Universidade Federal do Paraná
Setor de Ciências Exatas
Curso de Graduação em Matemática Industrial

Otimização

PRINCIPAIS PROPRIEDADES DE FUNÇÕES CONVEXAS

Vanessa Hlenka

Orientadora : Elizabeth Wegner Karas
Co-orientador: Ademir Alves Ribeiro
Área de Concentração: Matemática Aplicada

Curitiba
2006

RESUMO

Este trabalho de iniciação científica está inserido num projeto de pesquisa em Otimização Contínua e suas aplicações. Trataremos de modo geral o problema de minimizar uma função com restrições de igualdade e desigualdade. Em problemas de programação não linear, obter otimizadores globais é muito difícil, quando não impossível. Normalmente, ficamos satisfeitos com o estudo de otimizadores locais. No entanto é possível estudar condições sobre as características do problema de modo que o minimizador local seja global. Uma dessas principais características é a convexidade, tanto da região viável como das funções envolvidas.

A partir de conceitos de análise real e em $\mathbb{R}^n$, obteremos algumas propriedades importantes de convexidade, nosso foco de análise. Primeiramente estudaremos a convexidade na reta, depois estenderemos o estudo de algumas propriedades para o $\mathbb{R}^n$. Sempre que possível, ilustraremos este trabalho com exemplos e utilizaremo-nos do auxílio do Matlab para plotar alguns gráficos e então identificar essas propriedades geometricamente.

Sumário

Introdução

Um importante conceito no contexto de otimização é a convexidade, propriedade que, quando presente numa função, garante que seu mínimo local é global. Seja I um intervalo não vazio de $\mathbb{R}$. A função $f : I \to \mathbb{R}$ é dita convexa em $\mathbb{R}$ quando

$$f(\alpha x + (1-\alpha)x') \leq \alpha f(x) + (1-\alpha)f(x'),$$

para todos os pares de pontos $(x,\, x')$ em I e todo $\alpha \in\,]0,1[$.

Podemos reconhecer uma função convexa de outras maneiras, sendo uma delas observar o epigrafo de f, que é definido por

$$\operatorname{epi} f := \{(x,r) \mid x \in \operatorname{Dom} f,\ r \geq f(x)\},$$

é um conjunto convexo.

A desigualdade que define uma função convexa f pode ser generalizada para mais que dois pontos: Para qualquer coleção $\{x_1,\ldots,x_k\}$ de pontos em I e qualquer coleção de números $\{\alpha_1,\ldots,\alpha_k\}$ satisfazendo $\alpha_i \geq 0$ para $i = 1,\ldots,k$ e $\sum_{i=1}^{n} \alpha_i = 1$, vale a desigualdade de Jensen:

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i).$$

Definições básicas como as citadas acima serão apresentadas inicialmente. A seguir serão estudadas algumas das propriedades de funções convexas em $\mathbb{R}^2$, dentre as quais a continuidade no interior de seu domínio.

Se $f \in \operatorname{Conv} \mathbb{R}$, então f é contínua no interior do $\operatorname{Dom} f$. Além disso, para cada intervalo compacto $[a,b]$ contido no interior do $\operatorname{Dom} f$, existe $L \geq 0$ tal que

$$|f(x) - f(x')| \leq L|x - x'|,$$

para todos x e x' em $[a,b]$. A constante L é chamada constante de Lipschitz e f então é dita localmente Lipschitziana.

Dizemos que $f \in \operatorname{Conv} \mathbb{R}$ é fechada, ou semi-contínua inferior, se

$$\liminf_{x \to x_0} f(x) \geq f(x_0),$$

para todo $x_0 \in \mathbb{R}$. Funções convexas fechadas são de fundamental importância em análise convexa e otimização. Por exemplo, se a função f for semi-contínua inferior, e se o conjunto C for fechado, o problema $\min\{f(x) \mid x \in C\}$ admite uma solução.

Apresentaremos outras maneiras de identificar funções convexas, como através da diferenciabilidade, primeiramente para funções em $\mathbb{R}^2$ e logo após em $\mathbb{R}^n$.

Capítulo 1

Funções convexas de uma variável real

Funções convexas de uma variável real formam uma importante classe de funções no contexto geralmente chamado de análise real. Elas são muito utilizadas em otimização e também em diversas áreas de matemática aplicada.

As definições e propriedades da seção a seguir podem ser encontradas em [3].

1.1 Definições básicas e exemplos

Os intervalos formam os exemplos mais simples de subconjuntos de $\mathbb{R}$. Um subconjunto $I \in \mathbb{R}$ é um intervalo se e somente se, sempre que x_0 e x_1 pertencem a I, uma das seguintes propriedades vale:

(i) todo ponto entre x_0 e x_1 pertence a I (definição baseada na relação de ordem de $\mathbb{R}$);

(ii) todo α entre 0 e 1, o ponto $\alpha x_0 + (1-\alpha)x_1$ pertence a I (definição usando a estrutura de vetor de $\mathbb{R}$).

A seguinte classificação de intervalos não-vazios é conveniente:

- Os intervalos compactos: $I = [a, b]$, ($a, b \in \mathbb{R}$ com $a \leq b$);
- Os intervalos limitados mas não fechados: $[a, b), (a, b], (a, b)$ com $a, b \in \mathbb{R}$, $a < b$;
- Os intervalos ilimitados: $(-\infty, b]$ e $(-\infty, b)$, $[a, +\infty)$ e $(a, +\infty)$ com $a, b \in \mathbb{R}$;
- O próprio intervalo $\mathbb{R}$, que é um conjunto abrto e fechado.

Intervalos fechados são também chamados de segmentos, ou segmentos de reta. A seguinte representação paramétrica, ilustrada na Figura 1.2, é clássica para um ponto $u \in (a, b)$:

$$u = \alpha b + (1-\alpha)a = a + \alpha(b-a) \text{ com } \alpha = \frac{u-a}{b-a} \in (0,1). \tag{1.1}$$

Observação 1.1 *u definido desta maneira é dito combinação convexa de a e b.*

Finalmente, lembramos a definição básica de uma função $f : D \to \mathbb{R}$.

Definição 1.2 *O gráfico de f é o subconjunto de $D \times \mathbb{R}$*

$$\mathrm{gr}\, f := \{(x, r) |\ x \in D \ \text{ e } \ r = f(x)\}.$$

O epigrafo de f é "tudo o que está acima do gráfico":

$$\mathrm{epi}\, f := \{(x, r) |\ x \in \mathbb{R} \ \text{ e } \ r \geq f(x)\}.$$

O epigrafo estrito é definido igualmente, substituindo "$\geq$" por "$>$".

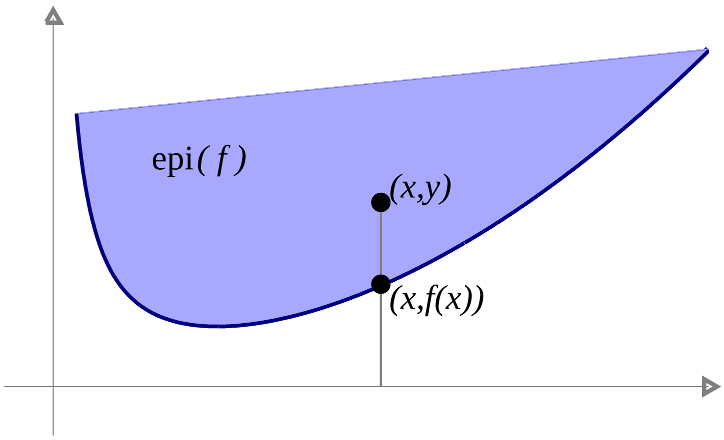

Figura 1.1: Epigrafo de uma função f. O epigrafo de uma função convexa é um conjunto convexo.

1.1.1 Primeiras definições de uma função convexa

A primeira definição de uma função convexa é a que segue:

Definição 1.3 (Analítica) *Seja I um intervalo não vazio de $\mathbb{R}$. A função $f : I \to \mathbb{R}$ é dita convexa em I quando*

$$f(\alpha x_0 + (1 - \alpha)x_1) \leq \alpha f(x_0) + (1 - \alpha)f(x_1), \tag{1.2}$$

para todos os pares de pontos $(x_0,\, x_1)$ em I e todo $\alpha \in\ (0, 1)$.

A função f é dita estritamente convexa quando vale a desigualdade estrita em (1.2) para $x_0 \neq x_1$.

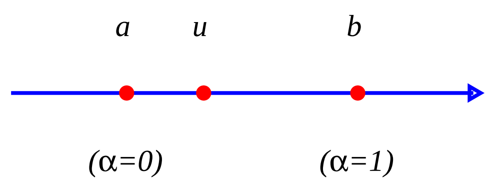

Figura 1.2: Parametrização de um intervalo

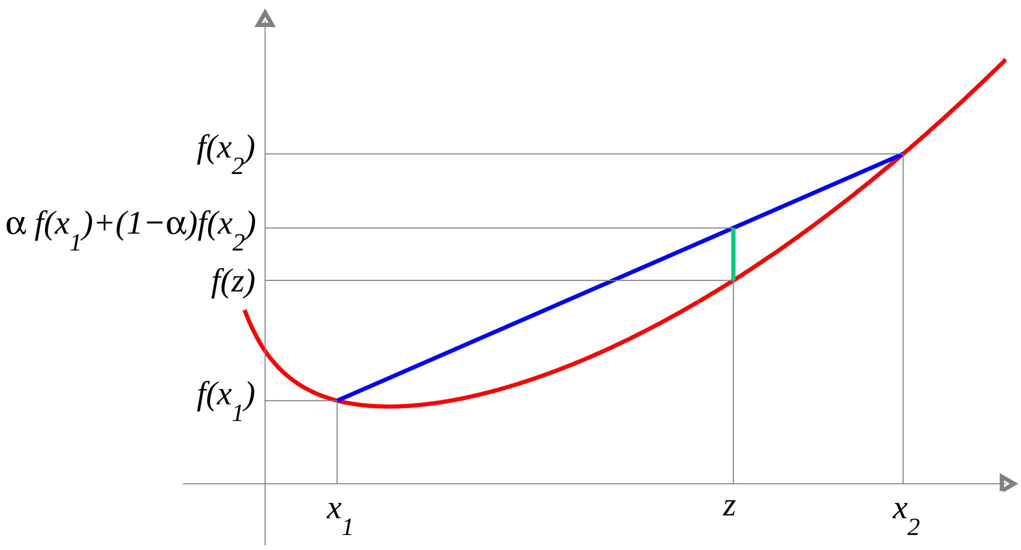

Figura 1.3: Interpretação geométrica da definição analítica de função convexa.

O significado geométrico de convexidade é claro: considere na Figura 1.3 o segmento que une o ponto $(x_1, f(x_1))$ ao ponto $(x_2, f(x_2))$. Dizer que f é convexa significa que, para todos x_1, x_2 em I e todo z em (x_1, x_2), o ponto $(z, f(z))$ do gráfico de f está abaixo do segmento que une $(x_1, f(x_1))$ e $(x_2, f(x_2))$.

Definição 1.4 (Geométrica) *Seja I um intervalo não vazio de $\mathbb{R}$. A função $f : I \to \mathbb{R}$ é convexa em I se e somente se* $\operatorname{epi} f$ *é um subconjunto convexo de $\mathbb{R}^2$.*

A Figura 1.4 sugere que, se u está entre x_0 e x_1 e P_u está abaixo de $P_{x_0}P_{x_1}$, a inclinação de $P_{x_0}P_u$ (ou melhor: da reta que une P_{x_0} e P_u) é menor que a inclinação de $P_{x_0}P_{x_1}$, que por sua vez é menor que a inclinação de $P_uP_{x_1}$. O próximo resultado de geometria elementar esclarece o argumento.

Proposição 1.5 *Sejam $P_{x_0} = (x_0, y_0)$, $P_u = (u, v)$, e $P_{x_1} = (x_1, y_1)$ três pontos sobre $f : \mathbb{R} \to \mathbb{R}$, onde f é convexa, com $u \in (x_0, x_1)$. Então as três propriedades seguintes são equivalentes:*

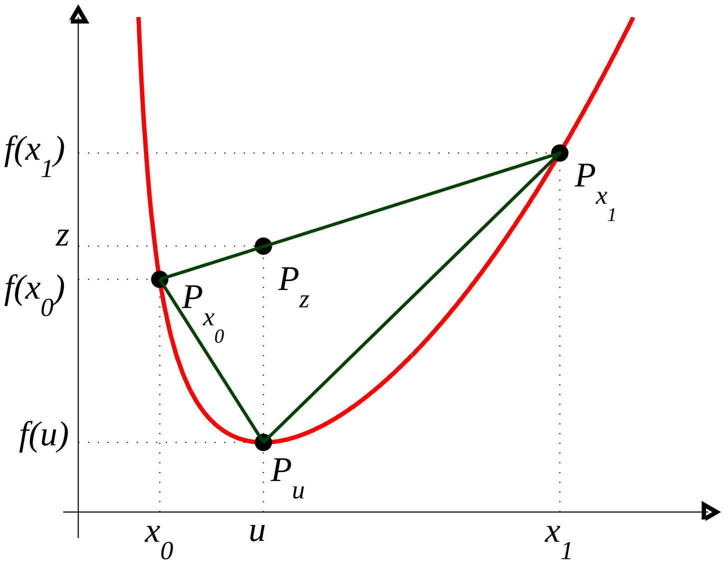

Figura 1.4: A propriedade fundametal de um epigrafo convexo

(i) P_u está abaixo de $P_{x_0}P_{x_1}$;

(ii) inclinação($P_{x_0}P_u$) $\leq$ inclinação($P_{x_0}P_{x_1}$);

(iii) inclinação($P_{x_0}P_{x_1}$) $\leq$ inclinação($P_uP_{x_1}$).

Prova. A Figura 1.4 ilustra esta demonstração.

(i)⇒(ii): A propriedade (i) quer dizer que a imagem de u por f está abaixo da reta que une P_{x_0} e P_{x_1} (e passa por P_z). A equação desta reta é dada por

$$y - z = \frac{y_1 - y_0}{x_1 - x_0}(x - u). \tag{1.3}$$

Como qualquer ponto da reta que une P_{x_0} e P_{x_1} satisfaz a equação (1.3), em particular o ponto P_{x_0}, façamos $x = x_0$ e $y = y_0$; então temos

$$y_0 - z = \frac{y_1 - y_0}{x_1 - x_0}(x_0 - u).$$

Logo, a imagem de u pela reta é dada por

$$z = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(u - x_0).$$

Sabemos que a imagem de u por f é menor que a imagem de u pela reta dada por (1.3), ou seja, $v \leq z$, ou

$$v \leq y_0 + \frac{y_1 - y_0}{x_1 - x_0}(u - x_0).$$

Como $u - x_0 > 0$, temos

$$\frac{v - y_0}{u - x_0} \leq \frac{y_1 - y_0}{x_1 - x_0},$$

que é justamente a propriedade (ii).

(ii)⇒(iii): Temos que

$$\frac{v-y_0}{u-x_0} \leq \frac{y_1-y_0}{x_1-x_0}.$$

Como $u-x>0$ e $x_1-x_0>0$, podemos escrever

$$(v-y_0)(x_1-x_0) \leq (y_1-y_0)(u-x_0),$$

$$vx_1 - vx_0 - y_0x_1 + y_0x_0 \leq uy_1 - uy_0 - x_0y_1 + x_0y_0,$$

$$uy_0 - uy_1 - y_0x_1 \leq vx_0 - vx_1 - x_0y_1.$$

Somando y_1x_1 em ambos os membros da desigualdade, obtemos

$$y_1x_1 - y_0x_1 + uy_0 - uy_1 \leq y_1x_1 - x_0y_1 - vx_1 + vx_0,$$

$$x_1(y_1-y_0) - u(y_1-y_0) \leq y_1(x_1-x_0) - v(x_1-x_0),$$

$$(x_1-u)(y_1-y_0) \leq (y_1-v)(x_1-x_0).$$

Como $x_1-x_0>0$ e $x_1-u>0$, então

$$\frac{y_1-y_0}{x_1-x_0} \leq \frac{y_1-v}{x_1-u}.$$

(iii)⇒(i): Temos que

$$\frac{y_1-y_0}{x_1-x_0} \leq \frac{y_1-v}{x_1-u}.$$

Como $x_1-x_0>0$ e $x_1-u>0$, então

$$(y_1-y_0)(x_1-u) \leq (y_1-v)(x_1-x_0),$$

$$y_1x_1 - y_1u - y_0x_1 + y_0u \leq y_1x_1 - y_1x_0 - vx_1 + vx_0,$$

$$vx_1 - vx_0 - x_1y_0 \leq -x_0y_1 + y_1u - y_0u.$$

Somando x_0y_0 em ambos os membros da desigualdade, obtemos

$$vx_1 - vx_0 - x_1y_0 + x_0y_0 \leq x_0y_0 - x_0y_1 + y_1u - y_0u,$$

ou seja,

$$v(x_1-x_0) - y_0(x_1-x_0) \leq u(y_1-y_0) - x(y_1-y_0).$$

Logo, como $x_1-x_0>0$, temos

$$v \leq y_0 + \frac{y_1-y_0}{x_1-x_0}(u-x_0).$$

□

Em outras palavras, (ii) e (iii) acima significam

$$\frac{f(u)-f(x_0)}{u-x_0} \leq \frac{f(x_1)-f(x_0)}{x_1-x_0} \leq \frac{f(x_1)-f(u)}{x_1-u}, \tag{1.4}$$

expressão que pode ser obtida através da representação (1.1) de $u \in (x_0, x_1)$ aplicada na definição (1.2) de convexidade. Temos que

$$u = \alpha x_0 + (1 - \alpha)x_1.$$

Então,

$$\alpha = \frac{x_1 - u}{x_1 - x_0}$$

e

$$1 - \alpha = \frac{u - x_0}{x_1 - x_0}.$$

Substituindo em (1.2), obtemos

$$f(u) \leq \frac{x_1 - u}{x_1 - x_0} f(x_0) + \frac{u - x_0}{x_1 - x_0} f(x_1).$$

Como $x_1 - x_0 > 0$, então

$$f(u)(x_1 - x_0) \leq (x_1 - u)f(x_0) + (u - x_0)f(x_1),$$

$$f(u)x_1 - f(u)x_0 \leq x_1 f(x_0) - uf(x_0) + uf(x_1) - x_0 f(x_1).$$

Somando $x_0 f(x_0)$ em ambos os membros da desigualdade, obtemos

$$f(u)x_1 - f(u)x_0 + x_0 f(x_0) \leq x_0 f(x_0) + x_1 f(x_0) - uf(x_0) + uf(x_1) - x_0 f(x_1),$$

$$f(u)(x_1 - x_0) - f(x_0)(x_1 - x_0) \leq f(x_1)(u - x_0) - f(x_0)(u - x_0),$$

$$(f(u) - f(x_0))(x_1 - x_0) \leq (f(x_1) - f(x_0))(u - x_0).$$

Portanto, como $x_1 - x_0 > 0$,

$$f(u) \leq f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(u - x_0) = f(x_1) + \frac{f(x_0) - f(x_1)}{x_1 - x_0}(x_1 - u),$$

o que mostra a conexão entre (1.2) e relações de valor médio semelhantes à (1.4).

Combinando a definição geométrica de função convexa com a equivalência enunciada na Definição 1.4, chegamos à seguinte caracterização de convexidade:

Proposição 1.6 (Critério de inclinações crescentes) *Seja I um intervalo não vazio de $\mathbb{R}$. A função $f : I \to \mathbb{R}$ é convexa em I se e somente se, para todo $x_0 \in I$, a função-inclinação*

$$x \mapsto \frac{f(x) - f(x_0)}{x - x_0} =: s(x) \tag{1.5}$$

é crescente em $I \setminus \{x_0\}$.

Prova. Sejam $x_1, x_2 \in I$ com $x_1 < x_2$. Seja $x_0 \in I$, com $x_0 \neq x_1$ e $x_0 \neq x_2$. Temos três casos a considerar:

(i) $x_0 < x_1 < x_2$;

(ii) $x_1 < x_0 < x_2$;

(iii) $x_1 < x_2 < x_0$.

Provemos primeiro, para cada caso, que, se s é crescente, então f é convexa. Como s é crescente, temos que

$$s(x_1) < s(x_2),$$

ou seja,

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} < \frac{f(x_2) - f(x_0)}{x_2 - x_0}. \tag{1.6}$$

(i): Para $\alpha \in (0, 1)$, temos que $x_1 = \alpha x_2 + (1 - \alpha)x_0$. Substituindo em (1.6), obtemos:

$$\frac{f(x_1) - f(x_0)}{\alpha x_2 + (1 - \alpha)x_0 - x_0} < \frac{f(x_2) - f(x_0)}{x_2 - x_0},$$

$$\frac{f(x_1) - f(x_0)}{\alpha(x_2 - x_0)} < \frac{f(x_2) - f(x_0)}{x_2 - x_0}.$$

Como $x_2 - x_0 > 0$, temos

$$f(x_1) - f(x_0) < \alpha f(x_2) - \alpha f(x_0),$$

$$f(x_1) < \alpha f(x_2) + (1 - \alpha)f(x_0),$$

ou seja, f é convexa para todo $x_0 < x_1$.

(ii): Para $\alpha \in (0, 1)$, temos que $x_0 = \alpha x_2 + (1 - \alpha)x_1$. Substituindo em (1.6), obtemos:

$$\frac{f(x_1) - f(x_0)}{x_1 - (\alpha x_2 + (1 - \alpha)x_1)} < \frac{f(x_2) - f(x_0)}{x_2 - (\alpha x_2 + (1 - \alpha)x_1)},$$

$$\frac{f(x_1) - f(x_0)}{\alpha(x_1 - x_2)} < \frac{f(x_2) - f(x_0)}{(1 - \alpha)(x_2 - x_1)}.$$

Como $\alpha > 0$, $(1 - \alpha) > 0$ e $x_1 - x_2 < 0$, temos

$$-(1 - \alpha)f(x_1) + (1 - \alpha)f(x_0) < \alpha f(x_2) - \alpha f(x_0),$$

$$(1 - \alpha)f(x_0) + \alpha f(x_0) < \alpha f(x_2) + (1 - \alpha)f(x_1).$$

Portanto,

$$f(x_0) < \alpha f(x_2) + (1 - \alpha)f(x_1),$$

ou seja, f é convexa para todo $x_0 \in (x_1, x_2)$.

(iii): Para $\alpha \in (0, 1)$, temos que $x_2 = \alpha x_0 + (1 - \alpha)x_1$. Substituindo em (1.6), obtemos:

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} < \frac{f(x_2) - f(x_0)}{\alpha x_0 + (1 - \alpha)x_1 - x_0},$$

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} < \frac{f(x_2) - f(x_0)}{(1 - \alpha)(x_1 - x_0)}.$$

Como $x_1 - x_0 < 0$, temos

$$f(x_2) - f(x_0) < (1 - \alpha f(x_1) - (1 - \alpha)f(x_0),$$

$$f(x_2) < \alpha f(x_0) + (1 - \alpha)f(x_1),$$

ou seja, f é convexa para todo $x_0 > x_2$.

Agora provemos para cada caso que, se f é convexa, então s é crescente.

(i): Para $\alpha \in (0,1)$, temos que

$$x_1 = \alpha x_2 + (1-\alpha)x_0. \tag{1.7}$$

Temos que f é convexa, então, por definição, vale a desigualdade:

$$f(x_1) < \alpha f(x_2) + (1-\alpha)f(x_0),$$

$$f(x_1) - f(x_0) < \alpha[f(x_2) - f(x_0)].$$

Como $\alpha > 0$ e $x_2 - x_0 > 0$, temos

$$\frac{f(x_1) - f(x_0)}{\alpha(x_2 - x_0)} < \frac{f(x_2) - f(x_0)}{x_2 - x_0}.$$

De (1.7), temos que $\alpha x_2 = x_1 - (1-\alpha)x_0$, logo

$$\frac{f(x_1) - f(x_0)}{x_1 - (1-\alpha)x_0 - \alpha x_0} < \frac{f(x_2) - f(x_0)}{x_2 - x_0},$$

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} < \frac{f(x_2) - f(x_0)}{x_2 - x_0},$$

ou seja, $s(x)$ é crescente para $x_0 < x_1$.

(ii): Para $\alpha \in (0,1)$, temos que

$$x_0 = \alpha x_2 + (1-\alpha)x_1. \tag{1.8}$$

Sabemos que f é convexa, então, por definição, vale a desigualdade:

$$f(x_0) < \alpha f(x_2) + (1-\alpha)f(x_1).$$

Somando $0 = \alpha f(x_0) - \alpha f(x_0)$ no primeiro membro da desigualdade acima, obtemos

$$\alpha f(x_0) + (1-\alpha)f(x_0) < \alpha f(x_2) + (1-\alpha)f(x_1),$$

$$-(1-\alpha)f(x_1) + (1-\alpha)f(x_0) < \alpha f(x_2) - \alpha f(x_0).$$

Como $x_2 - x_1 > 0$, temos

$$\frac{(1-\alpha)[f(x_1) - f(x_0)]}{-(x_2 - x_1)} < \frac{\alpha[f(x_2) - f(x_0)]}{x_2 - x_1}.$$

Sabemos que $1 - \alpha > 0$ e $\alpha > 0$, então

$$\frac{f(x_1) - f(x_0)}{\alpha(x_1 - x_2)} < \frac{f(x_2) - f(x_0)}{(1-\alpha)(x_2 - x_1)},$$

$$\frac{f(x_1) - f(x_0)}{\alpha x_1 - \alpha x_2} < \frac{f(x_2) - f(x_0)}{x_2 - \alpha x_2 - (1-\alpha)x_1}.$$

De (1.8), temos que $\alpha x_2 = x_0 - (1-\alpha)x_1$, logo

$$\frac{f(x_1)-f(x_0)}{\alpha x_1 - x_0 + (1-\alpha)x_1} < \frac{f(x_2)-f(x_0)}{x_2 - x_0 + (1-\alpha)x_1 - (1-\alpha)x_1},$$

$$\frac{f(x_1)-f(x_0)}{x_1 - x_0} < \frac{f(x_2)-f(x_0)}{x_2 - x_0},$$

ou seja, $s(x)$ é crescente para $x_0 \in (x_1, x_2)$.

(iii): Para $\alpha \in (0,1)$, temos que

$$x_2 = \alpha x_0 + (1-\alpha)x_1. \tag{1.9}$$

Sabemos que f é convexa, então, por definição, vale a desigualdade:

$$f(x_2) < \alpha f(x_0) + (1-\alpha)f(x_1).$$

Somando $0 = f(x_0) - f(x_0)$ no segundo membro da desigualdade acima, obtemos

$$f(x_2) - f(x_0) < (1-\alpha)f(x_1) - (1-\alpha)f(x_0).$$

Como $1-\alpha > 0$ e $x_1 - x_0 < 0$, então

$$\begin{aligned} \frac{f(x_1)-f(x_0)}{x_1 - x_0} &< \frac{f(x_2)-f(x_0)}{(1-\alpha)(x_1 - x_0)} \\ &= \frac{f(x_2)-f(x_0)}{(1-\alpha)x_1 - x_0 + \alpha x_0}. \end{aligned}$$

De (1.9), temos que $\alpha x_0 = x_2 - (1-\alpha)x_1$, então

$$\begin{aligned} \frac{f(x_1)-f(x_0)}{x_1 - x_0} &< \frac{f(x_2)-f(x_0)}{(1-\alpha)x_1 - x_0 + x_2 - (1-\alpha)x_1} \\ &= \frac{f(x_2)-f(x_0)}{x_2 - x_0}, \end{aligned}$$

ou seja, $s(x)$ é crescente para $x_0 > x_2$.

□

Sabemos que todo $P_u = (u, f(u)) \in \operatorname{gr} f$ está abaixo da linha $P_{x_0}P_{x_1}$ quando $u \in (x_0, x_1)$. Mas o que acontece fora do intervalo? A Proposição 1.5 implica que, para $w \notin [x_0, x_1]$, P_w está acima da linha $P_{x_0}P_{x_1}$. Para ver isso, troque u e x_1 na Figura 1.4.

Exemplo 1 *Se φ é uma função crescente em um segmento $[a,b]$, então a convexidade da função*

$$[a,b] \ni x \longmapsto f(x) := \int_a^x \varphi(u)du$$

é facilmente estabelecida da Definição 1.2.

De fato, tomemos

- $\alpha \in (0,1)$,

- $a \leq x < x' \leq b$,
- $x'' := \alpha x + (1-\alpha)x'$,

e chamemos

$$\Phi := f(x'') - \alpha f(x) - (1-\alpha)f(x'),$$

que deve ser não-positiva ($\Phi \leq 0$), para que tenhamos

$$f(x'') \leq \alpha f(x) + (1-\alpha)f(x'),$$

ou seja, queremos que f seja convexa.

Temos

$$\Phi = f(x'') - \alpha f(x) - (1-\alpha)f(x').$$

Substituimos $f(x)$ por $\int_a^x \varphi(u)du$:

$$\begin{aligned}
\Phi &= \int_a^{x''} \varphi(u)du - \alpha \int_a^x \varphi(u)du - (1-\alpha)\int_a^{x'} \varphi(u)du \\
&= \int_a^{x''} \varphi(u)du - \alpha \int_a^x \varphi(u)du - \int_a^{x'} \varphi(u)du + \alpha \int_a^{x'} \varphi(u)du \\
&= \int_{x'}^{x''} \varphi(u)du + \alpha \int_x^{x'} \varphi(u)du.
\end{aligned}$$

Como $x \leq x'' \leq x'$, temos

$$\begin{aligned}
\Phi &= \int_{x'}^{x''} \varphi(u)du + \alpha \left[\int_x^{x''} \varphi(u)du + \int_{x''}^{x'} \varphi(u)du\right] \\
&= \int_{x'}^{x''} \varphi(u)du + \alpha \int_x^{x''} \varphi(u)du + \alpha \int_{x''}^{x'} \varphi(u)du \\
&= \alpha \int_x^{x''} \varphi(u)du + \int_{x'}^{x''} \varphi(u)du + \alpha \int_{x''}^{x'} \varphi(u)du \\
&= \alpha \int_x^{x''} \varphi(u)du + (\alpha - 1)\int_{x''}^{x'} \varphi(u)du.
\end{aligned}$$

Temos que $a \leq x < x'' < x' \leq b$. Como φ é crescente em $[a,b]$, se $u \in (x, x'')$, então $\varphi(u) \leq \varphi(x'')$. Logo, vale

$$\begin{aligned}
\Phi &= \alpha \int_x^{x''} \varphi(u)du + (\alpha - 1)\int_{x''}^{x'} \varphi(u)du \\
&\leq \alpha \int_x^{x''} \varphi(x'')du + (\alpha - 1)\int_{x''}^{x'} \varphi(x'')du.
\end{aligned}$$

Como $\varphi(x'')$ é uma constante, então

$$\Phi \leq \alpha\varphi(x'')(x''-x) + (\alpha-1)\varphi(x'')(x'-x'').$$

Trocando o segundo e o quarto x'' da desigualdade acima por $\alpha x+(1-\alpha)x'$, obtemos

$$\begin{aligned}\Phi &\leq \alpha\varphi(x'')(\alpha x+(1-\alpha)x'-x)+(\alpha-1)\varphi(x'')(x'-(\alpha x+(1-\alpha)x'))\\ &= \alpha\varphi(x'')(\alpha x+x'-\alpha x'-x)+(\alpha-1)\varphi(x'')(x'-\alpha x-x'+\alpha x')\\ &= \alpha\varphi(x'')(\alpha-1)(x-x')-\alpha(\alpha-1)\varphi(x'')(x-x')\\ &= 0.\end{aligned}$$

Logo,

$$f(x'')-\alpha f(x)-(1-\alpha)f(x')\leq 0,$$
$$f(x'')\leq \alpha f(x)+(1-\alpha)f(x'),$$

ou seja, f é convexa.

Exemplo 2 *A função $f(x)=|x|$ é convexa em $\mathbb{R}$.*

Com efeito, tomemos

- $\alpha\in(0,1)$,
- $a\leq x<x'\leq b$,
- $x'':=\alpha x+(1-\alpha)x'$,

e chamemos

$$\Phi := f(x'')-\alpha f(x)-(1-\alpha)f(x'),$$

que deve ser não-positiva ($\Phi\leq 0$), para que tenhamos

$$f(x'')\leq \alpha f(x)+(1-\alpha)f(x'),$$

ou seja, queremos que f seja convexa.

Temos que

$$\Phi=|x''|-\alpha|x|-(1-\alpha)|x'|.$$

Substituindo x'' por $\alpha x+(1-\alpha)x'$, temos

$$\Phi=|\alpha x+(1-\alpha)x'|-\alpha|x|-(1-\alpha)|x'|.$$

Aplicando a desigualdade triangular, obtemos

$$\begin{aligned}\Phi &\leq |\alpha x|+|(1-\alpha)x'|-\alpha|x|-(1-\alpha)|x'|\\ &= \alpha|x|+(1-\alpha)|x'|-\alpha|x|-(1-\alpha)|x'|\\ &= 0.\end{aligned}$$

Como $\Phi\leq 0$, concluimos que f é convexa.

Teorema 1.7 *Seja f definida em $(0,+\infty)$. Então a função*

$$0<x\mapsto g(x):=xf\left(\frac{1}{x}\right)$$

é convexa em $(0,+\infty)$ se e somente se f também é convexa em $(0,+\infty)$.

Prova. Suponhamos que f é convexa em $(0,+\infty)$; seja $x_0 > 0$ e consideremos a função-inclinação

$$s_g(x) := \frac{g(x)-g(x_0)}{x-x_0} = \frac{xf\left(\dfrac{1}{x}\right) - x_0 f\left(\dfrac{1}{x_0}\right)}{x-x_0},$$

definida em $(0,+\infty)\backslash\{x_0\}$. Temos

$$\begin{aligned}
s_g(x) &= \frac{xf\left(\dfrac{1}{x}\right) - x_0 f\left(\dfrac{1}{x_0}\right) + xf\left(\dfrac{1}{x_0}\right) - xf\left(\dfrac{1}{x_0}\right)}{x-x_0} \\
&= \frac{x-x_0}{x-x_0} f\left(\frac{1}{x_0}\right) + \frac{x}{x-x_0}\left[f\left(\frac{1}{x}\right) - f\left(\frac{1}{x_0}\right)\right] \\
&= f\left(\frac{1}{x_0}\right) + \frac{\dfrac{x}{x}}{\dfrac{x-x_0}{x}}\left[f\left(\frac{1}{x}\right) - f\left(\frac{1}{x_0}\right)\right] \\
&= f\left(\frac{1}{x_0}\right) + \frac{1}{1-\dfrac{x_0}{x}}\left[f\left(\frac{1}{x}\right) - f\left(\frac{1}{x_0}\right)\right] \\
&= f\left(\frac{1}{x_0}\right) - \frac{1}{x_0\left(\dfrac{1}{x} - \dfrac{1}{x_0}\right)}\left[f\left(\frac{1}{x}\right) - f\left(\frac{1}{x_0}\right)\right] \\
&= f\left(\frac{1}{x_0}\right) - \frac{1}{x_0}\,\frac{f\left(\dfrac{1}{x}\right) - f\left(\dfrac{1}{x_0}\right)}{\dfrac{1}{x} - \dfrac{1}{x_0}} \\
&= f\left(\frac{1}{x_0}\right) - \frac{1}{x_0}\, s_f\left(\frac{1}{x}\right).
\end{aligned}$$

Quando x cresce, $\dfrac{1}{x}$ decresce; então $s_f\left(\dfrac{1}{x}\right)$ decresce, pela Proposição 1.6 e, portanto, s_g cresce, logo g é convexa.

Analogamente, podemos mostrar que, se g é convexa em $(0,+\infty)$, então g também é convexa em $(0,+\infty)$. Como

$$g(x) = xf\left(\frac{1}{x}\right),$$

temos

$$g\left(\frac{1}{x}\right) = \frac{1}{x}\, f(x).$$

Logo,

$$f(x) = xg\left(\frac{1}{x}\right).$$

Então as demais linhas desta demonstração seguem as mesmas da demonstração anterior. $\square$

A seguir, apresentamos dois exemplos em que aplicamos o Teorema 1.7.

Exemplo 3 *Sejam as funções*

$$f(x) = -\log x,$$

convexa em $(0, +\infty)$ *e*

$$g(x) := xf\left(\frac{1}{x}\right).$$

Então

$$\begin{aligned} g(x) &= x\left(-\log\left(\frac{1}{x}\right)\right) \\ &= -x\log\left(\frac{1}{x}\right) \\ &= x\log\left(\frac{1}{x}\right)^{-1} \\ &= x\log x. \end{aligned}$$

Logo, pelo Teorema 1.7, a função

$$g(x) = x\log x$$

também é convexa em $(0, +\infty)$*.*

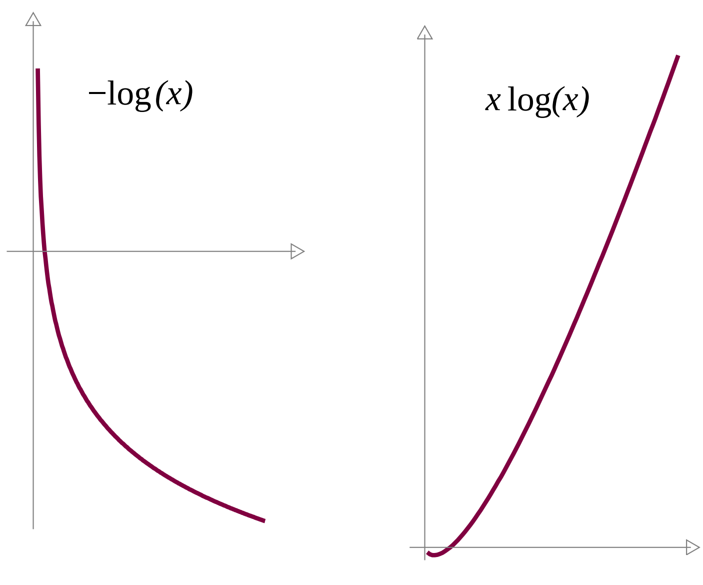

Figura 1.5: A convexidade das funções $-\log(x)$ e $x\log(x)$.

Exemplo 4 *Sejam as funções*

$$f(x) = \exp x$$

convexa em $(0, +\infty)$ *e*

$$g(x) := xf\left(\frac{1}{x}\right).$$

Então

$$g(x) = x\exp\left(\frac{1}{x}\right)$$

também é convexa em $(0, +\infty)$*, pelo Teorema 1.7.*

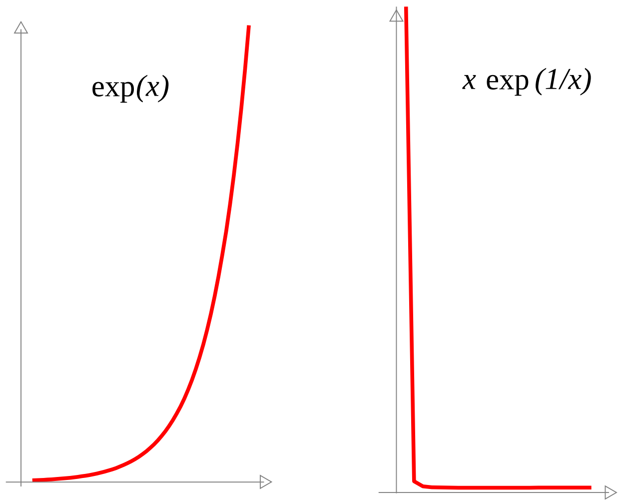

Figura 1.6: A convexidade das funções $\exp(x)$ e $x\exp(1/x)$.

1.1.2 Desigualdades com mais que dois pontos

Uma característica essencial da desigualdade básica (1.2) é que ela pode ser generalizada para mais que dois pontos. Nesta subseção, além de [3], temos [4] como referência.

Teorema 1.8 *Sejam* I *um intervalo não vazio de* $\mathbb{R}$ *e* f *uma função convexa em* I*. Então, para qualquer coleção* $\{x_1, \ldots, x_k\}$ *de pontos em* I *e qualquer coleção de números* $\{\alpha_1, \ldots, \alpha_k\}$ *satisfazendo*

$$\alpha_i \geq 0 \ \text{ para } \ i = 1, \ldots, k \ \text{ e } \ \sum_{i=1}^{k} \alpha_i = 1, \tag{1.10}$$

vale a desigualdade de Jensen:

$$f\left(\sum_{i=1}^{k} \alpha_i x_i\right) \leq \sum_{i=1}^{k} \alpha_i f(x_i).$$

Prova. Considere primeiro $k = 2$. A relação é trivial se α_1 ou α_2 é zero; se não, temos justamente (1.2).

Agora, suponha por indução que a relação é verdadeira para $k-1$; sejam as coleções $\{x_i\}$ e $\{\alpha_i\}$ como em (1.10). Se α_k é 0 ou 1, não há o que provar. Se não, fixemos

$$\bar{\alpha} := \sum_{i=1}^{k-1} \alpha_i \in (0,1).$$

Então, temos

$$\sum_{i=1}^{k} \alpha_i = 1,$$

$$\alpha_k + \sum_{i=1}^{k-1} \alpha_i = 1,$$

$$\bar{\alpha} + \alpha_k = 1.$$

Portanto,

$$\alpha_k = 1 - \bar{\alpha},$$

onde $\alpha_k \in (0,1)$. Seja $\bar{\alpha}_i := \dfrac{\alpha_i}{\bar{\alpha}}$, para $i = 1, ..., k-1$. Temos que

$$\bar{\alpha}_i \geq 0 \quad \text{e} \quad \sum_{i=1}^{k-1} \bar{\alpha}_i = 1. \tag{1.11}$$

Então

$$\sum_{i=1}^{k} \alpha_i x_i = \bar{\alpha} \sum_{i=1}^{k-1} \bar{\alpha}_i x_i + (1-\bar{\alpha}) x_k.$$

Nesta relação, o ponto $\bar{x} := \sum_{i=1}^{k-1} \bar{\alpha}_i x_i$ está em I, isto é, entre o menor x_i e o maior x_i. Aplicando (1.11), obtemos:

$$\begin{aligned} f(\sum_{i=1}^{k} \alpha_i x_i) &\leq \bar{\alpha} f(\bar{x}) + (1-\bar{\alpha}) f(x_k) \\ &= \bar{\alpha} f(\bar{x}) + \alpha_k f(x_k). \end{aligned}$$

Então o resultado segue da suposição de indução aplicada a $\bar{x}$:

$$\begin{aligned} \bar{\alpha} f(\bar{x}) &\leq \bar{\alpha} \sum_{i=1}^{k-1} \bar{\alpha}_i f(x_i) \\ &= \sum_{i=1}^{k-1} \alpha_i f(x_i). \end{aligned}$$

□

O conjunto descrito por (1.10) é chamado de *simplex unitário* de $\mathbb{R}^k$. Uma coleção de α_i's satisfazendo (1.10) é chamado de *multiplicadores convexos* e o correspondente $x = \sum_{i=1}^{k} \alpha_i x_i$ é uma *combinação convexa* dos x_i's.

Dizemos que a maioria das desigualdades úteis entre números reais são conseqüências da desigualdade de Jensen acima, mesmo que nem sempre seja fácil descobrir a função convexa latente. A seguir, alguns exemplos típicos.

Exemplo 5 *Suponhamos que sabemos que a função* $-\log x$ *é convexa em* $(0,+\infty)$*. Então, vale a desigualdade de Jensen:*

$$\begin{aligned}-\log\left(\sum_{i=1}^k \alpha_i x_i\right) &\leq -\sum_{i=1}^k \alpha_i \log x_i\\ &= -\sum_{i=1}^k \log x_i^{\alpha_i}\\ &= -\log\left(\prod_{i=1}^k x_i^{\alpha_i}\right),\end{aligned}$$

para todos $x_i > 0$ *e* $\alpha = (\alpha_1, \ldots, \alpha_k)$ *no simplex unitário. Através da monotonicidade da função exponencial, obtemos*

$$\prod_{i=1}^k x_i^{\alpha_i} \leq \sum_{i=1}^k \alpha_i x_i.$$

Observação 1.9 *No caso particular em que* $\alpha_i = \dfrac{1}{n}$, $i = 1, \ldots, n$*, temos a clássica desigualdade entre a média aritmética e a média geométrica:*

$$\sqrt[n]{x_1 x_2 \ldots x_n} \leq \frac{x_1 + \cdots + x_n}{n}.$$

Exemplo 6 *Seja a função* $x \log x$*, convexa em* $(0,+\infty)$*. Aplicando a desigualdade de Jensen, temos*

$$\left(\sum_{i=1}^k \alpha_i x_i\right)\log\left(\sum_{i=1}^k \alpha_i x_i\right) \leq \sum_{i=1}^k \alpha_i (x_i \log x_i),$$

$$\begin{aligned}\log\left(\sum_{i=1}^k \alpha_i x_i\right)^{\sum_{i=1}^k \alpha_i x_i} &\leq \sum_{i=1}^k \log x_i^{\alpha_i x_i}\\ &= \log\left(\prod_{i=1}^k x_i^{\alpha_i x_i}\right).\end{aligned}$$

Como a função $\log x$ *é crescente, temos*

$$\left(\sum_{i=1}^k \alpha_i x_i\right)^{\sum_{i=1}^k \alpha_i x_i} \leq \prod_{i=1}^k x_i^{\alpha_i x_i}.$$

1.1.3 Definição moderna de convexidade

Quando trabalhamos com convexidade, é conveniente considerar a função f como sendo definida *em todo o espaço* $\mathbb{R}$, levando-se em conta o valor $+\infty$ para $f(x)$. Até agora, convexidade envolvia um par (I, f), onde I era um intervalo não vazio e f uma função de I em $\mathbb{R}$, satisfazendo (1.2) em I. Podemos estender igualmente uma função f para além de I através da função

$$f_e(x) := \begin{cases} f(x) & \text{para} \quad x \in I, \\ +\infty & \text{para} \quad x \notin I. \end{cases}$$

A *função-estendida* f_e leva $\mathbb{R}$ ao conjunto $\mathbb{R} \cup \{+\infty\}$; é claro, o valor $+\infty$ já foi cuidadosamente escolhido: é a única maneira de preservar a relação da definição (1.2) fora de I. Daqui em diante e sem menção explícita, todas as (potenciais) funções convexas serão funções-estendidas: o "e" subscrito será, portanto, omitido e as definições de §1.1.1 serão conseqüentemente substituídas como segue:

Definição 1.10 *A função $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$, não identicamente igual a $+\infty$, é dita convexa quando a desigualdade em $\mathbb{R} \cup \{+\infty\}$*

$$f(\alpha x + (1-\alpha)x') \leq \alpha f(x) + (1-\alpha) f(x') \tag{1.12}$$

vale para todos os pares de pontos (x, x') em $\mathbb{R}$ e todos $\alpha \in (0, 1)$.

Equivalentemente, f é uma função cujo epigrafo é um conjunto não vazio em $\mathbb{R} \times \mathbb{R}$.

O conjunto de tais funções é denotado por $\operatorname{Conv} \mathbb{R}$.

Definição 1.11 *O domínio de $f \in \operatorname{Conv} \mathbb{R}$ é o conjunto não vazio*

$$\operatorname{Dom} f := \{x \in \mathbb{R} \mid f(x) \in \mathbb{R}\}.$$

A utilidade da Definição (1.10) é mais que notacional; ela é conveniente especialmente quando otimização está envolvida. A seguir, daremos três exemplos para ilustrar isso.

- Seja x um parâmetro real e considere o problema de otimização simples:

  $$\inf\{-y \mid y^2 \leq x\}. \tag{1.13}$$

  Isto não tem sentido para $x < 0$ mas, para $x \geq 0$, o valor ótimo é $-\sqrt{x}$, uma função convexa de x. Em vista da convenção $\inf \emptyset = +\infty$, temos uma função convexa no sentido da Definição (1.10). Isso é bom para saber que problemas do tipo (1.13) fornecem funções convexas de x, mesmo que eles possam não ter sentido para todos os valores de x.

- Associada a uma dada função f, temos a chamada *função conjugada*, dada por

  $$x \mapsto \sup\{xy - f(y) \mid y \in \mathbb{R}\} \ \text{ com } \ x \in \mathbb{R}$$

  Aqui, mais uma vez os valores de x para os quais o supremo é finito não são necessariamente conhecidos antecipadamente. Este supremo é, deste modo, uma função estendida de x, uma função que se torna de extrema importância.

- Suponha que uma função g, convexa em I, precisa ser minimizada em algum sub-intervalo $C \in \mathbb{R}$. A restrição $x \in C$ pode ser incluída na função objetiva pela definição:
$$f(x) := \begin{cases} g(x) & \text{se} \quad x \in C, \\ +\infty & \text{caso contrário.} \end{cases}$$
A função f resultante pertence a $\operatorname{Conv} \mathbb{R}$ e minimizá-la (em todo o $\mathbb{R}$) é equivalente ao problema original.

1.2 Primeiras propriedades

1.2.1 Estabilidade sob operações funcionais

Nesta seção, listamos algumas das operações que podem ser provadas para preservar convexidade, simplesmente em vista das próprias definições.

Proposição 1.12 *Sejam $f_1, \ldots, f_m$ m funções convexas em $\mathbb{R}$ e $t_1, \ldots, t_m$ números positivos. Se existe $x_0 \in \mathbb{R}$ tal que $f_j(x_0) < +\infty$, para todo $j = 1, \ldots, m$, então a função*
$$f := \sum_{j=1}^{m} t_j f_j$$
pertence a $\operatorname{Conv} \mathbb{R}$.

Prova. Para f_j $(j = 1, \ldots, m)$, temos
$$f_j(\alpha x + (1-\alpha)x') \leq \alpha f_j(x) + (1-\alpha) f_j(x').$$
Como t_j $(j = 1, \ldots, m)$ é positivo, a desigualdade se mantém ao multiplicarmos ambos os membros da desigualdade por t_j:
$$t_j f_j(\alpha x + (1-\alpha)x') \leq \alpha t_j f_j(x) + (1-\alpha) t_j f_j(x').$$
Somando as desigualdades
$$t_1 f_1(\alpha x + (1-\alpha)x') \leq \alpha t_1 f_1(x) + (1-\alpha) t_1 f_1(x'),$$
$$\vdots$$
$$t_m f_m(\alpha x + (1-\alpha)x') \leq \alpha t_m f_m(x) + (1-\alpha) t_m f_m(x'),$$
obtemos
$$\sum_{j=1}^{m} t_j f_j(\alpha x + (1-\alpha)x') \leq \alpha \sum_{j=1}^{m} t_j f_j(x) + (1-\alpha) \sum_{j=1}^{m} t_j f_j(x'),$$
ou seja, $f := \sum_{j=1}^{m} t_j f_j \in \operatorname{Conv} \mathbb{R}$. $\square$

Proposição 1.13 *Se $f : I \to \mathbb{R}$ e $g : J \to \mathbb{R}$ são funções convexas, com $f(I) \subset J$, e g monótona não-decrescente, então $g \circ f$ é convexa.*

Prova. Dados $x, x' \in I$ e $\alpha \in [0,1]$, temos que

$$f((1-\alpha)x + \alpha x') \leq (1-\alpha)f(x) + \alpha f(x'),$$

pois f é convexa, por hipótese. Temos que $f(I) \subset J$. Também temos que g é não decrescente, logo

$$g(f((1-\alpha)x + \alpha x')) \leq g((1-\alpha)f(x) + \alpha f(x')).$$

Sabemos que g é convexa, então

$$g((1-\alpha)f(x) + \alpha f(x')) \leq (1-\alpha)g(f(x)) + \alpha g(f(x')),$$

ou seja,

$$g \circ f((1-\alpha)x + \alpha x') \leq (1-\alpha)g \circ f(x) + \alpha g \circ f(x'),$$

o que significa que $g \circ f$ é convexa. $\square$

1.3 Propriedades de continuidade

Ao longo desta seção, utilizamos conceitos de análise que podem ser verificados em [5].

1.3.1 Continuidade no interior do domínio

Funções convexas se tornam úteis para aproveitar propriedades nítidas de continuidade. Desenhos simples sugerem que uma função convexa pode ter descontinuidades em pontos do final de seu intervalo de definição do $\operatorname{Dom} f$, mas tem um comportamento contínuo no seu interior. Isto é feito precisamente no resultado seguinte, conforme [3].

Teorema 1.14 *Seja uma função convexa f definida em (a,b). Então f é contínua.*

Prova. Dado $x \in (a,b)$, escolha $\delta > 0$ tal que $[x-\delta, x+\delta] \subset (a,b)$. Agora considere um ponto $z_1 \in (x-\delta, x)$. Aplicando a propriedade das inclinações crescentes para os pontos $x-\delta, z_1, x$, temos

$$\frac{f(z_1) - f(x-\delta)}{z_1 - (x-\delta)} \leq \frac{f(x) - f(x-\delta)}{\delta} \leq \frac{f(x) - f(z_1)}{x - z_1}.$$

Agora, aplicando a propriedades das inclinações crescentes para os pontos $z_1, x, x+\delta$, temos

$$\frac{f(x) - f(z_1)}{x - z_1} \leq \frac{f(x+\delta) - f(z_1)}{(x+\delta) - z_1} \leq \frac{f(x+\delta) - f(x)}{\delta}.$$

Logo,

$$\frac{f(x) - f(x-\delta)}{\delta} \leq \frac{f(x) - f(z_1)}{x - z_1} \leq \frac{f(x+\delta) - f(x)}{\delta}.$$

Seja agora um ponto $z_2 \in (x, x+\delta)$. Aplicando a propriedade das inclinações crescentes para os pontos $x-\delta, x, z_2$, temos

$$\frac{f(x)-f(x-\delta)}{\delta} \leq \frac{f(z_2)-f(x-\delta)}{z_2-(x-\delta)} \leq \frac{f(z_2)-f(x)}{z_2-x}.$$

Agora, aplicando a propriedade das inclinações crescentes para os pontos $x, z_2, x+\delta$, temos

$$\frac{f(z_2)-f(x)}{z_2-x} \leq \frac{f(x+\delta)-f(x)}{\delta} \leq \frac{f(x+\delta)-f(z_2)}{(x+\delta)-z_2}.$$

Logo,

$$\frac{f(x)-f(x-\delta)}{\delta} \leq \frac{f(z_2)-f(x)}{z_2-x} \leq \frac{f(x+\delta)-f(x)}{\delta}.$$

Então podemos concluir que, se $z \in [x-\delta, x+\delta]$,

$$\frac{f(x)-f(x-\delta)}{\delta} \leq \left|\frac{f(x)-f(z)}{x-z}\right| \leq \frac{f(x+\delta)-f(x)}{\delta}.$$

Seja

$$L = \frac{f(x+\delta)-f(x)}{\delta}.$$

Portanto, temos que

$$|f(x)-f(z)| \leq L|x-z|,$$

para qualquer $z \in [x-\delta, x+\delta]$. Logo, podemos passar o limite quando $z \to x$, obtendo

$$\lim_{z\to x} |f(x)-f(z)| = 0,$$

o que prova que f é contínua.

Observação 1.15 *Mais adiante, mostraremos que f é localmente Lipschitziana. Precisamos primeiro falar sobre as derivadas laterais de f, que serão de fundamental importância para encontrarmos a constante de Lipschitz.*

Proposição 1.16 *Seja o domínio de f ($f \in \operatorname{Conv} \mathbb{R}$) com interior não vazio e chamemos de $a \in \mathbb{R}$ seu extremo à esquerda. Então o limite à direita $f(a_+) := \lim_{x\to a^+} f(x)$ existe em $\mathbb{R} \cup \{+\infty\}$, e $f(a) \geq f(a_+)$.*

Similarmente, se $b \in \mathbb{R}$ é o extremo à direita do $\operatorname{Dom} f$, o limite à esquerda $f(b_-)$ existe em $\mathbb{R} \cup \{+\infty\}$ e $f(b) \geq f(b_-)$.

Prova. Sejam x_0 pertencente ao interior do $\operatorname{Dom} f$, $d := -1$, $t_0 := x_0 - a > 0$. A função crescente

$$0 < t \mapsto \frac{f(x_0+td)-f(x_0)}{t} =: q(t)$$

tem um limite $\ell \in \mathbb{R} \cup \{+\infty\}$ para $t \to t_o^+$, em cujo caso $x_0 + td \to a^+$.

Para provar que $q(t)$ é crescente, precisamos mostrar que

$$t_1 > t_2 \Rightarrow q(t_1) > q(t_2).$$

Tomemos, então,

$$t_1 > t_2$$

e multipliquemos ambos os membros da desigualdade por d para obter

$$dt_1 < dt_2,$$

pois $d < 0$, por hipótese. Somamos x_0 em ambos os lados e obtemos

$$x_0 + dt_1 < x_0 + dt_2.$$

Temos que

$$\frac{f(x_0 + dt_1) - f(x_0)}{(x_0 + dt_1) - x_0} < \frac{f(x_0 + dt_2) - f(x_0)}{(x_0 + dt_2) - x_0},$$

pois funções inclinação são crescentes. Multiplicando ambos os membros da desigualdade por d, chegamos em

$$\frac{f(x_0 + dt_1) - f(x_0)}{t_1} > \frac{f(x_0 + dt_2) - f(x_0)}{t_2},$$

logo

$$q(t_1) > q(t_2),$$

portanto $q(t)$ é função crescente. A partir de

$$\frac{f(x_0 + td) - f(x_0)}{t} = q(t),$$

temos

$$f(x_0 + td) = f(x_0) + tq(t) \to f(x_0) + (x_0 - a)\ell =: f(a_+) \in \mathbb{R} \cup \{+\infty\}.$$

Então coloquemos $t \to t_0^-$ na relação

$$q(t) \leq q(t_0) = \frac{f(a) - f(x_0)}{x_0 - a},$$

para todo $t \in]0, t_0[$, para obter

$$\ell = \frac{f(a_+) - f(x_0)}{x_0 - a} \leq \frac{f(a) - f(x_0)}{x_0 - a},$$

logo $f(a_+) \leq f(a)$.

Analogamente podemos mostrar que $f(b) \geq f(b_-)$. □

1.3.2 Semi-continuidade inferior: funções convexas fechadas

De acordo com a Proposição 1.16, o comportamento de uma função convexa nos pontos extremos de seu domínio assemelha-se a um dos casos ilustrados na Figura 1.7. Vemos que o caso (2) é um tanto “anormal”; ele está fora da definição seguinte, que é importante para a existência de soluções em problemas de minimização. Pode-se ler mais a respeito em [3]. Também temos [1] como referência e utilizamos ferramentas de análise que podem ser encontradas em [6].

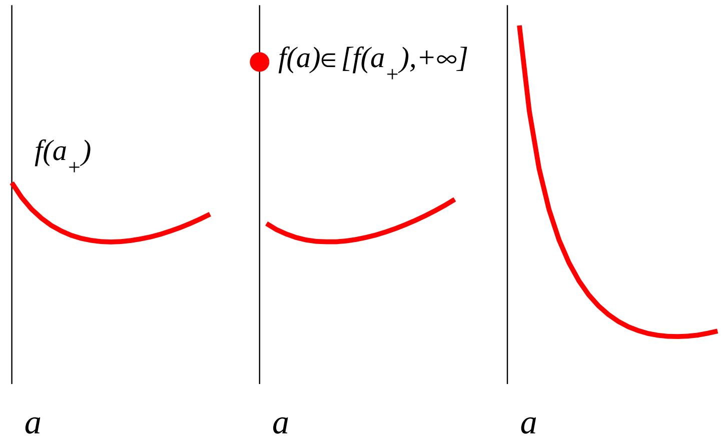

Figura 1.7: Propriedades de continuidade de funções convexas de uma variável

Definição 1.17 *Dizemos que $f \in \operatorname{Conv} \mathbb{R}$ é fechada, ou semi-contínua inferior, se*

$$\liminf_{x \to x_0} f(x) \geq f(x_0), \quad \text{para todo} \;\; x_0 \in \mathbb{R}. \tag{1.14}$$

O conjunto de funções convexas fechadas sobre $\mathbb{R}$ é denotado por $\overline{\operatorname{Conv}}\,\mathbb{R}$.

Teorema 1.18 *Seja $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$. As seguintes proposições são equivalentes:*

(i) f é semi-contínua inferior em $\mathbb{R}$;

(ii) $\operatorname{epi} f$ é um conjunto fechado em $\mathbb{R} \times \mathbb{R}$;

(iii) os conjuntos de nível $L_t(f) = \{x \in \mathbb{R} \mid f(x) \leq t, \; t \in \mathbb{R}\}$ são fechados para todo $t \in \mathbb{R}$.

Prova.

(i)→(ii): Seja $(x_k, t_k) \in \operatorname{epi} f$ uma seqüência convergindo para (x, t), quando $k \to \infty$. Como $f(x_k) \leq t_k$ para todo k, então

$$t = \lim_{k \to \infty} t_k \geq \liminf_{x_k \to x} f(x_k) \geq f(x),$$

ou seja, $(x, t) \in \operatorname{epi} f$. Como (x, t) é ponto de aderência de $\operatorname{epi} f$, temos que $\operatorname{epi} f$ é um conjunto fechado.

(ii)→(iii): Tomemos $x_k \in L_t(f)$ tal que $x_k \to a$. Precisamos provar que $a \in L_t(f)$. Seja a seqüência (x_k, t). Como $x_k \in L_t(f)$, temos que $f(x_k) \leq t$. Por definição, $(x_k, t) \in \operatorname{epi} f$, pois $t \geq f(x_k)$. Como $\operatorname{epi} f$ é um conjunto fechado, por hipótese, e $(x_k, t) \to (a, t)$, então $(a, t) \in \operatorname{epi} f$, ou seja, $t \geq f(a)$. Logo $L_t(f)$ é fechado para todo $t \in \mathbb{R}$.

(iii)→(i): Suponhamos agora, por contradição, que f não é semi-contínua inferior para algum x. Então existe uma seqüência $\{x_k\}$ convergindo para x tal que $f(x_k)$ converge para $r < f(x) \leq +\infty$. Consideremos $t \in (r, f(x))$. Quando k é suficientemente grande, temos $f(x_k) \leq t < f(x)$, ou seja, $L_t(f)$ não contém seu limite x. Então, $L_t(f)$ não é fechado. $\square$

Exemplo 7 *Seja f uma função convexa cujo domínio é todo o $\mathbb{R}$, e seja C um intervalo fechado não-vazio. Então a "restrição convexa" de f a C:*

$$f_C(x) := \begin{cases} f(x) & \textit{para} \quad x \in C, \\ +\infty & \textit{para} \quad x \notin C, \end{cases}$$

é fechada e convexa. Seu epigrafo é a intersecção do epi f *com a barra vertical gerada por C.*

Exemplo 8 *Seja C um intervalo não-vazio de $\mathbb{R}$. A função* indicadora *de C é*

$$I_C(x) := \begin{cases} 0 & \textit{para} \quad x \in C, \\ +\infty & \textit{para} \quad x \notin C. \end{cases}$$

I_C é uma função convexa se e somente se C é fechado (seus conjuntos de nível são vazios ou C).

Definição 1.19 *O fecho de $f \in \operatorname{Conv} \mathbb{R}$ é a função definida por*

$$\operatorname{cl} f(x) := \begin{cases} \liminf\limits_{y \to x} f(y) & \textit{para} \quad x \in \operatorname{cl} \operatorname{Dom} f, \\ +\infty & \textit{para} \quad x \notin \operatorname{cl} \operatorname{Dom} f. \end{cases}$$

Teorema 1.20 *Seja $f \in \operatorname{Conv} \mathbb{R}$. Para todo x_0 pertencente ao interior de seu domínio, f admite uma derivadas à direita e à esquerda finitas:*

$$D_- f(x_0) := \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} = \sup_{x < x_0} \frac{f(x) - f(x_0)}{x - x_0}, \tag{1.15}$$

$$D_+ f(x_0) := \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} = \sup_{x > x_0} \frac{f(x) - f(x_0)}{x - x_0}. \tag{1.16}$$

Elas satisfazem

$$D_- f(x_0) \leq D_+ f(x_0). \tag{1.17}$$

Prova. Aplicando o critério de inclinações crescentes, o quociente diferença envolvido em (1.15), (1.16) é justamente a função inclinação s. Para quaisquer dois pontos x, x' no interior do domínio de f satisfazendo $x < x_0 < x'$, $s(x)$ e $s(x')$ são números finitos satisfazendo $s(x) \leq s(x')$. Quando $x \to x_0^-$, $s(x)$ cresce e, quando $x \to x_0^+$, $s(x')$ decresce. Então, ambas convergem, como descrito pelas notações (1.15), (1.16); isto prova (1.17) ao mesmo tempo. $\square$

Teorema 1.21 *Se* $f \in \operatorname{Conv} \mathbb{R}$, *então* f *é contínua no* $\operatorname{int} \operatorname{Dom} f$. *Além disso, para cada intervalo compacto* $[a, b] \subset \operatorname{int} \operatorname{Dom} f$, *existe* $L \geq 0$ *tal que*

$$|f(x) - f(x')| \leq L|x - x'|, \tag{1.18}$$

para todos $x, x' \in [a, b]$.

Prova. Sejam $[a, b] \subset \operatorname{int} \operatorname{Dom} f$, $a < b$, e $a \leq x < x' \leq b$; se $a < x$, escrevemos (1.15) e (1.16) para pontos apropriados para obter

$$\begin{aligned} D_+f(a) \leq \frac{f(x) - f(a)}{x - a} &\leq D_-f(x) \leq D_+f(x) \\ \leq \frac{f(x') - f(x)}{x' - x} &\leq D_-f(x') \leq \cdots \leq D_-f(b). \end{aligned}$$

As desigualdades relevantes valem mesmo se $x = a$. Seja

$$L = \max\{-D_+f(a), D_-f(b)\}.$$

Então

$$|f(x) - f(x')| \leq L|x - x'|,$$

ou seja, que é a continuidade Lipschitziana de f em $[a, b]$. Em outras palavras, f é localmente Lipschitziana no interior de seu domínio. $\square$

1.4 Reconhecendo funções convexas

Podemos verificar a convexidade de uma dada função de várias maneiras. Estudaremos algumas dessas maneiras nesta seção, que é escrita conforme em [6].

1.4.1 Fórmula de Taylor

Teorema 1.22 *Considere uma função* $f : D \subset \mathbb{R} \to \mathbb{R}$ *de classe* C^2 *e a um ponto do interior de* D. *Então existe um único polinômio* p_2 *de grau* 2 *que satisfaz as condições:*

$$\begin{aligned} p_2(a) &= f(a), \\ p_2'(a) &= f'(a), \\ p_2''(a) &= f''(a). \end{aligned}$$

Este polinômio é dado por

$$p_2(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2.$$

Além disso,

$$\lim_{x \to a} \frac{f(x) - p_2(x)}{(x - a)^2} = 0,$$

isto é, a diferença entre $f(x)$ *e* $p_2(x)$ *vai para zero mais rapidamente do que* $(x - a)^2$. *Em outras palavras,*

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2 + R_2(a, x),$$

com
$$\lim_{x\to a}\frac{R_2(a,x)}{(x-a)^2}=0.$$

Lembramos que p_2 é o polinômio de Taylor de ordem 2 em torno do ponto a.

Prova. Seja
$$p_2=\alpha+\beta x+\gamma x^2 \tag{1.19}$$
um polinômio do segundo grau. Então
$$p_2'(x)=\beta+2\gamma x$$
e
$$p_2''(x)=2!\gamma.$$
Queremos que p_2 satisfaça
$$p_2''(a)=f''(a),$$
ou seja,
$$2!\gamma=f''(a).$$
Então
$$\gamma=\frac{f''(a)}{2!}.$$
Também queremos que p_2 satisfaça
$$p_2'(a)=f'(a),$$
ou seja,
$$\beta+2\gamma a=f'(a).$$
Logo, concluimos que
$$\beta=f'(a)-f''(a)a.$$
E, por fim, desejamos que p_2 satisfaça
$$p_2(a)=f(a),$$
isto é,
$$\alpha+\beta a+\gamma a^2=f(a).$$
Então encontramos
$$\alpha=f(a)-[f'(a)-f''(a)a]a-\frac{f''(a)}{2!}a.$$
Substituindo os valores encontrados para α,β,γ em (1.19), chegamos em
$$p_2(x)=f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2.$$

Agora, utilizando duas vezes a regra de L'Hôpital, temos

$$\begin{aligned}\lim_{x\to a}\frac{f(x)-p_2(x)}{(x-a)^2} &= \lim_{x\to a}\frac{f'(x)-p_2'(x)}{2(x-a)}\\ &= \lim_{x\to a}\frac{f''(x)-p_2''(x)}{2}\\ &= \frac{f''(a)-p_2''(a)}{2}\\ &= 0,\end{aligned}$$

isto é, o erro $R_2(a,x) = f(x) - p_2(x)$ vai para zero mais rapidamente do que a expressão $(x-a)^2$. □

Exemplo 9 *Seja a função* $f(x) = \exp(x)$. *O polinômio de Taylor de ordem* 1 *(aproximação linear) em torno do ponto* 0 *é*

$$p_1(x) = 1 + x.$$

O polinômio de Taylor de ordem 2 *(aproximação quadrática) em torno do ponto* 0 *é*

$$p_2(x) = 1 + x + \frac{x^2}{2}.$$

A Figura 1.8 mostra os gráficos de f, p_1 *e* p_2.

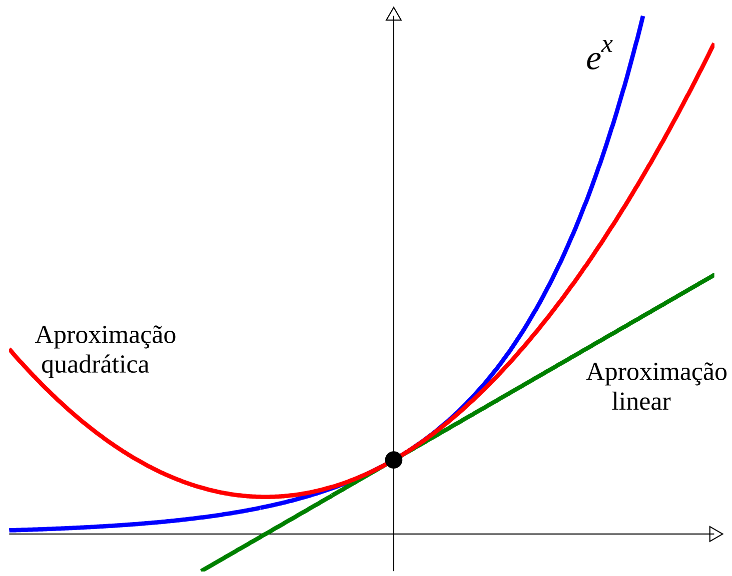

Figura 1.8: Aproximações de $\exp(x)$ pelos polinômios de Taylor de ordens 1 e 2.

Até agora, estudamos apenas propriedades de funções convexas na reta. Deste ponto em diante, passaremos a estudar algumas das propriedades de funções convexas em $\mathbb{R}^n$.

Teorema 1.23 *Considere uma função* $f : D \subset \mathbb{R}^n \to \mathbb{R}$ *de classe* C^2 *e* a *um ponto interior de* D*. Então existe um único polinômio* p_2 *de grau 2 de* n *variáveis que satisfaz as condições:*

$$p_2(a) = f(a),$$
$$\nabla p_2(a) = \nabla f(a),$$
$$\nabla^2 p_2(a) = \nabla^2 f(a).$$

Este polinômio é dado por

$$p_2(x) = f(a) + \nabla f^t(a)(x-a) + \frac{1}{2}(x-a)^t \nabla^2 f(a)(x-a).$$

Além disso,

$$f(x) = p_2(x) + R_2(a,x), \tag{1.20}$$

onde o erro $R_2(a,x)$ *satisfaz a propriedade:*

$$\lim_{x\to a} \frac{R_2(a,x)}{||x-a||^2} = \lim_{x\to a} \frac{f(x)-p_2(x)}{||x-a||^2} = 0,$$

isto é, o erro $R_2(a,x)$ *vai para zero mais rapidamente do que o quadrado da distância entre* a *e* x*. A equação* (1.20) *é a chamada "fórmula de Taylor com resto infinitesimal".*

Prova. Seja o polinômio

$$p_2(x) = \alpha + u^t(x-a) + \frac{1}{2}(x-a)^t A(x-a).$$

Fazendo $x = a$ em p_2, temos

$$p_2(a) = \alpha + u^t(a-a) + \frac{1}{2}(a-a)^t A(a-a),$$

logo,

$$p(a) = \alpha = f(a).$$

Temos que

$$\nabla p(x) = u + A(x-a),$$

então, para $x = a$,

$$\nabla p(a) = u = \nabla f(a).$$

Também temos que

$$\nabla^2 p(x) = A = \nabla^2 f(a).$$

Portanto,

$$p(x) = f(a) + \nabla f(a)(x-a) + \frac{1}{2}(x-a)^t \nabla^2 f(a)(x-a).$$

A prova de que

$$\lim_{x\to a} \frac{R_2(a,x)}{||x-a||^2} = \lim_{x\to a} \frac{f(x)-p_2(x)}{||x-a||^2} = 0,$$

resulta do lema seguinte. $\square$

Lema 1.24 *Seja $r : U \to \mathbb{R}$ uma função p vezes diferenciável no ponto $0 \in U$. Se r, juntamente com todas as suas derivadas parciais de ordem menor que, ou igual a p, se anula no ponto 0, então*

$$\lim_{v\to 0}\frac{r(v)}{|v|^p}=0.$$

Prova. Usaremos indução em p. O resultado é óbvio se $p = 0$ (interpretando "derivada parcial de ordem zero" como o valor da função). Supondo-o válido para p, seja r uma função $p+1$ vezes diferenciável no ponto 0, com todas as derivadas parciais de ordem menor que, ou igual a p+1, nulas na origem. Então, para cada $i = 1, \ldots, n$, a função $\dfrac{\partial r}{\partial x_i} : U \to \mathbb{R}$ é p vezes diferenciável e goza de propriedade semelhante, com p em vez de $p+1$. Pela hipótese de indução, temos

$$\lim_{v\to 0}\frac{\dfrac{\partial r}{\partial x_i}(v)}{|v|^p}=0.$$

Seja

$$f(t) = r(tv),$$

para $0 \le t \le 1$. Então

$$\begin{aligned} f'(t) &= \nabla f(tv)v \\ &= \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(tv)\alpha_i, \end{aligned}$$

onde $v = (\alpha_1, \ldots, \alpha_n)$. Pelo Teorema do Valor Médio, temos que existe $\theta \in [0, 1]$ tal que

$$f(1) - f(0) = f'(\theta)\,1,$$

$$r(v) - r(0) = \sum_{i=1}^{n}\frac{\partial r}{\partial x_i}(\theta v)\alpha_i.$$

Então

$$r(v) = \sum_{i=1}^{n}\frac{\partial r}{\partial x_i}(\theta v)\alpha_i$$

e, portanto,

$$\frac{r(v)}{|v|^{p+1}}=\sum_{i=1}^{n}\frac{\dfrac{\partial r}{\partial x_i}(\theta v)\alpha_i}{|v|^{p+1}}=\sum_{i=1}^{n}\frac{\dfrac{\partial r}{\partial x_i}(\theta v)}{|v|^p}\,\frac{\alpha_i}{|v|},$$

Temos que $\dfrac{\alpha_i}{|v|}$ é sempre menor que ou igual a 1, em valor absoluto. (Basta tomar em $\mathbb{R}^n$ a norma da soma.) Segue-se então que

$$\lim_{v\to 0}\frac{r(v)}{|v|^{p+1}}=0,$$

o que conclui a demonstração do lema.

$\square$

A seguir, é dada, para funções diferenciáveis, uma caracterização alternativa de convexidade.

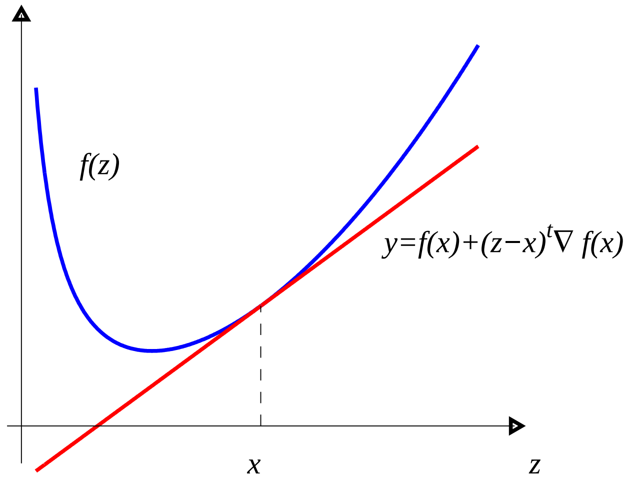

Figura 1.9: Caracterização de convexidade em termos de derivadas primeiras.

1.4.2 Caracterização de funções convexas diferenciáveis

Esta subseção foi escrita tendo [2] como base e o auxílio de [1].

Proposição 1.25 *Sejam C um subconjunto convexo de $\mathbb{R}^n$ e $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ diferenciável sobre $\mathbb{R}^n$.*

(i) f é convexa sobre C se e somente se

$$f(z) \geq f(x) + (z-x)^t \nabla f(x), \tag{1.21}$$

para todos $x, z \in C$, conforme ilustra a Figura 1.9;

(ii) f é estritamente convexa sobre C se e somente se a desigualdade acima é estrita para todo $x \neq z$.

Prova. Primeiro, suponhamos que a desigualdade (1.21) vale. Tomemos quaisquer $x, y \in C$ e $\alpha \in [0,1]$, e seja $z = \alpha x + (1-\alpha)y$. Usando a desigualdade (1.21) duas vezes, obtemos

$$f(x) \geq f(z) + (x-z)^t \nabla f(z),$$
$$f(y) \geq f(z) + (y-z)^t \nabla f(z).$$

Multiplicamos a primeira desigualdade por α,

$$\alpha f(x) \geq \alpha f(z) + \alpha (x-z)^t \nabla f(z),$$

e a segunda por $(1-\alpha)$,

$$(1-\alpha) f(y) \geq (1-\alpha) f(z) + (1-\alpha)(y-z)^t \nabla f(z).$$

Somando estas duas últimas desigualdades obtidas, temos

$$\begin{aligned}
\alpha f(x)+(1-\alpha)f(y) &\geq f(z)+[\alpha(x-z)^t+(1-\alpha)(y-z)^t]\nabla f(z)\\
&= f(z)+(\alpha x^t-\alpha z^t+y^t-z^t-\alpha y^t+\alpha z^t)\nabla f(z)\\
&= f(z)+(\alpha x^t+y^t-\alpha y^t-z^t)\nabla f(z)\\
&= f(z)+(\alpha x+(1-\alpha)y-z)^t\nabla f(z).
\end{aligned}$$

Como $z=\alpha x+(1-\alpha)y$, temos

$$\alpha f(x)+(1-\alpha)f(y)\geq f(z)+(z-z)^t\nabla f(z),$$

ou seja,

$$\alpha f(x)+(1-\alpha)f(y)\geq f(z).$$

Logo, f é convexa, por definição.

Se a desigualdade (1.21) é estrita como enunciado na parte (ii) da proposição, então, ao tomarmos $x\neq y$ e $\alpha\in(0,1)$ acima, as desigualdades precedentes se tornam estritas, assim mostrando a desigualdade estrita de f.

Reciprocamente, suponha que f é convexa. Sejam x e z vetores quaisquer em C, com $x\neq z$, e para $\alpha\in(0,1)$, considere a função

$$g(\alpha)=\frac{f(x+\alpha(z-x))-f(x)}{\alpha}. \tag{1.22}$$

Mostraremos que $g(\alpha)$ é crescente com α e estritamente crescente se f é estritamente convexa. Isto implicará que

$$(z-x)^t\nabla f(x)=\lim_{\alpha\to 0^+}g(\alpha)\leq g(1)=f(z)-f(x),$$

com desigualdade estrita se g é estritamente crescente, assim mostrando que a desejada desigualdade (1.21) vale, e vale estritamente se f é estritamente convexa. De fato, considere quaisquer $\alpha_1,\ \alpha_2$, com $0<\alpha_1<\alpha_2<1$, e sejam

$$\bar{\alpha}=\frac{\alpha_1}{\alpha_2},\quad \bar{z}=x+\alpha_2(z-x). \tag{1.23}$$

Como f é convexa, por hipótese, temos

$$f(\bar{\alpha}\bar{z}+(1-\bar{\alpha})x)\leq\bar{\alpha}f(\bar{z})+(1-\bar{\alpha})f(x),$$

$$f(x+\bar{\alpha}(\bar{z}-x))\leq\bar{\alpha}(f(\bar{z})-f(x))+f(x),$$

$$\frac{f(x+\bar{\alpha}(\bar{z}-x))-f(x)}{\bar{\alpha}}\leq f(\bar{z})-f(x). \tag{1.24}$$

Substituindo (1.23) em (1.24), obtemos, por cálculo direto,

$$\frac{f\left(x+\dfrac{\alpha_1}{\alpha_2}(x+\alpha_2(z-x)-x)\right)-f(x)}{\dfrac{\alpha_1}{\alpha_2}}\leq f(x+\alpha_2(z-x))-f(x),$$

$$\frac{f(x+\alpha_1(z-x))-f(x)}{\alpha_1} \leq \frac{f(x+\alpha_2(z-x))-f(x)}{\alpha_2},$$
$$\frac{f(x+\alpha_1(z-x))-f(x)}{\alpha_1} \leq \frac{f(x+\alpha_2(z-x))-f(x)}{\alpha_2},$$

ou seja,

$$g(\alpha_1) \leq g(\alpha_2).$$

Agora, seja $x \neq y$. Queremos provar que

$$\lim_{\alpha \to 0^+} g(\alpha) \leq g(1) = f(z) - f(x). \tag{1.25}$$

Suponhamos, por absurdo, que

$$\ell =: \lim_{\alpha \to 0^+} g(\alpha) > g(1).$$

Seja

$$\varepsilon < \ell - g(1). \tag{1.26}$$

Temos que $\ell - g(1) > 0$. Pela definição de limite, existe $\delta > 0$ tal que, se $\alpha \in (0, \delta)$, então $\ell - \varepsilon < g(\alpha) < \ell + \varepsilon$. Por (1.26), temos $\ell - \varepsilon > g(1)$. Conseqüentemente, $g(\alpha) > \ell - \varepsilon > g(1)$, isto é, $g(\alpha) > g(1)$, o que contradiz (1.25). $\square$

A seguir, a ilustração geométrica das idéias que estão nas entrelinhas da prova da Proposição 1.25. Na Figura 1.10, aproximamos linearmente f por $z = \alpha x + (1-\alpha)y$. A desigualdade (1.21) implica que

$$f(x) \geq f(z) + (x-z)'\nabla f(z),$$

$$f(y) \geq f(z) + (y-z)'\nabla f(z).$$

Como pode ser visto na figura, segue que $\alpha f(x) + (1-\alpha)f(y)$ está acima de $f(z)$, então f é convexa.

Na Figura 1.11, assumimos que f é convexa, e da geometria da figura notamos que

$$f(x) + \frac{f(x+\alpha(z-x))-f(x)}{\alpha},$$

a qual permanece abaixo de $f(z)$, é monotonicamente não crescente para $\alpha \to 0^+$, e converge para $f(x) + (z-x)'\nabla f(x)$. Segue que $f(z) \geq f(x) + (z-x)'\nabla f(x)$.

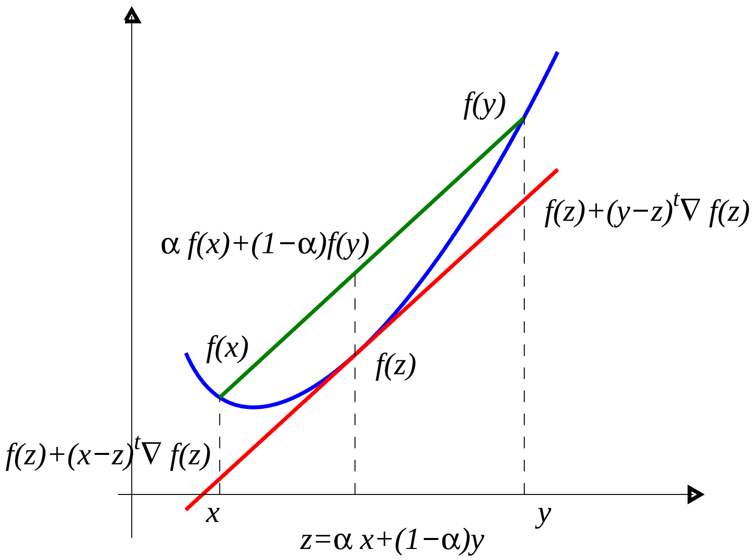

Figura 1.10: Aproximação linear de f por $z = \alpha x + (1-\alpha)y$.

Note uma simples conseqüência da Proposição 1.25. (i): se $f : \mathbb{R}^n \to \mathbb{R}$ é uma função convexa e $\nabla f(x^*) = 0$, então x^* minimiza f sobre R^n. Esta é uma condição clássica suficiente para otimalidade sem restrições, originalmente formulada (em uma dimensão) por Fermat em 1637.

Para funções convexas duas vezes diferenciáveis, há outra caracterização de convexidade, como mostra a proposição seguinte.

Proposição 1.26 *Seja C um subconjunto convexo de R^n e seja $f : \mathbb{R}^n \to \mathbb{R}$ continuamente duas vezes diferenciável sobre R^n.*

(i) Se $\nabla^2 f(x)$ é semidefinido positivo para todo $x \in C$, então f é convexa sobre C;

(ii) Se $\nabla^2 f(x)$ é definido positivo para todo $x \in C$, então f é estritamente convexa sobre C;

(iii) Se C é aberto e f é convexa sobre C, então $\nabla^2 f(x)$ é semidefinido positivo para todo $x \in C$.

Prova.

(i) Suponhamos que $\nabla^2 f(x)$ seja semidefinido positivo para todo $x \in C$. Então, por definição, $p^t \nabla^2 f(x) p \geq 0$, para todo $p \in \mathbb{R}^n$. Sejam $x, y \in C$.Como f é continuamente duas vezes diferenciável sobre R^n, existe $\alpha \in (0,1)$ tal que

$$f(x) = f(y) + \nabla f^t(y)(x-y) + \frac{1}{2}(x-y)^t \nabla^2 f(z)(x-y),$$

onde z é um ponto no segmento de reta que liga x em y, isto é,

$$z = \alpha x + (1-\alpha)y.$$

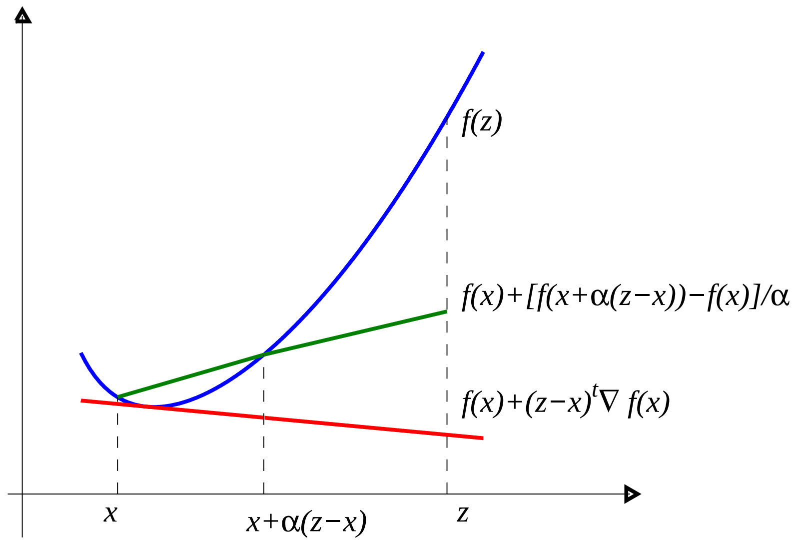

Figura 1.11: Demonstração da Proposição 1.25

Como C é convexo, temos que $z \in C$. Por hipótese, temos que

$$(x-y)^t \nabla^2 f(z)(x-y) \geq 0.$$

Portanto,

$$f(x) \geq f(y) + \nabla f^t(y)(x-y).$$

Logo f é convexa, pelo teorema anterior.

(ii) Suponhamos que $\nabla^2 f(x)$ seja definido positivo para todo $x \in C$. Então, por definição, $p^t \nabla^2 f(x) p > 0$, para todo $p \neq 0$. Sejam x, $y \in C$, com $x \neq y$. Como f é continuamente duas vezes diferenciável sobre R^n, existe $\alpha \in (0,1)$ tal que

$$f(x) = f(y) + \nabla f^t(y)(x-y) + \frac{1}{2}(x-y)^t \nabla^2 f(z)(x-y),$$

onde z é um ponto no segmento de reta que liga x em y, isto é,

$$z = \alpha x + (1-\alpha)y.$$

Como C é convexo, temos que $z \in C$. Por hipótese, temos que

$$(x-y)^t \nabla^2 f(z)(x-y) > 0,$$

para todos x, y tais que $x \neq y$. Portanto,

$$f(x) > f(y) + \nabla f^t(y)(x-y).$$

Logo f é convexa, pela Proposição 1.25

(iii) Suponhamos, por absurdo, que existem algum $x \in C$ e algum $d \in \mathbb{R}^n$ tais que

$$d^t \nabla^2 f(x) d < 0.$$

Consideremos a função

$$g(\lambda) = f(y + \lambda d).$$

Então

$$g'(\lambda) = \nabla f^t(y + \lambda d) d$$

e

$$g''(\lambda) = d^t \nabla^2 f(y + \lambda d) d.$$

Para $\lambda = 0$, temos $g(0) = f(y)$, o que implica que g é convexa no ponto 0. Logo,

$$g''(0) = d^t \nabla^2 f(y) d < 0.$$

Pela conservação de sinal, temos que existe $\delta > 0$ tal que $g''(\lambda) < 0$ para todo $\lambda \in (-\delta, \delta)$. Seja

$$x = y + \frac{\delta}{2} d.$$

Por Taylor, temos que existe $\alpha \in (0, 1)$ tal que

$$f(x) = f(y) + \nabla f^t(y)(x - y) + \frac{1}{2}(x - y)^t \nabla^2 f(z)(x - y),$$

onde

$$z = y + \alpha(x - y) = y + \alpha \frac{\delta}{2} d.$$

Então

$$f(x) = f(y) + \nabla f^t(y)(x - y) + \frac{1}{2}\left(\frac{\delta}{2}\right) d^t \nabla^2 f\left(y + \alpha \frac{\delta}{2} d\right)\left(\frac{\delta}{2} d\right).$$

Temos que

$$\frac{1}{2}\left(\frac{\delta}{2}\right) d^t \nabla^2 f\left(y + \alpha \frac{\delta}{2} d\right)\left(\frac{\delta}{2} d\right) < 0,$$

logo

$$f(x) < f(y) + \nabla f^t(y)(x - y),$$

o que, pela Proposição 1.25, contradiz a hipótese de que f é convexa sobre C. □

Exemplo 10 *Considere a função quadrática*

$$f(x) = x^t Q x + a^t x,$$

onde Q é uma matriz $n \times n$ simétrica e b é um vetor em $\mathbb{R}^n$. Como $\nabla^2 f(x) = 2Q$, segue, usando a Proposição 1.26, que f é convexa se e somente se Q é positiva semidefinida, e estritamente convexa se e somente se Q é positiva definida.

Referências Bibliográficas

[1] N. Andrei. Convex Functions. CAMO research report. `http://www.ici.ro/camo/neculai/convex.pdf`

[2] D. Bertsekas, A. Nedić, A.E. Ozdaglar. *Convex Analysis and Optimization.* Athena Scientific, USA, 2003.

[3] J.B. Hiriart-Urruty, C. Lemaréchal. *Convex Analysis and Minimization Algorithms I: Fundamentals (Grundlehren Der Mathematischen Wissenschaften, No 305).* Springer-Verlag, Berlin, 1993.

[4] E.L. Lima. *Análise Real Volume 1: Funções de Uma Variável Real.* Coleção Matemática Universitária. IMPA, Rio de Janeiro, 2006.

[5] E.L. Lima. *Curso de Análise Volume 1.* Coleção Projeto Euclides. IMPA, Rio de Janeiro, 2006.

[6] E.L. Lima. *Curso de Análise Volume 2.* Coleção Projeto Euclides. IMPA, Rio de Janeiro, 2006.