SENIOR 2010 - ADVANCED - A

Titolo nota 08/09/2010

Più moltiplicatori di Lagrange

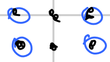
$$V = \{x \in \mathbb{R}^n : \phi_1(x) = \dots = \phi_k(x) = 0\}$$ k equazioni $(k < n)$

$f : \mathbb{R}^n \to \mathbb{R}$ $\qquad \max / \min \{ f(x) : x \in V \}$

Teorema Se $x_0 \in V$ è p.to di max/min, allora succede una di queste 3 cose

- ∇f o qualche $\nabla \phi_i$ non esiste in x_0
- Costruiamo la matrice che ha come righe i k gradienti

$$\begin{matrix} \nabla \phi_1 \\ \nabla \phi_2 \\ \vdots \\ \nabla \phi_k \end{matrix} \qquad \text{Matrice } k \times n$$



[Rango di una matrice = dimensione della massima sottomatrice quadrata $r \times r$ con $\det \neq 0$
= max numero di righe lin. indip.
= " " " colonne " "
= dimensione dell'immagine]

In x_0 il rango è $< k$ (cioè tutte le sottomatrici $k \times k$ hanno determinante $= 0$ $\binom{n}{k}$ condizioni)

- In x_0 si ha che ∇f è combinazione lineare dei $\nabla \phi_i$, cioè $\exists \lambda_1, \dots, \lambda_k$ t.c.

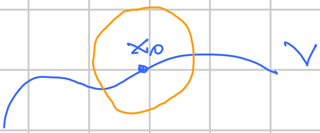
$$\nabla f(x_0) = \boxed{\lambda_1} \nabla \phi_1(x_0) + \dots + \boxed{\lambda_k} \nabla \phi_k(x_0)$$

moltiplicatori

— o — o —

Idea per una dim. con 1 moltiplicatore:

$V = \{x \in \mathbb{R}^n : \phi(x) = 0\}$ $\qquad x_0 \in V$ p.to di min per $f(x)$ ristretta a V



Prendo $B(x_0, 1)$ chiusa

Per ogni n calcolo $\min \{ f(x) + n\,\phi^2(x) : x \in B(x_0, 1) \}$

Esiste ed è raggiunto in un p.to x_n. Si può dimostrare che $x_n \to x_0$.

In x_n si ha che $\nabla (f(x) + n\,\phi^2(x)) = 0$, cioè $\nabla f(x) + 2n\,\phi(x) \nabla \phi(x) = 0$

cioè in x_n $\nabla f(x_n)$ è multiplo di $\nabla \phi(x_n)$. Passo al limite

Lemma sui vettori $\quad v_n \to v_\infty \quad w_n \to w_\infty \quad \forall n$ si ha che v_n è multiplo di w_n

Allora $\to$ o $w_\infty = 0$

o v_∞ è multiplo di w_∞

— o — o —

Con k equazioni la dim. è la stessa

$$\min \{ f(x) + n\,\Phi_1^2(x) + \dots + n\,\Phi_k^2(x) \}$$

Nel p.to di min x_n si ha che

$$\nabla f(x_n) + 2n\,\Phi_1(x_n)\nabla\Phi_1(x_n) + \dots + 2n\,\Phi_k(x_n)\nabla\Phi_k(x_n) = 0$$

cioè $\nabla f(x_n)$ è comb. lineare dei $\nabla\Phi_i(x_n)$

Lemma Date $k+1$ succ. di vettori

$$v_n \to v_\infty$$
$$w_n^1 \to w_\infty^1$$
$$\vdots$$
$$w_n^k \to w_\infty^k$$

Hp: per ogni n si ha che v_n è comb lin. dei w_n^i

Tesi: $\to$ o il limite è ancora comb lin.

$\to$ o i w_∞^i non sono lin. indip., cioè la matrice da essi formata non ha rango max.

Dim. Ipotesi dice che la matrice $\begin{pmatrix} v_n \\ w_n^1 \\ \vdots \\ w_n^k \end{pmatrix}$ $(k+1) \times n$

ha rango $\leq k$, cioè tutte le s. matrici $(k+1)\times(k+1)$ hanno det $=0$

Quindi anche la matrice limite ha rango $\leq k$, cioè

$\to$ o i k sotto sono dip. tra di loro

$\to$ o il primo dipende dagli altri.

— o — o —

A meno di cambiare $f(x)$ con $f(x) + \mathrm{dist}^{10}(x, x_0)$ possiamo supporre x_0 p.to di minimo stretto (basta il $\square$).

— o — o —

Problemi ad applicare i moltiplicatori

$\to$ sapere che max/min esistono (guadagnare compattezza)

$\to$ bordi dei bordi

$\to$ risolvere il sistema!

— o — o —

CONVESSITÀ

Def. Sia $A \subseteq \mathbb{R}^n$ un insieme convesso (per ogni coppia di p.ti contiene tutto il segmento)

$f: A \to \mathbb{R}$ si dice convessa se

$$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y) \qquad \forall x \in A \ \forall y \in A \ \ \forall \lambda \in [0,1]$$

Oss. In una variabile si può verificare con $f''(x) \geq 0$
In più variabili ci sarebbe la matrice Hessiana, oppure sfruttando che è somma di funzioni convesse semplici (distanza da un p.to, distanza^2 da un p.to, distanza da una retta, ...)

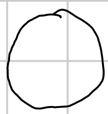
Oss. f è convessa $\iff$ il sopragrafico $\{(x,y) \in A \times \mathbb{R} : y \geq f(x)\}$ è un insieme convesso di $\mathbb{R}^{n+1}$

—o—o—

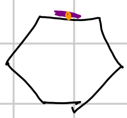
Teorema Se $A \subseteq \mathbb{R}^n$ è un convesso compatto, allora
$\max\{f(x) : x \in A\}$,
se f è convessa e continua, esiste e tra i p.ti di max c'è almeno un p.to estremale di A (solo estremali se f è stretta convessa)

Def. $x_0 \in A$ è p.to estremale $\iff$ quando $x_0 \in$ segmento $\subseteq A$ si ha che x_0 è un estremo del segmento

Esempi



p.ti estremali = bordo



p.ti estremali = vertici

—o—o—

Esempio Trovare in un triangolo il p.to per cui è massima la somma delle distanze dai vertici

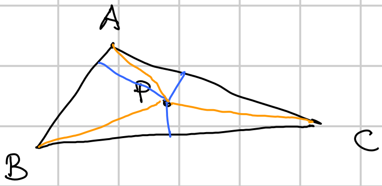
Soluzione La somma delle distanze è convessa (anzi strettamente convessa perché i 3 vertici non sono allineati ...), quindi il max si realizza in un p.to estremale, cioè in un vertice (quello opposto al lato minore)

Come si dim. per bene che è strett. convessa

$$f(x) = \text{dist}(x, A) + \text{dist}(x, B) + \text{dist}(x, C).$$

La disug. "$\le$" nella definizione è facile. Se ci fosse "$=$" i p.ti x e y dovrebbero essere allineati con A, B, C (il che è impossibile). Infatti dovrebbe esserci "$=$" sui 3 addendi.

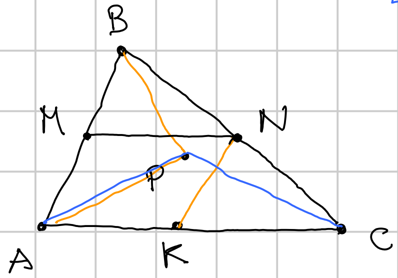
<u>Esempio 2</u> La somma delle distanze dai lati di un Δ è convessa, anzi è affine nel triangolo (il grafico è un piano, perché somma di 3 piani).
Quindi il max è almeno in un vertice, quindi è l'altezza + lunga



<u>Esempio 3</u> IMOSL 99 Dim. che in un triangolo ABC se prendo un p.to P si ha che

$$\underbrace{AP + BP + CP + \min\{AP, BP, CP\}}_{LHS} \le AB + BC + CA$$

<u>Dim.</u>



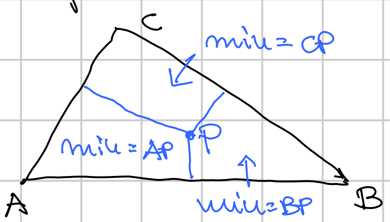
Facile: $AP + CP \le AM + MN + NC$
$AP + BP \le BN + NK + AK$

Sommo:

$$LHS \le 2AP + BP + CP \le \tfrac{1}{2}AB + \tfrac{1}{2}AC + \tfrac{1}{2}BC + \tfrac{1}{2}BC + \tfrac{1}{2}AB + \tfrac{1}{2}AC = \text{Perim.}$$

Osservazione: è sempre vero che P è dentro almeno 2 "trapezi"

$\boxed{\text{Dim. "bovina"}}$ Pongo $f(P) = AP + BP + CP + \min\{AP, BP, CP\}$
È convessa? FORGET IT!! Su tutto il triangolo NO, ma lo è nei 3 pezzi in cui il minimo so chi è.



Nelle 3 zone $f(P)$ è convessa (strett.) perché somma di convesse

$2AP + BP + CP$

Basta quindi controllare i p.ti estremali delle 3 zone, cioè
→ vertici
→ p.ti medi dei lati
→ circocentro
Da qui si finisce.

— o — o —

Equivalente in + variabili di $f''(x)$. Derivate parziali seconde! (sono n^2)

$$\begin{pmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{pmatrix}$$

La matrice è simmetrica

<u>Def.</u> Una Matrice simmetrica $n \times n$ si dice <u>semidefinita positiva</u> se

$$x \cdot Ax \geq 0 \qquad \forall x \in \mathbb{R}^n$$

<u>Def.</u> Forma quadratica in $\mathbb{R}^n$ = somma di monomi di 2° grado in n variabili

Esempio $(n=2)$ $\quad q(x,y) = ax^2 + by^3 + 2c\,xy$

$(n=3)$ $\quad q(x,y,z) = ax^2 + by^2 + cz^2 + 2dxy + 2eyz + 2fzx$

Una forma è semidefinita pos. se viene ≥ 0 per ogni $x \in \mathbb{R}^n$

Forme ⟺ matrici simmetriche

$$ax^2 + by^2 + 2cxy \qquad (x \;\; y) \begin{pmatrix} a & c \\ c & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

<u>Teorema</u> f è convessa in A "⟺" la matrice delle derivate seconde è semidefinita positiva

se ci sono le derivate

Come stabilire se una forma è positiva, negativa o "indefinita"
→ autovalori della matrice
→ ogni forma è somma / differenza di quadrati

$$x^2 + 3xy - y^2 = x^2 + 2x\frac{3y}{2} + \frac{9}{4}y^2 - \frac{9}{4}y^2 - y^2 = \left(x + \frac{3y}{2}\right)^2 - \frac{13}{4}y^3$$

→ condizione sufficiente (ma non necessaria) per essere positiva

tutte det > 0

—— o —— o ——

DISUG. DI JENSEN $\quad I \subseteq \mathbb{R}$ convesso $\quad f: I \to \mathbb{R}$ convessa

$$f(\lambda_1 x_1 + \dots + \lambda_n x_n) \le \lambda_1 f(x_1) + \dots + \lambda_n f(x_n)$$

$x_1, \dots, x_n \in I$

$\lambda_1 + \dots + \lambda_n = 1$

$\lambda_i \ge 0$

Dim. Induzione UP and DOWN

$\boxed{1}$ Vera per $n \Rightarrow$ vera per $2n$ $\quad$ ($n=2$: def. di funzione convessa)

$\boxed{2}$ Vera per $n \Rightarrow$ vera per $(n-1)$: basta usare un λ nullo

Dim. $\boxed{1}$ $\quad f(\overbrace{\lambda_1 x_1 + \dots + \lambda_n x_n}^{\lambda x} + \overbrace{\lambda_{n+1} x_{n+1} + \dots + \lambda_{2n} x_{2n}}^{(1-\lambda) y}) =$

$$f\left((\lambda_1 + \dots + \lambda_n)\frac{\lambda_1 x_1 + \dots + \lambda_n x_n}{\lambda_1 + \dots + \lambda_n} + (\lambda_{n+1} + \dots + \lambda_{2n})\frac{\lambda_{n+1} x_{n+1} + \dots}{\lambda_{n+1} + \dots + \lambda_{2n}}\right)$$

$$\le (\lambda_1 + \dots + \lambda_n) f(\quad) + (\lambda_{n+1} + \dots + \lambda_{2n}) f(\quad)$$

Ora uso Jensen su n vero per ipotesi induttiva (i coefficienti sono $\frac{\lambda_1}{\lambda_1 + \dots + \lambda_n}, \dots, \frac{\lambda_n}{\lambda_1 + \dots + \lambda_n}$ che fanno somma 1

Occhio a trattare il caso della divisibilità per zero ...

Oss. Caso particolare $\quad f\left(\frac{x_1 + \dots + x_n}{n}\right) \le \frac{f(x_1) + \dots + f(x_n)}{n}$

Volendola dim. direttamente faccio $\boxed{1}$ come prima, ma per $\boxed{2}$ è più complicato:
sono dati $x_1, \dots, x_{n-1}$ e devo scegliere $x_n = \frac{x_1 + \dots + x_{n-1}}{n-1}$.
Mi ritrovo

$$\underbrace{f\left(\frac{x_1 + \dots + x_{n-1}}{n-1}\right)}_{LHS} = f\left(\frac{x_1}{n} + \dots + \frac{x_{n-1}}{n} + \frac{x_n}{n}\right)$$

$$x_1\left(\frac{1}{n} + \frac{1}{n(n-1)}\right) = x_1 \frac{\cancel{n}}{\cancel{n}(n-1)}$$

$$\le \frac{1}{n} f(x_1) + \dots + \frac{1}{n} f(x_{n-1}) + \underbrace{\frac{1}{n} f(x_n)}_{\frac{1}{n} LHS}$$

Porto a sx e viene

LHS $\frac{n-1}{n}$ $\leq$ RHS

— o —

Applicazione 1 AM – QM (Jensen su $f(x) = x^2$)
(N.B. vale senza Hp di segno)

Applicazione 2 GM – AM (Jensen sulla concava $f(x) = \log x$;
qui servono $x_i > 0$)

$$\log\left(\frac{x_1 + \dots + x_n}{n}\right) \overset{\text{concava}}{\geq} \frac{1}{n}\left(\log x_1 + \dots + \log x_n\right)$$

$$= \log\left[x_1 \cdot \ldots \cdot x_n\right]^{\frac{1}{n}}$$

Usando i λ si ottengono versioni pesate.

Applicazione 3 $M_p(x_1, \dots, x_n) \leq M_q(x_1, \dots, x_n)$ $p \leq q$
ci si riduce facilmente a $0 < p < q$ e $p = 1$ $q = \alpha > 1$
Ora è Jensen su x^α che è convessa per $\alpha \geq 1$ $(x \geq 0)$

Applicazione 4 $ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q$ se $a > 0, b > 0, p > 0, q > 0$
$\frac{1}{p} + \frac{1}{q} = 1$

Jensen su $\log x = f(x)$ $x = a^p$ $y = b^q$ $\lambda = \frac{1}{p}$ $(1-\lambda) = \frac{1}{q}$

$$\log(\lambda x + (1-\lambda) y) \geq \lambda f(x) + (1-\lambda) f(y)$$

$$\log\left(\frac{1}{p}a^p + \frac{1}{q}b^q\right) \geq \cancel{\frac{1}{p}} \log(a^{\cancel{p}}) + \cancel{\frac{1}{q}} \log(b^{\cancel{q}})$$

$$= \log a + \log b = \log(ab)$$

Analogo $abc \leq \frac{1}{p}a^p + \frac{1}{q}b^q + \frac{1}{r}c^r$ $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$

Serve per dimostrare C–S a 3 specie (Hölder)

$$\sum a_i b_i c_i \leq \left(\sum a_i^p\right)^{\frac{1}{p}} \left(\sum b_i^q\right)^{\frac{1}{q}} \left(\sum c_i^r\right)^{\frac{1}{r}} \qquad \frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$$

Per omogeneità posso assumere
$\sum a_i^p = \sum b_i^q = \sum c_i^r = 1$ e poi sommo sui i le YOUNG.

Caso speciale : $i \in \{1,2\}$ $\qquad p_1 = p_2 = \dots = p_n = n$

$$(a_1 \cdot \ldots \cdot a_n) + (b_1 \cdot \ldots \cdot b_n) \leq (a_1^n + b_1^n)^{\frac{1}{n}} \cdot \ldots \cdot (a_n^n + b_n^n)^{\frac{1}{n}}$$

$$= \sqrt[n]{(a_1^n+b_1^n) \cdot \ldots \cdot (a_n^n + b_n^n)}$$

(a_1, b_1)
(a_2, b_2)
$\vdots$
(a_n, b_n)

Dimostrazione diretta : UP and DOWN

Applicazione 5 $\qquad \min \{ f(x_1) + \dots + f(x_n) : x_1 + \dots + x_n = A \}$

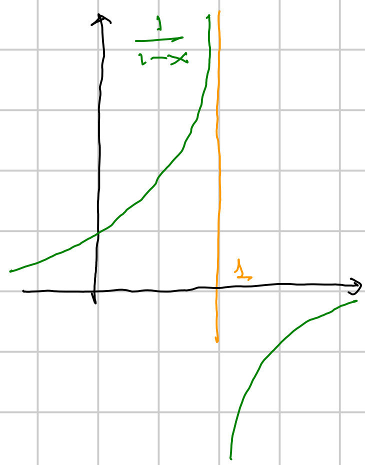
Se f è convessa, allora il minimo è quando sono tutti uguali

$$f(x_1) + \dots + f(x_n) \geq n f\left(\frac{x_1 + \dots + x_n}{n}\right) = n f\left(\frac{A}{n}\right)$$

Nesbit $\qquad \frac{a}{b+c} + \frac{b}{c+a} + \frac{c}{a+b} \geq \frac{3}{2} \qquad a, b, c > 0$

Per omogeneità assumo $a+b+c = 1$ $\qquad \frac{a}{1-a} + \frac{b}{1-b} + \frac{c}{1-c} \geq \frac{\frac{1}{3}}{1-\frac{1}{3}} 3$

Se $f(x) = \frac{1}{1-x}$ è convessa ho finito



Rigoroso : $f''(x) = \dots$ oppure

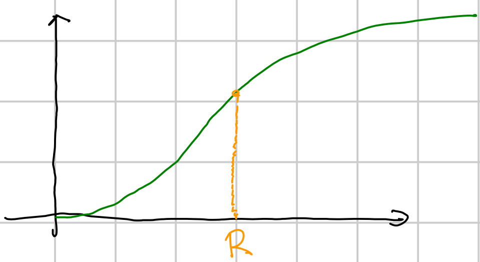
cambio variabile e diventa $\frac{1}{x}$ per la quale uso la definizione.

Achtung! Precisare sempre DOVE si lavora.
In questo caso : $[0, 1)$

— o — o —

$\boxed{\text{LEMMA CONVEX-CONCAVE}}$ $\qquad \min \{ f(x_1) + \dots + f(x_n) : x_1 + \dots + x_n = A, \; x_i \geq 0 \}$

Supponiamo f convex-concave. Per il minimo se la giocano 2 configurazioni



→ $x_1 = \dots = x_n$ nella zona convessa

→ $x_1 = \dots = x_{n-1}$ nella zona convessa e x_n nella zona concava per aggiustare la somma

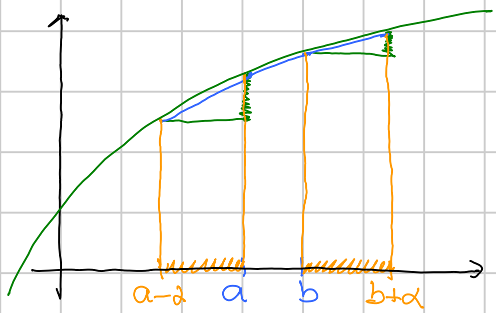
Ponendo $x = x_n$ mi riduco a studiare

$$f(x) + (n-1)\, f\left(\frac{A-x}{n-1}\right)$$

oppure, ponendo $x_i = x$: $(n-1)\, f(x) + f(A-(n-1)x)$
e si può fare studiando la funzione.
Idea alla base del convex-concave:

$\boxed{1}$ min esiste se f è continua in $[0,A]$: Weierstrass.
$\boxed{2}$ tutti gli x_i che stanno nella zona convessa si possono rendere uguali conservando la somma ed abbassando $\sum f(x_i)$
$\boxed{3}$ A max ce n'è uno nella zona concava: se ce ne sono 2 li "allargo" conservando la somma fino a quando il sx cade nella zona convessa (il dx non può arrivare ad A senza che tutti gli altri siano zero)

<u>Lemma di allargamento</u> (versione concava)



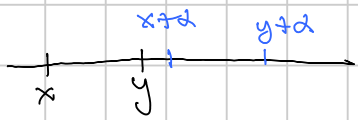
$$f(a-\alpha) + f(b+\alpha) \le f(a) + f(b)$$

il che equivale a

$$\underbrace{f(b+\alpha) - f(b)}_{\alpha f'(c_1)} \le \underbrace{f(a) - f(a-\alpha)}_{\alpha f'(c_2)}$$

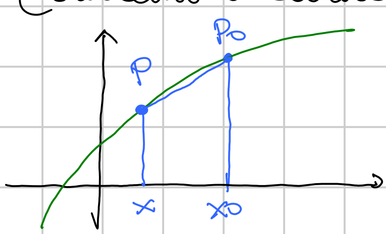
Più in generale: se $y \ge x$ e $\alpha \ge 0$, allora

$$f(y+\alpha) - f(y) \le f(x+\alpha) - f(x)$$



Non ho garanzia di ordinamento fra c_1 e c_2.

Idea: fissato un punto x, il rapporto incrementale è monotono (crescente o decrescente a seconda di convessità concavità)



il coeff. angolare della retta PP_0 diminuisce quando P sale (fissato P_0)
(si dimostra a partire dalla def. di convessità)

DISUGUAGLIANZA DI KARAMATA

$f: I \to \mathbb{R}$ convessa

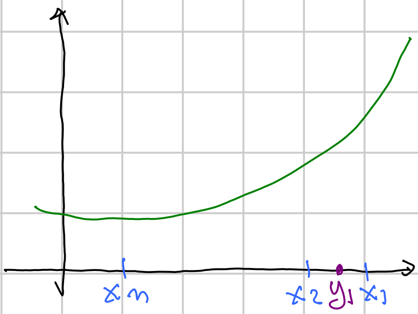
$(x_1, \dots, x_n) \succ (y_1, \dots, y_n)$ in stile bunching, cioè

- $x_1 \geq y_1$
- $x_1 + \dots + x_k \geq y_1 + \dots + y_k \qquad k = 1, \dots, n-1$
- $x_1 + \dots + x_n = y_1 + \dots + y_n$ (controllare !!!!)
- $x_1 \geq x_2 \geq \dots \geq x_n, \quad y_1 \geq y_2 \geq \dots \geq y_n$

Allora: $\quad f(x_1) + \dots + f(x_n) \geq f(y_1) + \dots + f(y_n)$

KARAMATA $\Rightarrow$ JENSEN $\qquad$ Ordino gli x_i e poi uso $y_1 = \dots = y_n = \dfrac{x_1 + \dots + x_n}{n}$

Idea della dim: induzione + Lemma di allargamento



Prendo x_1 e lo sposto verso y_1 e contemporaneamente alzo x_n finché non ho "contatti" tra x e y.
Dopo il contatto semplifico e scendo di uno. Così però magari la somma $x_1 + x_2 + x_3$ è scesa ...
Forse è meglio lavorare con x_1 e x_2 ...
Alternativa: spostare x_1 e y_1 a sx fino a quando ...

— o — o —

NB. Le funzioni convesse stanno sopra le proprie rette tangenti !!

Idea: derivata = rapporto incrementale con $P = P_0$

Applicazioni (Bernoulli con esponente $\alpha \geq 1$)

$$(1+x)^\alpha \geq 1 + \alpha x \qquad \forall x \geq -1 \quad \forall \alpha \geq 1$$

$f(x) = (1+x)^\alpha$ è convessa per $\alpha \geq 1$ $\qquad f(x) \geq \underbrace{1 + \alpha x}_{\text{retta tg. in } x=0}$

$f(0) = 1 \qquad f'(x) = \alpha (1+x)^{\alpha - 1} \qquad f'(0) = \alpha$