Oral Qualifying Exam

Nilava Metya

09/11/2024

A : OK so let's start. Why don't we start with probability?
C : Can you please state the Borel-Cantelli lemmas?
Me: Yes. *Starts stating verbally.*
C : Could you write on the board?
Me: *Goes to board.* The first Borel-Cantelli lemma states that if $\{A_i\}_{i=1}^{\infty}$ is a collection of events such that $\sum_i \mathbb{P}(A_i) < \infty$ then $\mathbb{P}(\limsup A_n) = 0$. The second Borel-Cantelli lemma states that if $\{A_i\}_{i=1}^{\infty}$ is a collection of independent events such that $\sum_i \mathbb{P}(A_i) = \infty$ then $\mathbb{P}(\limsup A_n) = 1$.
C : Very good. Now answer this question. You have a fair coin and you're tossing it infinitely many times. What is the probability that only finitely many heads turn up?
Me: Are you suggesting that we could use the Borel-Cantelli lemmas for this?
C : They might be useful.
Me: So we want to define these events A_i judiciously. Let's see. We want to find $\mathbb{P}(\text{finitely many heads})$. What if we compare it with $\mathbb{P}(\text{infinitely many tails})$ and define $A_i = \{i^{\text{th}} \text{ coin is tail}\}$ and $\sum \mathbb{P}(A_i) = \infty$ so that $\mathbb{P}(\text{infinitely many tails}) = 1$. Also these A_i are independent. *Makes some trivial mistake about getting the inequality correct between $\mathbb{P}(\textit{finitely many heads})$ and $\mathbb{P}(\textit{infinitely many tails})$. But got it finally.* But this gives a tautology that the probability is at most 1.
C : Yes so you have to define A_i as something else so that you can extract more information.
Me: Maybe let's write it differently $\mathbb{P}(\text{finitely many heads}) = 1 - \mathbb{P}(\text{infinitely many heads})$. Ah so we take $A_i = \{i^{\text{th}} \text{ coin is head}\}$ which still makes the sum of probabilities diverge. By Borel-Cantelli, the latter proabbility is 1, so the probability we want is 0.
C : Very good. Does it matter that it's a fair coin?
Me: No, it can be any $p-$coin with $p > 0$.
C : Do we need independence?
Me: I don't remember a counterexample on the top of my head, but I know that the lemma fails without the independence assumption.
C : Those were all my questions.
B : Maybe I can ask now. Can you give an example of a convex set and a non-convex set?
Me: *Draws the unit circle in $\mathbb{R}^2$ and writes the set description.* This is a convex set. A non-convex set could be $\{-1, 1\}$ in $\mathbb{R}$.
B : Good. Consider a closed convex set A in $\mathbb{R}^n$. Consider the function $f(x) = d(x, A)^2$. Is it convex? Can you tell me some regularity properties?
Me: Sure, let's try maybe an example. *Draws the graph in $\mathbb{R}^2$ for $A = [0, 1]$.* In this case, it seems to be convex and differentiable.
B : What happens if A is not convex? Can you show the graph for your non-convex example?
Me: Sure. *Draws the graph for $A = \{-1, 1\}$.* This is clearly not convex, and not differentiable. Look at 0.
B : Yes so there can be kink, right? Now what about closed convex A?
Me: *Thinks about it, completely lost because don't know 'yes' or 'no'.*
B : Ok so I tell you that it's convex and differentiable. Can you prove it.
Me: Starts proving convexity using the first definition of convexity.
B : Maybe we're running out of time now. I'll tell you my next question. Is it strongly convex?
Me: *Starts thinking, writes definition.* Well, inside A the function is constant 0, so it can't be strongly convex.
B : What about its behaviour outside A?
Me: *Starts thinking, lost again.*
B : What happens if A is the square?
Me: *Draws the square.* If we go outward, it's increasing and behaves quadratically.
B : But it's not really about only this direction right? We need to consider all directions. What if you walk parallel to a side?
Me: *Makes a very bad mistake again, to find the projection of each x on A.*

A : Where does the distance minimize? It's like a perpendicular, right?

Me: Oh, my bad. Yes, along this direction, the function stays constant because the lengths of the perpendiculars are the same. So not strongly convex.

B : OK. Maybe we can move onto someone else now.

D : I could ask next. Can you tell me about solving polynomial equations?

Me: The setting is in $R = \mathbb{C}[x_1, \cdots, x_n]$. We are given polynomials $f_1, \cdots, f_k \in R$ and we want to find common solutions. From high school the way to solve, say, a system of simultaneous linear equations in variables x, y is we eliminate the x and solve for y, then plug the value of y into an equation to recover x. This is analogous to polynomial division in one variable, where we 'cancelled' the x. So we need a formal notion of polynomial division in multiple variables. This gives us the notion of monomial ordering which is really a refinement of the division poset on all monomials. Now, we need a reduced form of the collection of these polynomials to do something analogous to the division algorithm in one variable. This is resolved by something called a Gröbner basis. A Gröbner basis is $\cdots$

D : Let's say we know about Gröbner bases.

Me: Alright. So we want to focus on the ideal $I = \langle f_1, \cdots, f_n \rangle \subseteq R$. I'll talk about elimination theory. Let's fix a lex order with $x_1 > \cdots > x_n$. What I told earlier about eliminating x to get a linear equation is y is exactly starting with the ideal in $\mathbb{C}[x, y]$ generated by the two linear polynomials and intersecting it with $\mathbb{C}[y]$. In this sense, we define the ℓ^{th} elimination ideal as $I_\ell = I \cap \mathbb{C}[x_{\ell+1}, \cdots, x_n]$. A helpful fact to deal with elimination ideals is that if G is a Gröbner basis for I with the given lex ordering, then $G_\ell = G \cap \mathbb{C}[x_{\ell+1}, \cdots, x_n]$ is a Gröbner basis for I_ℓ. So what do we do now. Let's use x, y, z instead of $x_1, \cdots, x_n$.

D : Your z being x_n?

Me: Yes, and $n = 3$. So the idea was that we keep on eliminating variables, until we reduce to an ideal with only one variable. Say we solved the equations simultaneously and we got to $I \cap \mathbb{C}[z]$. Then we find the root of the corresponding equation. Or the simultaneous equations in z.

D : Can there be multiple equations in the last elimination ideal?

Me: Oh! There will be exactly one because $\mathbb{C}[z]$ is a PID. In fact, this is related to one of the characterizations of a Gröbner basis. So we solve this equation in z. Then we can go to the equations in $I \cap \mathbb{C}[y, z]$, substitute those values of z and then solve for y. This would give solutions for (y, z). This happens for some specific values of z, which I'll talk about in a minute (the Extension theorem). After getting those values of y we go to I and solve for x after substituting for y, z. Now the question was: which solutions z of $I \cap \mathbb{C}[z]$ extend to solutions (y, z) of $I \cap \mathbb{C}[y, z]$? The extension theorem says that if a partial solution $\boldsymbol{a} = (a_2, \cdots, a_n)$ of $I \cap \mathbb{C}[x_2, \cdots, x_n]$ extends to a solution $(a_1, \boldsymbol{a})$ of $I \subseteq \mathbb{C}[x_1, \cdots, x_n]$ if the leading terms of a Gröbner basis of I with an ordering where any term containing x_1 is more than any term containing the other variables (equivalently, leading terms of I) do not simultaneously vanish at $\boldsymbol{a}$.

D : Can you mention some issues that might come up with this method of solving equations?

Me: The first thing that comes to my mind is that usually the equations have high degree and due to Abel/Galois we cannot solve the equations exactly. This means we need to use numerical methods. So we solve for z, and then plug in to solve for y. Already the approximate values of z make the equations for y approximate, whence the errors in the solutions of y have even more error. Usually there are more variables, so the errors accumulate together to give very inaccurate solutions.

D : How do you overcome this? Any alternatives? Something to do with eigenvalues?

Me: Ah right. So let me tell you the general setup. We are only interested in zero dimensional ideals I, that is, there are finitely many solutions. Now I is zero dimensional iff $A = R/I$ is finite dimensional. In fact, we can find this dimension. A basis is given by the monomials not divisible by leading terms of I.

D : Say that again?

Me: A basis of A is given by all monomials in R which are not divisible by any leading monomial of I (or a Gröbner basis of I). Now this is useful because if I is a zero-dimensional ideal then fixing a polynomial $f \in R$, the eigenvalues of the linear map $m_f : A \to A$ (where $m_f([g]) = [fg]$) exactly coincides with $f(V(I))$.

D : Any assumptions?

Me: Yes we need that I is radical. The way we use this to solve systems of polynomial equations is taking

f to be x_i, the projection onto the i^{th} coordinate. This way we can recover all coordinates of $V(I)$. Well, modulo repeated solutions.

D : What do you mean repeated solutions?

Me: I mean double roots. *Looks at professor*$\cdots$ Oh! There's not double roots because I is radical.

D : How many f do you need.

Me: n of them. We can also take random linear forms.

D : But why? You still need n of them. Something to do with eigenvectors?

Me: Oh I remember there was a section about that, but I don't remember anything from there honestly.

D : Maybe we can move on to someone else now.

A : Can you write down a typical optimization problem? With the inequalities and equalities.

Me: I write $\min_{x\in\mathbb{R}^n} f(x)$ subject to $g_i(x) \le 0 \forall 1 \le i \le m, h_i(x) = 0 \forall 1 \le i \le k$.

A : When do we say that the problem is convex?

Me: When g_i are all convex and h_i are affine. But this is only for computational purposes. In general the constaint set can be convex without all constraints being of this type.

A : Uhhh, what about f?

Me: Oh oh oh! f also has to be convex.

A : Can you write down the KKT conditions?

Me: I don't remember them on the top of my head.

A : That's ok. Can you write down a typical SDP problem?

Me: Yes, it looks like $\min_{X\in S^n} \operatorname{Tr}(CX)$ subject to the linear constraints $\operatorname{Tr}(A_iX) = b_i \forall 1 \le i \le m$ and $X \succeq 0$. Here C, A_i are all symmetric and S^n stands for the space of $n \times n$ symmetric matrices.

A : Can you tell me why SDP is a generalization of LP?

Me: I'll take a generic LP and write an SDP which has the same optimal value and the optimal solutions can be recovered. Are you happy with the following generic LP: $\min_{x\in\mathbb{R}^n} c^\top x$ subject to $Ax = b, x \ge 0$ where $c \in \mathbb{R}^n, A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$?

A : Sure that's ok.

Me: So the SDP we want to form will have the variable matrix X modelling the parameter x in LP, the obvious way is to put x in the diagonals of X and force the off-diagonal entries of X to be 0 via the linear equalities $\operatorname{Tr}(E_{ij}X) = 0$ for each $i < j$. The constraint $x \ge 0$ is equivalent to $X \succeq 0$ with the above understanding. So the equivalent SDP is $\min_{X\in S^n} \operatorname{Tr}(diag(c)X)$ subject to $\operatorname{Tr}(E_{ij}X) = 0$ for each $i < j$, where $diag(c)$ is the $n \times n$ matrix with diagonal entries c.

A : Uhh what about $Ax = b$?

Me: Oh sorry! I will add the additional constraints that $\operatorname{Tr}(diag(a_i)X) = b_i$ where a_i is the i^{th} row of A.

A : Can you write down a generic SOCP problem?

Me: For the second order cone problem, I first will write the Laurent cone which is $\mathcal{L}_{n+1} = \{(x,t) \in \mathbb{R}^{n+1} : ||x||_2 \le t\}$. And now the optimization problem is: $\min_{x\in\mathbb{R}^n} f^\top x$ subject to $(A_ix+b_i, c_i^\top x+d_i) \in \mathcal{L}_{n+1} \forall 1 \le i \le m$.

A : Can you write it as an SDP?

Me: Ahh, let me try to look at the general constraint $||Ax + b||_2 \le c^\top x + d$. Now we square it $(Ax + b)^\top(Ax + b) \le (c^\top x + d)^2$. What if we group the quadratic and linear terms together. So we maybe try to do something with $A^\top A - cc^\top$?

A : Something about Schur complements?

Me: Uhhh... *Thinks how to manipulate.* Maybe $\begin{bmatrix} I & A \\ A^\top & I \end{bmatrix}$?

A : Not identity but ssomething else. It's okay, let it be. It takes time. Let's see if someone else has any questions?

D : Yes I can ask. Can you tell me what's a multipolynomial discriminant?

Me: I do not know what a multipolynomial discriminant is, but I can tell you about multipolynomial resultants.

D : Oh yes right, that's what it was called. Sure tell me about it.

Me: Let me define it using a theorem. Given $n+1$ homogeneous polynomials $F_0, \cdots, F_n \in \mathbb{C}[x_0, \cdots, x_n]$ of total degrees $d_0, \cdots, d_n$ respectively, say $F_i = \sum_{|\alpha|=d} u_{i,\alpha}x^\alpha$, there is a unique polynomial $\operatorname{Res}_{d_0,\cdots,d_n} \in$

$\mathbb{Z}$[all $u_{i,\alpha}$] satisfying the following:

- F_i's have a common nontrivial solution iff Res is zero.
- $\text{Res}(x_0^{d_0}, \cdots, x_n^{d_n}) = 1$.
- Res is irreducible.

D : What can you tell me about the geometry of resultants?

Me: Well, the point of resultants is to eliminate $n+1$ variables from $n+1$ homogeneous equations. *Thinks what to say*

D : So you have this polynomial residing in this huge space of coefficients. Where does the corresponding variety reside? What about the variety of the original polynomials?

Me: If we consider the polynomials as a general polynomial, that is, the $u_{i,\alpha}$'s and x_i's are all treated as variables. The corresponding variety lies in $\mathbb{C}^N \times \mathbb{P}^n$ where N is the total number of coefficients. If we want to eliminate the $n+1$ variables $x_0, \cdots, x_n$, on the geometry side we are only taking the coordinates corresponding to the $\mathbb{C}^N$ part.

D : What is that called?

Me: Projection. And the projection is given by the ... the closure of the projection is cut out by the equation given by Res.

D : Do you know about complete varieties?

Me: No.

D : It's a variety X for which the projective morphism $X \times Y \to Y$ is closed for any variety Y. Projective varieties are complete.

Me: Okay, so in this case the image is closed. So Res gives the image of the projection actually.

D : What's the codimension of this projected variety?

Me: It's ideal is princiaplly generated, so codimension one.

D : How many connected components does it have?

Me: Uhhh I'm not sure? *Looks at the board*. Oh Res is irreducible, so one component.

D : Anybody else wants to ask anything? *everybody looking at each other.*

E : I think Nilava has answered his fair share of questions.

C : Now we ask Nilava to wait outside while we discuss.

Me: *Walks out of the door and waits.*

A : You can come in now.

Everyone: Congratulations!