint. *Write the value function as the composition of the objective function with the decision rule. Use the chain rule to differentiate that, and use the first order conditions to obtain the desired result.*

Lecture 9: Optimization III

Problem 70. Consider the problem of maximizing a utility function $u : \mathbb{R}^L_+ \to \mathbb{R}$ subject to a budget constraint ,

$$\begin{aligned} \text{maximize}_{x\in\mathbb{R}^L} \quad & u(x) \\ \text{subject to} \quad & \sum_{i=1}^{L} p_i x_i \leqslant w \\ & x_i \geqslant 0 \qquad \forall i \in \{1,\dots,L\}, \end{aligned}$$

where $p \in \mathbb{R}^L_{++}$ is a vector of prices, and $w \in \mathbb{R}_{++}$ is the agent's wealth. Solve this problem for the utility functions below. If needed, make additional assumptions to solve the problem, but always be explicit about them. For all items below, write down the solution x, the optimal value of the utility function, the Lagrange multipliers of each constraint, and list which constraints bind and which don't. Do all that as a function of the parameters p, w. Include the numerical results for a particular case, setting the parameters p, w to specific numbers. Feel free to use a computer to aid you, but remember you must justify why the procedures you did are legitimate. Keep in mind these examples are simple enough that you shouldn't *need* a computer to solve them. *Please divide your solution in steps, write clearly, remember to justify your statements. Solve it in a separate sheet of paper and only after that write down the final solution. Be organized.*

a. $u(x) = \prod x_i^{\alpha_i}$.

b. $u(x) = x_1 + \sqrt{x_2}$.

c. $u(x) = x_1 + x_2$.

d. $u(x) = 3x_1 + x_2$.

e. $u(x) = \min\{x_1 + 4x_2\}$.

f. $u(x) = 1 - \exp(x_1 + 3x_2)$.

Problem 71. Suppose a consumer is deciding how much to consume of two goods. However, to consume something she has not only to pay for it, but also to spend some time to actually consume it. Having limited money and time, our consumer faces both a budget and a time constraint. The monetary price of one unit of the first good is 10, and the monetary price of one unit of the second good is 1. Her budget is 21. The time price of good 1 is 1, and the time price of good 2 is 8; her time budget is 10. Her utility of consuming amounts x of the first good and y of the second good is xy.

a. Solve the optimization problem of the consumer described above. Report not only the maximizers, but also the slacks in the constraints, and the multipliers. Are the multipliers unique? Why or why not?

b. Suppose that now our consumer needs also to carry the goods home, and thus has limited carrying capacity. Set her carrying capacity constraint to $x + y \leqslant 3$. Report not only the maximizers, but also slacks in the constraints and multipliers, if they exist. In case multipliers exist, are the multipliers unique? Why, or why not?

Problem 72. Solve the following optimization problem.

$$\begin{aligned} \text{maximize}_{(x_1,x_2)\in\mathbb{R}^2} \quad & x_2 - x_1^2 \\ \text{subject to} \quad & -(10 - x_1^2 - x_2)^3 \leqslant 0 \\ & -x_1 \leqslant -2 \\ & x_1, x_2 \geqslant 0 \end{aligned}$$

Problem 73. Solve the following optimization problems.

a.

$$\begin{aligned} \text{maximize}_{(x_1,x_2)\in\mathbb{R}^2} \quad & x_1 \\ \text{subject to} \quad & x_2 - (1 - x_1)^3 \leqslant 0 \\ & x_1, x_2 \geqslant 0 \end{aligned}$$

b. Same problem as above, but with the added constraint $2x_1 + x_2 \leqslant 2$.

Problem 74. Solve the following problem.

$$\begin{aligned} \text{minimize}_{(x,y)\in\mathbb{R}^2} \quad & x^2 + 3y^2 \\ \text{subject to} \quad & x + 2y \geqslant 1 \\ & 2x + y \geqslant 1. \end{aligned}$$

In your solution, include the minimizer (x, y), the slacks in each constraint and the lagrange multipliers of each variable.

Problem 75. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. Show that there exists x in the unit ball in $\mathbb{R}^n$ and $t \in \mathbb{R}_+$ such that $\nabla f(x) = tx$.

Hint. *The equation we have represents the first order conditions of a certain optimization problem.*

Problem 76. State and prove the envelope theorem for constrained maximization problems. Under the light of this theorem, interpret the Lagrange multipliers of the inequality constraints in a nonlinear program in standard form.

Hint. *Write the value function as the composition of the Lagrangean with the decision rule. Use the chain rule to differentiate that, and use the first order KKT conditions to obtain the desired result.*

Lecture 10: Correspondences, Fixed Points and Systems of Equations

Problem 77. Show that a compact-valued correspondence is upper hemicontinuous if and only if it has closed graph.

Problem 78. State and prove the Maximum Theorem. If you can't prove the general case, assume that the objective function is strictly concave on the decision variables for all parameters (that will give you a unique solution for each problem, and then the decision rule will be a single-valued correspondence, that is, a function).

Problem 79. A *game in normal form* is given by a finite set N of players, a set of actions A_i for each player $i \in N$, and a set of utility functions $u_i : A \to \mathbb{R}$, where $A = \prod_{i=1}^{n} A_i$. We assume the action sets are convex compact subsets of some euclidean space and that the utility functions are continuous and quasiconcave. An action profile $a \in A$ is a *Nash equilibrium* when for every player $i \in N$

$$u_i(a_i, a_{-i}) \geqslant u_i(a_i', a_{-i}) \qquad \forall \quad a_i' \in A_i.$$

Prove that under these assumptions, every game in normal form has a Nash equilibrium.

Hint. *Define a "best-reply correspondence" from A into A, and see how Nash equilibria relate to fixed points of this correspondence. Use the maximum theorem + Kakutani's fixed point theorem to guarantee existence of Nash equilibria.*

Lecture 11: Probability

Problem 80. Consider a congress with n congressmen, where bills are passed by the following *quota rule*: a bill is passed when *more* than nq congressmen vote "yea". What is the probability of a "pivotal event", that is, what is the probability that there exists a swing voter? (A swing voter, or a pivotal voter exists when given the vote count, the outcome could be reversed by reversing the vote a of a single voter/congressman).

Problem 81. Consider a class with n students, and suppose that that for any student in the class and any two days of the 365 days in a year, the probability of that student having his birthday in one day or the other is the same. What is the probability that there are at least two students in the class whose birthdays coincide? How many students do you need for that probability to be greater than .5? What about greater than .99?

Problem 82. An urn contains three balls, numbered 1, 2 and 3. Two balls are drawn from the urn, one after the other, randomly, without replacement. Let X be the number of the frist ball, and Y the number of the second ball.

a. What is the joint distribution of X and Y?

b. Compute $P(X < Y)$.

c. Determine the marginal distributions of X and Y.

d. Are X and Y independent?

Problem 83. Suppose that the lifetime of a certain light bulb is given by an exponential distribution with parameter λ.

a. If T is the lifetime of the light bulb, show that

$$P(T > t + s | T > t) = P(T > s) \qquad \forall s, t > 0$$

This is the *memoryless property* of the exponential distribution.

b. Suppose $\lambda = 3$ and that lifetime is expressed in days. A light bulb is turned on in a room at instant $t = 0$. A day later, you get in the room and stay there for $8 hours$, leaving by the end of this period.

   i. What is the probability that you enter the room with a burned light bulb?

   ii. What is the probability that you enter the room with a functional light bulb, but exit with a burned one?

Problem 84. Consider a lottery that is run in the following way. There are M tickets numbered $1, 2, \ldots, M$, of which n, numbered $1, \ldots, n$ win prizes. Assume $M \geqslant 2n$.

a. If you buy n tickets, what is the probability of winning at least one prize?

b. Suppose $M = n^2$. What is the limit, as $n \to \infty$, of the probability of winning at least one prize?

Problem 85. The Dodgers are playing against the Yankees in a world series. The Dodgers win each game with probability 0.6. What is the probability that the Dodgers will win the series? (The series is won by the first team to win four games)

Problem 86. Bob wants to send a letter to Alice this month, but he is a lazy guy. Alice thinks that the probability that Bob will write the letter is 0.8. The probability that the mail service does not lose the letter is 0.9. The probability that the postman delivers it is 0.9. Given that Alice did not receive a letter, what is the probability that Bob did not write it?

Problem 87. Two players have 200 dollars each. They flip a coin with probability $p \in (0, 1)$ of "heads" coming up. If "heads" comes up, player 1 gets 100 dollars from player 2; if the outcome is "tails" then player 1 pays 100 dollars to player 2. They keep doing this until one of the players loses all his money. The coin flips are independent from each other. Let N be the number of coin flips necessary to end the game. Compute the expected value of N.

Problem 88. Show that if X and Y are two variables in L^2, then $Y - E(Y|X)$ is always orthogonal to X.

Problem 89. Consider the following experiment: Alice flips an "honest" coin n times, obtaining k heads. After that, Bob flips the same coin k times. All coin flips are independent from each other. Compute the expected value of X, where X is the number of heads obtained by Bob.

Problem 90. Consider a random variable X that assumes strictly positive values with probability one. Show that:

a. $E(\frac{1}{X}) \geqslant \frac{1}{EX}$.

b. $E(logX) \leqslant log(EX)$.

Problem 91. In a multiple choice exam, the probability that a student knows the answer is $p \in (0, 1)$. For each problem in the exam, there are m choices. If the student knows the answer of a certain problem, he marks the right alternative with probability one; otherwise, he chooses one of the alternatives randomly, with uniform probability. For a given problem, what is the probability that he knew the answer, given that he marked the correct alternative? Compute the limit of this probability when

a. $m \to \infty$ with fixed p;

b. $p \to 0$ with fixed m.

Problem 92. Out of three prisoners scheduled to be executed, A, B, and C, one of them will be pardoned. A asks the warden to tell him the name of one of the others who will be executed. As the question is not directly about A's fate, the warden obliges — either naming the other prisoner to be executed, in case A was too, or secretly flipping a coin to decide which of the remaining names to give A if A is the one being pardoned. Assuming the warden's truthfulness, there are now only two possibilities for who will be pardoned: A, and whichever of B or C the warden did not name. Did A gain any information as to his own fate, that is, does he change his estimate of the chances he will be pardoned? If warden says "B will be executed" and A could switch fates with C, should he?

Hint. *Suppose that instead of 3 prisoners, there are 1000, and the warden gives A the name of 999 prisoners that are going to die. If you still can't solve this, look up the "Three Prisoners' Problem", or "Monty Hall's Problem".*