Pattes pysselhörna, del 1: Svar

Patte

February 1, 1996

Vad var frågan nu igen??

Frågan löd:

1. Vilka funktioner definierade på $\Re^n$ är både konvexa och konkava?

Vad svarade vi??

Standardsvaret (som alla bör känna till från sin egen grundkurs) är att alla affina funktioner, dvs. klassen av funktioner som skiljer sig från en linjär funktion med högst en konstant, är både konvexa och konkava. Det är en del av sanningen, men det är långt ifrån hela sanningen, och med undantag för α-β (vars grekiska jag ännu inte lyckats tyda) har ingen gjort något allvarligt försök att passera utanför de affina funktionernas värld.

Var det så åt peppar'n fel?? Vad är det rätta svaret då?? Håll oss inte på sträckbänken här!!

Eftersom det affina fallet är det som man direkt tänker på passar det bäst att utgå från det, försöka karakterisera klassen av affina funktioner och med det som utgångspunkt gå vidare till en större klass.

De affina funktionerna

En affin funktion på $\Re^n$ är en funktion på formen

$$f(x) = a^{\mathrm{T}} x - b, \quad a \in \Re^n, \quad b \in \Re, \qquad x \in \Re^n.$$

För att kunna titta närmare på dess egenskaper rekapitulerar vi definitionen av epigraf och konvex funktion. Nedan betyder $\overline{\Re}$ mängden $\Re \cup \{-\infty\} \cup \{+\infty\}$, dvs. intervallet $[-\infty, +\infty]$; detta brukar kallas mängden av utvidgade reella tal (extended real values på danska).

Definition 1 (epigraf, hypograf) Epigrafen *av en funktion* $f : \Re^n \mapsto \overline{\Re}$ *är mängden*

$$\texttt{epi}\, f = \{\, (x, \alpha) \in \Re^{n+1} \mid f(x) \leq \alpha \,\}.$$

Hypografen *är mängden*

$$\texttt{epi}\, f = \{\, (x, \alpha) \in \Re^{n+1} \mid f(x) \geq \alpha \,\}.$$

Epigrafen är alltså alla punkter som ligger på eller ovanför funktionen f:s graf, och hypografen alla punkter som ligger på eller under densamma.

Definition 2 (konvexitet, konkavitet) *En funktion* $f : \Re^n \mapsto \overline{\Re}$ *är* konvex *om* $\texttt{epi}\ f$ *är en konvex mängd. Funktionen är* konkav *om* $\texttt{hypo}\ f$ *är en konvex mängd.*

Alltså kan man karakterisera en funktion f som varande både konvex och konkav om både dess epigraf och hypograf är konvexa mängder.

I det affina fallet har vi:

$$\begin{aligned} \texttt{epi}\, f &= \{\, (x, \alpha) \in \Re^{n+1} \mid a^{\mathrm{T}} x - b \leq \alpha \,\}, \\ \texttt{hypo}\, f &= \{\, (x, \alpha) \in \Re^{n+1} \mid a^{\mathrm{T}} x - b \geq \alpha \,\}, \end{aligned}$$

dvs. epigrafen och hypografen är halvrum i $\Re^{n+1}$. Vi noterar också att eftersom $\texttt{epi}\ f$ och $\texttt{hypo}\ f$ är konvexa mängder separeras de av ett hyperplan, som i detta fall är

$$\texttt{epi}\, f \cap \texttt{hypo}\, f = \{\, (x, \alpha) \in \Re^{n+1} \mid a^{\mathrm{T}} x - b = \alpha \,\} =: H.$$

Låt oss betrakta mängden H en stund. (OK, nu kan ni sluta stirra.) Kan en affin funktion ha ett snitt $\texttt{epi}\ f \cap \texttt{hypo}\ f$ som *inte* är ett hyperplan i $\Re^{n+1}$? Nej, det är en 1-1 korrespondens mellan hyperplan på formen H och klassen av affina funktioner. (En affin funktion kan representeras med ett hyperplan på formen H, och, omvänt, varje hyperplan på formen H är snittet $\texttt{epi}\ f \cap \texttt{hypo}\ f$ för en affin funktion.)

Vi kan alltså karakterisera klassen av affina funktioner sålunda:

$$f \text{ är affin}$$
$$\Longleftrightarrow$$
$$\texttt{epi}\, f \cap \texttt{hypo}\, f \text{ är ett hyperplan i } \Re^{n+1} \text{ på formen } H.$$

Nu tänker vi en stund till. För en allmän funktion f definerad på $\Re^n$ är

$$\texttt{epi}\, f \cap \texttt{hypo}\, f = \{\, (x, \alpha) \in \Re^{n+1} \mid f(x) = \alpha \,\}.$$

Denna mängd kan uppenbarligen inte vara konvex om funktionen f uppträder icke-linjärt någonstans där den är ändlig. Det tycks alltså vara så att det bara är affina funktioner som klarar biffen. Nej, det är inte sant, och det beror på att det finns hyperplan i $\Re^{n+1}$ som *inte* är på formen H, men som representerar funktioner som är både konvexa och konkava.

Improper konvexa och konkava funktioner

Hyperplanen i $\Re^{n+1}$ representeras av de linjära funktionerna i $\Re^{n+1}$, som i sin tur representeras via

$$(x, \alpha) \mapsto a^{\mathrm{T}} x + b_0 \alpha, \quad a \in \Re^n, \quad b_0 \in \Re.$$

Eftersom en multiplicering av en godtycklig (noll-skild) konstant inte förändrar ett hyperplan kan vi utan att förlora i allmängiltighet betrakta fallen $b_0 = 0$ och $b_0 = -1$. Hyperplanen motsvarande $b_0 = -1$ är på formen

$$\{\, (x, \alpha) \in \Re^{n+1} \mid a^{\mathrm{T}} x - \alpha = b \,\}, \quad a \in \Re^n, \quad b \in \Re,$$

dvs. hyperplanen av typen H ovan. Hyperplanen motsvarande $b_0 = 0$ är på formen

$$\{\, (x, \alpha) \in \Re^{n+1} \mid a^{\mathrm{T}} x = b \,\}, \quad 0 \neq a \in \Re^n, \quad b \in \Re;$$

i Rockafellars konvex analys-bok (där distinktionen mellan olika typer av hyperplan är hämtad) kallas dessa hyperplan för *vertikala*, därför att riktningen $(0^{\mathrm{T}}, -1)^{\mathrm{T}}$ ligger i hyperplanet.

De funktioner vi kan representera med vertikala hyperplan kan inte vara ändliga överallt (varför?). De funktioner vi kan addera till de affina som svar på frågan kommer därför samtliga att vara funktioner som är oändliga någonstans.

Håller vi oss till *ändliga* funktioner är alltså de affina funktionerna de enda som är konvexa och konkava. Det sammanfattar vi:

$$\begin{gathered} f \text{ är ändlig, konvex och konkav} \\ \Longleftrightarrow \\ f \text{ är affin} \\ \Longleftrightarrow \\ \texttt{epi}\ f \cap \texttt{hypo}\ f \text{ är ett icke-vertikalt hyperplan i } \Re^{n+1}. \end{gathered}$$

Inte nog med det, en funktion som är konvex men är lika med $-\infty$ någonstans *kan* inte vara ändlig i stort någonstans alls. Funktioner som är $-\infty$ någonstans kallas *improper*, och enligt en sats i Rockafellars konvex analys-bok (7.2) kan en sådan funktion bara vara ändlig på den relativa randen av det område där funktionen är $< \infty$.

Nu går vi djupare än vi kanske behöver, och det viktiga här är att det beroende på att det finns två typer av hyperplan i $\Re^{n+1}$, varav en är känd som representant för de affina funktionerna, så finns det alltså ytterligare en grupp av konvexa och konkava funktioner. Låt oss studera dessa funktioner i $\Re^1$, och sedan anger vi karakteriseringen på $\Re^n$.

Exempel i $\Re^1$

Betrakta följande funktion definierad på $\Re^1$:

$$f(x) = \begin{cases} -\infty, & x < a, \\ b, & x = a, \\ +\infty, & x > a. \end{cases}$$

Här är

$$\begin{aligned} \texttt{epi}\, f &= \{\, (x, \alpha) \in \Re^2 \mid x \leq a \,\} \setminus \{\, (x, \alpha) \in \Re^2 \mid x = a; \quad \alpha < b \,\}, \\ \texttt{hypo}\, f &= \{\, (x, \alpha) \in \Re^2 \mid x \geq a \,\} \setminus \{\, (x, \alpha) \in \Re^2 \mid x = a; \quad \alpha > b \,\}. \end{aligned}$$

I det här fallet är epigrafen inte en sluten mängd beroende på att f inte är kontinuerlig (ja, egentligen att den inte ens är halv-kontinuerlig [lower semi-continuous]). Epigrafen är ett halvrum med en halvlinje på randen borttagen; det är fortfarande en konvex mängd. Samma resonemang visar att hypografen är konvex mängd, och alltså är f både konvex och konkav.

För att definiera det hyperplan som karakteriserar f sluter vi epigrafen och hypografen; vi får:

$$\texttt{cl}\ (\texttt{epi}\ f) \cap \texttt{cl}\ (\texttt{hypo}\ f) = \{\, (x, \alpha) \in \Re^2 \mid x = a \,\},$$

vilket är ett vertikalt hyperplan i $\Re^2$.

Det allmänna fallet

I det allmänna fallet har vi följande resultat:

$$\begin{gathered} f \text{ är improper, konvex och konkav (med minst ett ändlig funktionsvärde)} \\ \iff \\ \texttt{cl}\ (\texttt{epi}\ f) \cap \texttt{cl}\ (\texttt{hypo}\ f) \text{ är ett vertikalt hyperplan i } \Re^{n+1}. \end{gathered}$$

Har vi glömt något??

Ja, det finns två funktioner till som är både konvexa och konkava. De motsvarar fallet ovan, där snittet mellan epigraf och hypograf är tomma mängden.

Fall 1: $f \equiv -\infty$.

Fall 2: $f \equiv \infty$.

Sammanfattning

Det finns tre subfamiljer av konvexa och konkava funktioner. Den första och mest kända är familjen av ändliga konvexa och konkava funktioner, även kallade de affina funktionerna, som karakteriseras av de icke-vertikala hyperplanen i $\Re^{n+1}$. Den andra är familjen av oändliga (improper) konvexa och konkava funktioner som har (minst) ett ändligt funktionsvärde, och karakteriseras av de vertikala hyperplanen i $\Re^{n+1}$. Den tredje är familjen av improper konvexa och konkava funktioner som inte är ändliga någonstans (en familj bestående av två funktioner.)

En komplett karakterisering av klassen av konvexa och konkava funktioner på $\Re^n$ sammanfattar detta, och besvarar Pysselfråga 1:

$$
\begin{gathered}
f \text{ är konvex och konkav på } \Re^n \\
\Longleftrightarrow \\
\texttt{cl}\,(\texttt{epi}\; f) \cap \texttt{cl}\,(\texttt{hypo}\; f) \text{ är tom eller ett hyperplan i } \Re^{n+1}.
\end{gathered}
$$

Så nu vet Ni vad Ni ska svara när en student ställer frågan!

Vi noterar att analysen hade varit svårare om vi hade använt den traditionella definitionen av konvexitet som vi använder i vår undervisning, eftersom den bara är giltig för ändliga funktioner. Vill vi ta hand om utvidgade funktioner, måste vi införa konventionen $\infty - \infty = \infty$ (i fallet konvexitet; $\infty - \infty = -\infty$ i fallet konkavitet), eller göra om definitionen så att sådana fall inte dyker upp. Se Rockafellars konvax analys-bok (Sektion 4) för precisa definitioner. Det var naturligtvis enklare att använda epigrafen också därför att det var möjligt att använda geometriska argument.

Det blev mer text och stök än jag hade räknat med. Men jag hoppas att den första pysselhörnan rynkade pannorna på Er, och att Ni hade intressanta diskussioner i “Not the Coffee Room.”

Vi avslutar som vanligt med att tillsammans sjunga den glada signaturmelodin. Vinka med Era roliga hattar till tittarna! Vi ses nästa vecka. Hej då!