Programação Não-Linear Irrestrita

Prof. Alexandre Salles da Cunha

Universidade Federal de Minas Gerais
Departamento de Ciência da Computação
Belo Horizonte, Brasil

acunha@dcc.ufmg.br

2022/1

UNIVERSIDADE FEDERAL
DE MINAS GERAIS

UFMG - ICEx
DEPARTAMENTO DE CIÊNCIA DA
COMPUTAÇÃO

O Problema de Otimização

$$
\begin{aligned}
q^* = \min f_0(x) & \\
f_i(x) \leq 0 \qquad & i = 1, \ldots, m \\
h_i(x) = 0 \qquad & i = 1, \ldots, p
\end{aligned}
$$

é um problema de otimização convexo se:

1. A função objetivo $f_0(x) : \mathbb{R}^n \to \mathbb{R}$ é convexa
2. As restrições $f_i(x) : \mathbb{R}^n \to \mathbb{R}$ são convexas, para todo $i = 1, \ldots, m$.
3. As restrições $h_i(x) = 0 : \mathbb{R}^n \to \mathbb{R}^q$ são afim, para todo $i = 1, \ldots, p$ (ou seja, podem ser escritas na forma $Ax = b, A \in \mathbb{R}^{p \times n}, b \in \mathbb{R}^p$.)

$\mathcal{X} = \{x \in \mathbb{R}^n : f_i(x) \leq 0, i = 1, \ldots, m; h_i(x) = 0, i = 1, \ldots, p\}$ é a interseção sublevel sets de funções convexas com p conjuntos afins.

$$\begin{aligned} q^* = &\min f_0(x) \\ & x \in \mathcal{X} \end{aligned}$$

- $\mathcal{X}$ é o conjunto de soluções viáveis. Se $\mathcal{X} \neq \emptyset$, o problema é viável, caso contrário, é inviável ($q^* = \emptyset$).
- Se $\mathcal{X} = \mathbb{R}^n$, o problema é irrestrito.
- Se $\mathcal{X} \neq \emptyset$ e $q^* = -\infty$, o problema é ilimitado inferiormente.
- Pode também ocorrer do problema ser viável, mas nenhum $x \in \mathcal{X}$ atingir q^*. Neste caso, há ínfimo, mas não há minimizador.
- $\mathcal{X}_{opt} := \{x \in \mathcal{X} : f_0(x) = q^*\} = \arg\min\{f_0(x) : x \in \mathcal{X}\}$.
  (Veja que podemos escrever $\mathcal{X}_{opt} := S_{q^*} = \{x \in \mathcal{X} : f_0(x) \leq q^*\}$ - sublevel set - portanto $\mathcal{X}_{opt}$ é cvx e PGO é cvx.)
- Se $\mathcal{X}_{opt} \neq \emptyset$, dizemos que o mínimo é atingido.

Problema Geral de Otimização

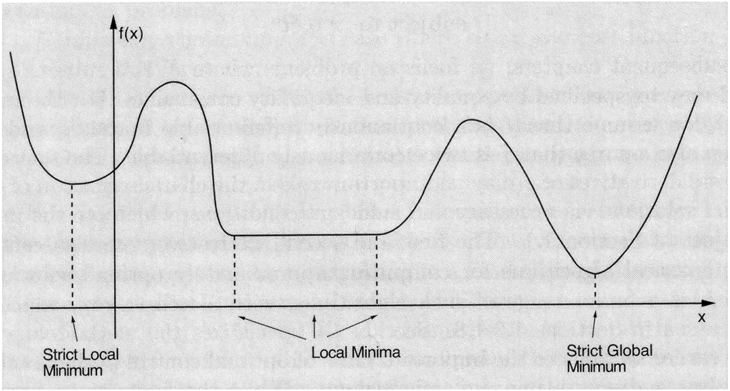
$$\min f_0(x) \tag{1}$$
$$x \in \mathcal{X} \tag{2}$$

Dependendo das propriedades de f_0 e $\mathcal{X}$, o PGO pode ser classificado de diversas formas:

- Problemas contínuos, discretos ou mistos.
- Problemas diferenciáveis, não diferenciáveis
- Problemas convexos, não convexos
- Problema determinístico, estocástico

Foco deste curso: problemas contínuos, determinístico, diferenciáveis, envolvendo uma função objetivo não linear e/ou a caracterização de $\mathcal{X}$ por meio de restrições não lineares, convexas ou não.

Ótimo local, local estrito, global



Todo mínimo local é global se o problema é convexo

Teorema

Considere $\min_{x \in \mathcal{X}} f_0(x)$. Se f_0 é uma função cvx e $\mathcal{X}$ é um conjunto cvx, então qualquer ótimo local é também um ótimo global e $\mathcal{X}_{opt}$ é cvx.

Prova

Seja x^* um otimizador local, $q^* = f_0(x^*)$, $y \in \mathcal{X}$ um ponto viável qualquer e $x(\theta) = \theta y + (1-\theta)x^*$ a combinação cvx destes. Temos que provar que $f_0(y) \geq f(x^*) = q^*$.

$$f_0(x(\theta)) \leq \theta f_0(y) + (1-\theta) f_0(x^*) \qquad f_0 \text{ é cvx}$$

$$f_0(x(\theta)) - f_0(x^*) \leq -f_0(x^*) + \theta f_0(y) + (1-\theta) f_0(x^*)$$

$$f_0(x(\theta)) - f_0(x^*) \leq \theta(f_0(y) - f_0(x^*))$$

Para $\theta > 0$ suficientemente pequeno, $0 \leq f_0(x(\theta)) - f_0(x^*)$ já que x^* é mínimo local. Logo, $f_0(y) \geq f_0(x^*)$.

Teorema de Weirstrass

Toda função $f : \mathbb{R}^n \to \mathbb{R}$ contínua em um conjunto compacto (fechado + limitado) atinge seus valores limites (sup e inf) no conjunto.

Lema

Se $\mathcal{X} \subseteq \textit{dom}\, f_0$ é não vazio e compacto e f_0 é contínua em $\mathcal{X}$, então PGO possui uma solução ótima $x^* \in \mathcal{X}$.

Corolário

Se f_0 é convexa e $\mathcal{X} \subseteq \textit{int dom}\, f_0$ é compacto, as hipóteses do Lema acima valem (convexidade de f_0 garante que é contínua em $\mathcal{X}$).

Estes resultados ainda respondem pouco. E se $\mathcal{X}$ for não compacto, por exemplo $\mathcal{X} = \mathbb{R}^n$?

Definição

Uma função $f : \mathbb{R}^n \to \mathbb{R}$ é coerciva se, para qualquer sequência $\{x^k\} \subset$ *int dom f* que tende para *bd*(*dom f*), o valor da sequência $\{f(x^k)\}$ tende para $+\infty$.

Lema

Uma função contínua $f : \mathbb{R}^n \to \mathbb{R}$ cujo domínio é aberto é coerciva se e somente se os subconjuntos de nível $S_\alpha = \{x \in \mathbb{R}^n : f(x) \leq \alpha\}$ são compactos, para qualquer $\alpha \in \mathbb{R}$.

Lema

Se $\mathcal{X} = \mathbb{R}^n$ e f_0 é contínua e coerciva, então min $f_0(x) : x \in \mathcal{X}$ admite um minimizador x^*.

- Caso convexo: otimalidade global.
- Caso não convexo: otimalidade local.

$$(PGO) \quad \min f_0(x) \qquad\qquad x \in \mathcal{X}$$

Definição

Dado $\overline{x} \in \mathcal{X}, d \in \mathbb{R}^n$, d é direção viável em $\overline{x}$, **se existe** $\alpha > 0$ **tal que** $\overline{x} + \alpha d \in \mathcal{X}$**.** Ou seja: um pequeno deslocamento ao longo de d não inviabiliza o ponto obtido.

- $\mathcal{X} = \mathbb{R}^n$: qualquer $d \in \mathbb{R}^n$ é uma direção viável.
- $\mathcal{X}$ é convexo: $d := (y - \overline{x})$ é viável para qualquer $y \in \mathcal{X}$.
- $\mathcal{X} = \{x \in \mathbb{R}^n : f_j(x) \leq 0, j = 1, \ldots, m\}$: d é direção viável em $\overline{x}$ se e somente se $d^T \nabla f_j(\overline{x}) \leq 0$ para todo $j \in J(\overline{x}) = \{j : f_j(\overline{x}) = 0\}$.

A ideia central em vários métodos é encontrar uma direção d, viável e de descida em x $(\nabla f(x)^T d < 0)$ e implementar $x \leftarrow x + \alpha d$ para $\alpha > 0$.

Proposição - condições necessárias e suficientes de otimalidade

Dado o problema de otimização convexo min $f_0(x) : x \in \mathcal{X}$ onde f_0 é diferenciável em *int dom* f_0. Então:

$$x \in \mathcal{X}_{opt} \iff \nabla f_0(x)^T(y - x) \geq 0, \ \forall y \in \mathcal{X}$$

Pela convexidade de f_0 em *dom* f_0

$$f_0(y) \geq f_0(x) + \nabla f_0(x)^T(y - x), \quad \forall x, y \in dom\ f_0$$

Prova ($\leftarrow$)

$$f_0(y) - f_0(x) \geq \nabla f_0(x)^T(y - x) \geq 0 \ \ \forall y \in \mathcal{X}$$

Prova ($\rightarrow$)

$x \in \mathcal{X}_{opt} \rightarrow f_0(x) \leq f_0(y),\ \forall y \in \mathcal{X}$

- Se $\nabla f_0^T(x) = 0$, a condição $\nabla f_0^T(x)(y - x) \geq 0$ vale trivialmente.
- $\nabla f_0^T(x) \neq 0$, tomamos $x(\theta) = \theta y + (1 - \theta)x$ para $\theta \in [0, 1]$ e assumimos, por absurdo, que $\nabla f_0^T(x)(y - x) < 0$.

$$
\begin{aligned}
f_0(x(\theta)) &= f_0(x) + \nabla f_0(x)^T(x(\theta) - x) + o(\|x(\theta) - x\|) \\
&= f_0(x) + \theta \nabla f_0^T(x)(y - x) + o(\theta\|y - x\|) \\
&= f_0(x) + \text{quant. negativa para } \theta \text{ suf. pequeno}
\end{aligned}
$$

Então teríamos $f_0(x(\theta)) < f_0(x)$: contradição.

Caso convexo: caso irrestrito ou ótimo em $int(\mathcal{X})$

Condição necessária e suficiente para mínimo em x

$$\nabla f_0(x) = 0$$

Caso irrestrito

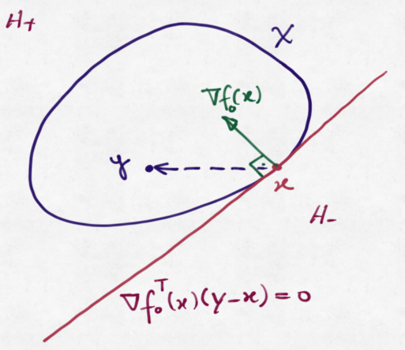
Para $x \in \mathbb{R}^n$ candidato a ótimo, devemos ter para qualquer $y^1 \in \mathbb{R}^n$, $\nabla f_0(x)^T(y^1 - x) \geq 0$. Tomando $y^2 = 2x - y^1$, $\nabla f_0(x)^T(y^2 - x) \geq 0 \to \nabla f_0(x)^T(y^1 - x) \leq 0$ e logo $\nabla f_0(x)^T(y^1 - x) = 0$. Portanto, $\nabla f_0(x) = 0$, uma vez que y^1 é qualquer.

$x \in int(\mathcal{X}) \neq \mathbb{R}^n$

Se $x \in int(\mathcal{X})$ e $\nabla f_0(x) \neq 0$, $-\nabla f_0(x)$ é uma direção viável de descida. Logo existe $y = x - \alpha \nabla f_0(x)$ para $\alpha > 0$ suficientemente pequeno tal que $f_0(y) < f_0(x)$.

- Para o caso convexo, qualquer direção viável em x pode ser representada por $y - x$ para algum $y \in \mathcal{X}$.
- Então x é ótimo se e somente se qualquer direção viável em x for uma direção de crescimento de $f_0(x)$ a partir de x.

Ponto ótimo na fronteira



Caso não convexo

Precisamos de garantias de segunda ordem

As condições anteriores, de primeira ordem, não são suficientes para otimalidade.

Uma questão prática

Não podemos representar uma direção viável pela dirença entre $y \in \mathcal{X}$ e x candidato a ótimo.

(para otimalidade local)

Caso 1: ponto x^* é um ponto interior de $\mathcal{X}$ ou $\mathcal{X} = \mathbb{R}^n$

Por uma aproximação de primeira ordem de uma série de Taylor, temos:

$$\begin{aligned} f_0(x^* + d) - f_0(x^*) &\approx \nabla f_0(x^*)^T d \\ &= \sum_{i=1}^{n} \frac{\partial f_0(x^*)}{\partial x_i} d_i \\ &\geq 0 \end{aligned}$$

Como d é qualquer (x^* é interior), $\frac{\partial f_0(x^*)}{\partial x_i} = 0, i = 1, \ldots, n$

Caso 2: ponto x^* é um ponto na fronteira de $\mathcal{X}$

Para uma direção d viável:

$$\begin{aligned} f_0(x^* + d) - f_0(x^*) &\approx \nabla f_0(x^*)^T d \\ &= \sum_{i=1}^{n} \frac{\partial f_0(x^*)}{\partial x_i} d_i \\ &\geq 0 \end{aligned}$$

Logo $\nabla f_0(x^*)^T d \geq 0$

Exemplo 1 - Problema Irrestrito

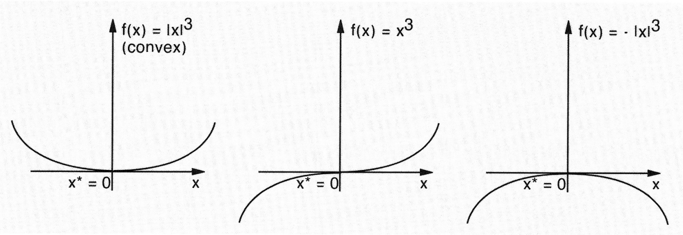
$$\text{minimize} \quad f_0(x_1, x_2) = x_1^2 - x_1 x_2 + x_2^2 - 3x_2 \tag{3}$$
$$x \in \mathbb{R}^2 \tag{4}$$

Exemplo 2 - Problema Restrito, candidato na fronteira

Sabe-se que o ponto mínimo global do Problema abaixo é $(\frac{1}{2}, 0)$. Observe que $\nabla f_0((\frac{1}{2}, 0)) \neq 0$, mas que o ponto atende às CNPO.

$$\begin{aligned} \text{minimize} \quad f_0(x_1, x_2) &= x_1^2 - x_1 + x_2 + x_1 x_2 && (5) \\ x &\in \mathbb{R}^2_+ && (6) \end{aligned}$$

Veja que qualquer direção viável $d = (d_1 \;\; d_2)^T$ deve satisfazer $d_2 \geq 0$ no ponto considerado.


f(x) = |x|3
(convex)
x* = 0
x
f(x) = x3
x* = 0
x
f(x) = - |x|3
x* = 0
x

- As CNPO foram estabelecidas fazendo-se aproximações de primeira ordem do comportamento de $f_0(x)$ nas vizinhanças de um ponto candidato a ótimo local.

- Podemos fazer aproximações de segunda ordem, empregando a matriz Hessiana $\nabla^2 f_0(x)$ da função objetivo.

CNSO

Seja $f_0 \in C^2$ uma função definida em $\mathcal{X} \subseteq \mathbb{R}^n$. Então, se x^* é um ponto de mínimo de f_0 em $\mathcal{X}$ e $d \in \mathbb{R}^n$ é uma direção viável qualquer em x^*, temos:

- $\nabla f_0(x^*)^T d \geq 0$ (CNPO)
- se $\nabla f_0(x^*) = 0$, então $d^T \nabla^2 f_0(x^*) d \geq 0$ (cond. 2a ordem)

Proposição

Seja $f \in C^2$ uma função definida em $\mathcal{X} \subseteq \mathbb{R}^n$. Então, se x^* é um ponto de mínimo de f_0 no interior de $\mathcal{X}$, temos:

- $\nabla f_0(x^*) = 0$
- $d^T \nabla^2 f_0(x^*) d \geq 0, \forall d \in \mathbb{R}^n$, *isto é, a matriz $\nabla^2 f_o(x^*)$ é semipositiva definida.*

Exemplo

$$\begin{array}{rrcl} \text{minimize} & f_0(x_1, x_2) & = & x_1^3 - x_1^2 x_2 + 2x_2^2 \\ & x & \in & \mathbb{R}_+^2 \end{array}$$

Assumindo ponto ótimo no interior, temos as CNPO

$$\begin{aligned} 3x_1^2 - 2x_1 x_2 &= 0 \\ -x_1^2 + 4x_2 &= 0 \end{aligned}$$

Soluções:

- $x^* = (0, 0)$, que ocorre na fronteira, satisfaz CNSO.
- $\overline{x} = (6, 9)$, ponto interior ao domínio, **não satisfaz CNSO** porque a matriz Hessiana neste ponto não é Positiva Semi-Definida.

Insuficiência das Condições Necessárias de 2a ordem

$$\nabla^2 f_0(x_1, x_2) = \begin{pmatrix} 6x_1 - 2x_2 & -2x_1 \\ -2x_1 & 4 \end{pmatrix}$$

Para $\overline{x}^T = (6 \quad 9)$ temos que a Hessiana é H:

```
H  =
    18.   - 12.
  - 12.     4.
-->spec(H)'
 ans  =
 - 2.892444    24.892444
```

Se tomarmos o autovetor $d^T = (-0,49806, -0,86714)$ associado ao autovalor -2.8924 verificamos que d é de descida:
$(\alpha d)^T \nabla^2 f_0(6,9)(\alpha d) = -2.8924(\alpha^2)$.

Seja $f_0 \in C^2$ uma função definida em $\mathcal{X} \subseteq \mathbb{R}^n$. Suponha que x^* seja um ponto no interior de $\mathcal{X}$, satisfazendo:

- $\nabla f_0(x^*) = 0$
- $\nabla^2 f_0(x^*) \succ 0$

Então, x^* é um ponto de mínimo local de $f_0(x)$.

Como $\nabla^2 f_0(x^*) \succ 0$, existe $\alpha > 0$ tal que para qualquer $d \in \mathbb{R}^n$: $d^T \nabla^2 f_0(x^*) d \geq \alpha^2 \|d\|^2$. Então:

$$
\begin{aligned}
f_0(x^* + d) - f_0(x^*) &= \frac{1}{2} d^T \nabla^2 f_0(x^*) d + o(\|d\|^2) \\
&\geq \frac{\alpha}{2} \|d\|^2 + o(\|d\|^2) \\
&\geq 0 \text{ (para } \|d\| \text{ suf. pequeno)} \qquad (7)
\end{aligned}
$$

$$\min f_0 : x \in \mathcal{X}$$

Ponto estacionário

Um ponto x^* que satisfaz a condição $\nabla f_0(x^*) = 0$ é dito estacionário.

Ponto singular

Um ponto de mínimo local x^* que não satisfaz as condições suficientes de otimalidade $(\nabla f_0(x^*) = 0, \nabla^2 f_0(x^*) \succ 0)$ é chamado singular. Caso satisfaça, é chamado de não-singular. Pontos singulares são mais difíceis de se lidar:

- Quando f_0 não é convexa, a sua otimalidade não pode ser assegurada usando-se argumentos suficientemente fáceis.
- Nas vizinhanças destes pontos, a maioria dos métodos de otimização tem convergência lenta ou apresenta comportamento errático.

Problema Quadrático Irrestrito

$$\text{minimize} \quad f_0(x) = \frac{1}{2}x^T Qx - b^T x \tag{8}$$

$$x \in \mathbb{R}^n \tag{9}$$

- Q : matriz $n \times n$ simétrica
- $b \in \mathbb{R}^n$

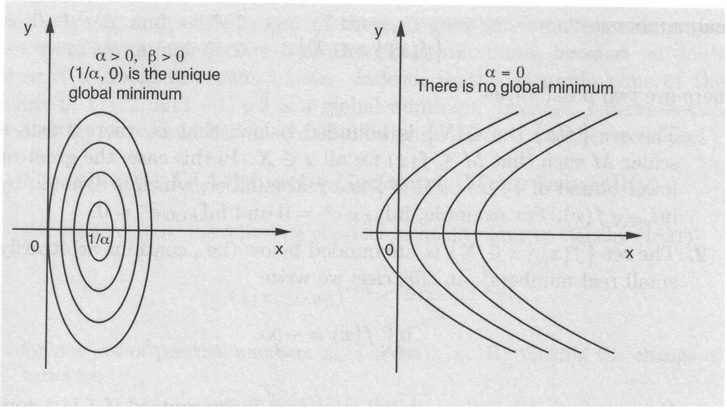
As condições necessárias de primeira e segunda ordem impõem

- $\nabla f_0(x) = 0 \rightarrow Qx - b = 0 \Rightarrow Qx = b$
- $\nabla^2 f_0 = Q \succeq 0.$

- Se $Q \not\succeq 0$, o problema não admite mínimo. Basta se mover ao longo de um autovetor associado a um autovalor negativo que f_0 diminui sem limites.
- Se $Q \succeq 0$, o problema é convexo e, então, qualquer solução de $Qx = b$ é um mínimo do problema. Entretanto, uma solução para este sistema pode não existir, caso $b \notin \mathcal{R}(Q)$.
- Se $Q \succ 0$, Q admite inversa e assim sendo, o único ponto mínimo (global) pode ser obtido através de $x^* = Q^{-1}b$.

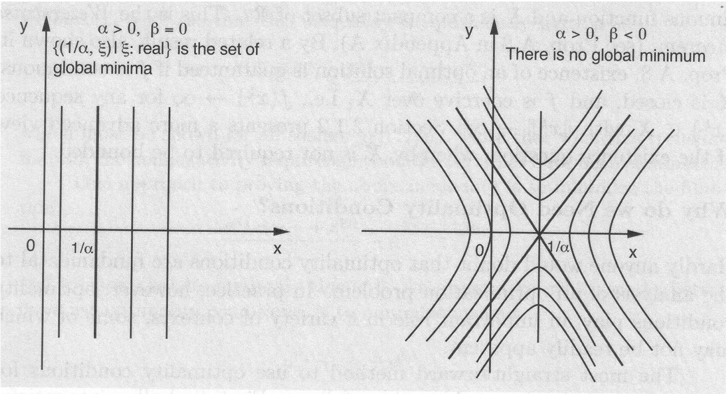
$$f_0(x,y) = \frac{1}{2}(\alpha x^2 + \beta y^2) - x$$

Caso a) Mínimo local é minimo global, $\nabla^2 f_0(x) \succ 0$.
Caso b) Não é possível satisfazer CNPO.



$$f_0(x, y) = \frac{1}{2}(\alpha x^2 + \beta y^2) - x$$

Caso c) $\nabla^2 f_0(x) \succeq 0$, com autovalor nulo. Infinitas soluções ótimas globais (ótimos singulares).
Caso d) $\nabla^2 f_0(x)$ indefinida. Não há mínimo.



- Resolver $\nabla f_0(x) = 0$ é pelo menos tão complicado quanto resolver o problema original.
- As condições de otimalidade fornecem entretanto a base para o desenvolvimento de algoritmos iterativos. Em particular, os algoritmos reconhecem soluções, verificando várias destas condições e terminam quando estas condições são *suficientemente satisfeitas*.
- Em particular, o comportamento (velocidade de convergência, por exemplo) dos algoritmos, nas vizinhanças de um ponto mínimo local, depende das condições de otimalidade serem ou não satisfeitas naquele mínimo.

$$\min f_0(x) : \quad x \in \mathbb{R}^n$$

Descida iterativa

- Dado uma solução inicial qualquer $x^0 \in \mathbb{R}^n$
- Gerar uma seqüencia $\{x^1, x^2, \ldots\}$ de pontos, de forma que:

$$f_0(x^{k+1}) < f_0(x^k), \;\; k = 0, 1, \ldots$$

- Uma vez que a cada iteração k a função objetivo melhore, **esperamos** que o valor de f decresca para o seu valor mínimo. Isto de fato acontece ?

Iteração típica

Dado um ponto $x \in \mathbb{R}^n : \; \nabla f_0(x) \neq 0$, opera diante da seguinte iteração típica:

$$x(\alpha) = x - \alpha \nabla f_0(x), \;\; \alpha \in \mathbb{R}_+$$

- Ou seja, a cada iteração, dado o ponto candidato atual x para a resolução do problema, gera um novo candidato $x(\alpha)$, que corresponde a um deslocamento a partir de x na direção de $-\nabla f_0(x)$.
- Por que ?

A direção contrária ao gradiente é de descida

Pela Expansão em Série de Taylor de Primeira Ordem em torno de x temos:

$$\begin{aligned} f_0(x(\alpha)) &= f_0(x) + \nabla f_0(x)^T (x(\alpha) - x) + o(\|x(\alpha) - x\|) \\ &= f_0(x) - \alpha \|\nabla f_0(x)\|^2 + o(\alpha \|\nabla f_0(x)\|) \\ &= f_0(x) - \alpha \|\nabla f_0(x)\|^2 + o(\alpha) \end{aligned}$$

- Uma vez que $\alpha \|\nabla f_0(x)\|^2$ domina $o(\alpha)$ para valores suficientemente pequenos de α, temos que:

$$f_0(x(\alpha)) < f_0(x)$$

Iteração típica

Dado um ponto $x \in \mathbb{R}^n$: $\nabla f_0(x) \neq 0$ e uma direção de descida, isto é, $d \in \mathbb{R}^n$ satisfazendo $\nabla f_0(x)^T d < 0$,

$$x(\alpha) = x + \alpha d, \quad \alpha \in \mathbb{R}_+$$

Por argumentos semelhantes, temos:

$$\begin{aligned} f_0(x(\alpha)) &= f_0(x) + \nabla f_0(x)^T (x(\alpha) - x) + o(\|x(\alpha) - x\|) \\ &= f_0(x) + \alpha \nabla f_0(x)^T d + o(\alpha \|d\|) \\ &= f_0(x) + \alpha \nabla f_0(x)^T d + o(\alpha) \end{aligned}$$

e como $\alpha \nabla f_0(x)^T d$ domina $o(\alpha)$ para α suficientemente pequeno, temos:

$$f_0(x(\alpha)) < f_0(x)$$

Iteração típica

$$x^{k+1} = x^k + \alpha^k d^k, \;\; k = 0, 1, \ldots$$

onde, a cada iteração k:

- $\nabla f_0(x^k) \neq 0$,
- $\nabla f_0(x^k)^T d^k < 0$ (a direção d^k é de descida)
- $\alpha^k > 0$

Variações do método:

- Como determinar direções d^k ?
- Dada a direção d^k, como determinar α^k ?

- Dependendo da função objetivo (não raro !) apresenta comportamento errático: $\nabla f_0(x^{k+1})$ e $\nabla f_0(x^k)$ são quase paralelos.
- Consequentemente, pode apresentar (com frequëncia !) baixa velocidade de convergência
- Uma alternativa consiste em fazer uma deflexão na direção do gradiente.

Determinação do passo α: tradeoff entre redução grande de $f_o(x^k + \alpha d^k)$ e realização de poucas aliações da função objetivo em valores candidatos de α.

- Enorme impacto. Presente em quase todos os métodos em PNL !

- Busca exata de α. Encontrar $\alpha : \nabla f_0(x^{k+1})^T d^k = \frac{df_0(x^k + \alpha d^k)}{d\alpha} = 0$ (na prática, não pode ser feita com número finito de passos).

- Busca inexata, mas que atenda à algumas condições técnicas (redução suficiente na função objetivo e na norma do gradiente, visando evitar convergência para um ponto não estacionário.)

- Envolvem pelo menos duas fases: *bracketing* e *sectioning*.

- Às vezes, podem envolver uma terceira fase, de interpolação.

Minimização exata ao longo da direção d^k

Calculamos o valor exato que minimiza a $f_0(\cdot)$, ao longo da linha $x^k + \alpha d^k$:

$$\alpha^k = \text{arg min}_{\alpha \geq 0}\{f_0(x^k + \alpha d^k)\}$$

Técnicas de redução de intervalo: Fibonacci ou Seção Áurea.

- Embora tentem reduzir o número de avaliações da função objetivo, são computacionalmente caras.
- Por que gastar muito esforço nas primeiras iterações, quando provavelmente os pontos iniciais estão distantes do ótimo ?

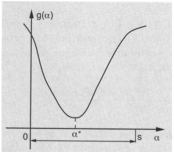
Hipótese

Vamos assumir que $g(\alpha) = f_0(x^k + \alpha d^k), \alpha \geq 0$ é unimodal para $\alpha \in [0, s]$, isto é, $g(\alpha)$ possui um único mínimo no intervalo $[0, s]$ e que α^* é este minimizador.

Fato

Se $x(\alpha^*)$ é um minimizador de $f_0 \in C^2$ ao longo da linha $x(\alpha) = x^k + \alpha d^k$, no ponto de mínimo $x(\alpha^*)$, devemos observar $g'(\alpha^*) = 0$ e $g''(\alpha^*) \geq 0$.

- Desejamos determinar α^* (busca exata) ou um valor de α suficientemente bom (busca inexata).



Dados x e d, $f_0(x + \alpha d)$ é unimodal

Como sucessivamente escolher n pontos $\{\alpha_k : k = 1, \ldots, n\}$ no intervalo $[0, s]$ para avaliar $f_0(x + \alpha_k d)$, de forma que possamos determinar a menor região (subintervalo) possível de $[0, s]$ onde o mínimo deve permanecer ?

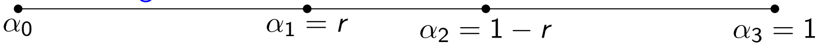
- $0 = \alpha_1 < \alpha_2 < \cdots < \alpha_{n-1} < \alpha_n = s$.
- Intervalo de incerteza: $[\alpha_{k-1}, \alpha_{k+1}]$ onde $\alpha_k = \arg\min\{f_0(x + \alpha_i d) : i = 1, \ldots, n\}$.
- Métodos bastante empregados:
  - Fibonacci
  - Seção Áurea

Emprega-se algum método de redução de intervalo até que, para o intervalo de incerteza, seja verificado

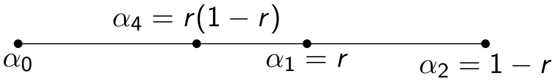
$$\alpha_{k+1} - \alpha_{k-1} \leq \epsilon,$$

onde ϵ é um parâmetro de implementação suficientemente pequeno.

Intervalo original:



Intervalo reduzido (por exemplo eliminando o subintervalo $(\alpha_2, \alpha_3]$)



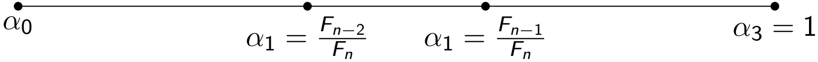
Logo a razão procurada deve satisfazer $r = (1 - r)(1 - r)$.
Particionamos o intervalo segundo a razão áurea:
$r = \frac{3-\sqrt{5}}{2} \approx 0.381966$

Método de Fibonacci

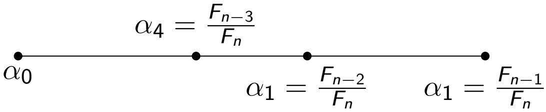
Números de Fibonacci: $\{1, 1, 2, 3, 5, 8, 13, \dots\}$

- $F_0 = F_1 = 1$
- para $n \geq 2$ temos $F_n = F_{n-1} + F_{n-2}$

Dada uma escolha de n (na figura $n = 9$), temos o intervalo original:



Após a redução:



$$0.618 \approx 1 - r = \lim_{n \to \infty} \frac{F_{n-1}}{F_n}$$

Para evitar o elevado custo computacional associado à determinação exata do passo, uma idéia natural (que pode não funcionar) consiste em:

Redução sucessiva do intervalo

- Arbitramos $\alpha^k = s$ e avaliamos $f_0(x^k + sd^k)$
- Se $f_0(x^k) > f_0(x^k + sd^k)$ aceitamos o ponto, isto é, fazemos $x^{k+1} = x^k + sd^k$.
- Caso contrário, reduzimos s por um certo fator, e o processo se repete até que $f_0(x^k) > f_0(x^k + sd^k)$ para o valor de s em mãos.

- Embora este método possa funcionar em muitas situações práticas, carece de sustentação teórica, uma vez que a redução de custo observada em cada iteração pode não ser suficiente para garantir convergência para o mínimo de $f_0(x)$.

Falha deste método de reduções sucessivas

Exemplo

$f_0(x)$ estritamente convexa, continuamente diferenciável, mínimo em $x^* = 0$:

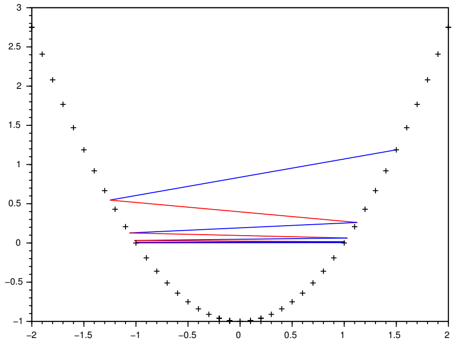
$$f_0(x) = \begin{cases} \frac{3(1-x)^2}{4} - 2(1-x) & \text{se} \quad x > 1 \\ \frac{3(1+x)^2}{4} - 2(1+x) & \text{se} \quad x < -1 \\ x^2 - 1 & \text{se} \quad -1 \le x \le 1 \end{cases}$$

$$\nabla f_0(x) = \begin{cases} \frac{3x+1}{2} & \text{se} \quad x > 1 \\ \frac{3x-1}{2} & \text{se} \quad x < -1 \\ 2x & \text{se} \quad -1 \le x \le 1 \end{cases}$$

Observe que:

1. $f_0(x)$ é par: $f_0(x) = f_0(-x)$ e que para $\overline{x}$ e $\hat{x}$, temos $f_0(\overline{x}) < f_0(\hat{x}) \leftrightarrow |\overline{x}| < |\hat{x}|$.
2. para $x > 1$, $x - \nabla f_0(x) = -\left(\frac{x+1}{2}\right) < -1$, e então:
   - $|x - \nabla f_0(x)| > 1$
   - $|x| > |x - \nabla f_0(x)|$
   - $f_0(x) > f_0(x - \nabla f_0(x))$
3. para $x < -1$, $x - \nabla f_0(x) = \frac{-x+1}{2} > 1$, mesmas conclusões.
4. O que acontece então se iniciarmos o método com um ponto $x^0 : |x^0| > 1$ e um valor de s que, por substituições sucessivas alcance o valor de 1 ?

- $\{x^k\}$ satisfaz $|x^k| > 1, \forall k$, apresentando dois pontos limites $x = 1$ e $x = -1$, não podendo convergir para $x = 0$.



Decréscimo suficiente de f_0 (condição de Armijo)

A redução observada deve ser proporcional ao deslocamento α e também à magnitude do gradiente no ponto x^k.

$$f_0(x^k + \alpha d^k) \leq f_0(x^k) + c_1 \alpha \nabla f_0(x^k)^T d^k \qquad (10)$$

onde $c_1 \in (0, 1)$ é um escalar.

- O lado direito de (10) é uma função linear em α com taxa de variação negativa.
- Na prática, os valores adotados para c_1 são pequenos, da ordem de 10^{-4} ou 10^{-3}.

Condição de curvatura (condições de Wolfe)

A condição de decréscimo mínimo não é suficiente para garantir que o método tenha um progresso substancial na direção de um ponto estacionário, porque é satisfeita por valores extremamente pequenos de α. Para evitar deslocamentos muito pequenos, usa-se a condição adicional:

$$\nabla f_0(x^k + \alpha d^k)^T d^k \geq c_2 \nabla f_0(x^k)^T d^k \tag{11}$$

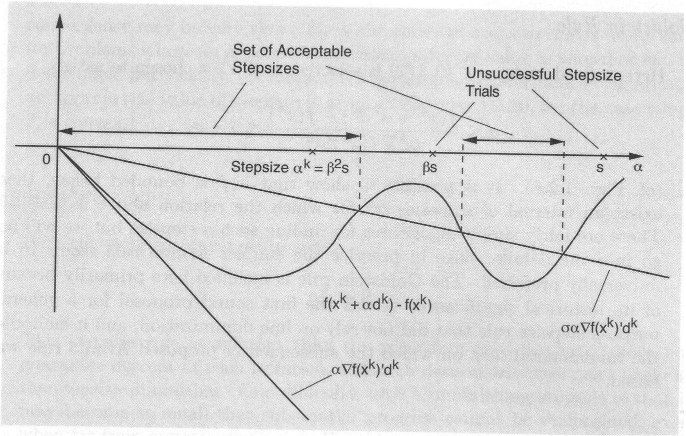
onde $c_2 \in (c_1, 1)$ é um escalar.

- Lado esquerdo de (11) é $g'(\alpha)$.
- Se $g'(\alpha) <<< 0$, temos a indicação de que f_0 pode ser reduzida substancialmente com um deslocamento adicional ao longo da mesma direção d^k. Devemos rejeitar o ponto.
- Se $g'(\alpha)$ é pouco negativa ou mesmo positiva, faz sentido interromper a busca unidirecional, pois há a indicação de pequeno ou nenhum progresso adicional ao longo de d^k.

- Assume que a direção d^k é de descida.
- É um método que estabelece um decréscimo mínimo para que o ponto seja aceito, e evita o aceite de pontos muito pequenos por implementar uma estratégia de backtracking, garantindo assim a convergência.

- Fixamos s, β, σ, escolhendo $0 < \beta < 1, 0 < \sigma < 1$ (muitas vezes, $s = 1$ caso exista a garantia de que o α mínimo está entre $[0, 1]$, caso contrário, implementa-se uma estratégia de bracketing).
- Fazemos $\alpha^k = \beta^{m_k} s$, onde m_k é o primeiro inteiro não negativo m para o qual:
  $f_0(x^k) - f_0(x^k + (\beta^m s)d^k) \geq -\sigma(\beta^m s)\nabla f_0(x^k)^T d^k$
- Ou seja: avaliamos os pontos $(x^k + sd^k), (x^k + \beta sd^k), (x^k + \beta^2 sd^k), \ldots$ até que o primeiro deles forneça o decréscimo suficiente.

Pontos onde $\alpha \approx 0$ só são aceitos se $\|\nabla f_0(x^k)\|$ for pequena, pois a tendência é avaliar pontos mais afastados primeiro.



- *Busca unidirecional exata* (no sentido de garantir que α^k seja bastante próximo de α^*): desde que o conjunto de nível $S_{f_0(x^0)}$ seja compacto, $\{x^k\}$ possui pelo menos um ponto limite. Alguns resultados mostraram que o método converge para um ponto estacionário (Curry, 1944 & Cauchy, 1847).
- Isto ocorre porque os pontos de mínimo locais tendem a atrair as sequências (Teorema da Captura). Isto vale para métodos que tenham em algum passo uma busca unidirecional exata na direção do gradiente.
- Busca inexata: Alguns resultados garantem que o método converge para um ponto estacionário caso α^k promova um decréscimo suficiente da função objetivo, no espírito do método de Armijo.
  (se a direção não se tornar quase ortogonal a $\nabla f_0(x^k)$.)

Teorema - Zoutendijk

Considerando a iteração típica $x^{k+1} \leftarrow x^k + \alpha^k d^k$ onde $\nabla f_0(x^k)^T d_k < 0$ e α^k satisfaz as condições (10) e (11). Suponha que:

1. f_0 seja limitada inferiormente em $\mathbb{R}^n$
2. f_0 seja continuamente diferenciável em um conjunto aberto N contendo o conjunto de subnível $S_{f_0(x^0)} = \{x \in \mathbb{R}^n : f_0(x) \leq f_0(x^0)\}$
3. ∇f_0 seja Lipschitz contínua em N, isto é, existe uma constante $L > 0$ tal que $\|\nabla f_0(x) - \nabla f_0(\hat{x})\| \leq L\|x - \hat{x}\|$ para qualquer $x, \hat{x} \in N$.

Então $\sum_{k \geq 0} cos^2(\theta_k)\|\nabla f_0(x^k)\|^2 < \infty$ onde $\cos(\theta_k) = \frac{-\nabla f_0(x^k)^T d^k}{\|\nabla f_0(x^k)\|\|d^k\|}$

Consequências:

$$cos^2(\theta_k)\|\nabla f_0(x^k)\|^2 \rightarrow 0 \tag{12}$$

Se $\theta_k << \frac{\pi}{2}$ ou $\cos(\theta_k) \geq \delta$, $\lim_{k\rightarrow\infty}\|\nabla f_0(x^k)\| = 0$. (estacionário)

- De (11): $(\nabla f_0(x^{k+1}) - \nabla f_0(x^k))^T d^k \geq (c_2 - 1)\nabla f_0(x^k)^T d^k$
- Condição Lipschitz: $(\nabla f_0(x^{k+1}) - \nabla f_0(x^k))^T d^k \leq \alpha^k L \|d^k\|^2$
- Combinando os dois resultados: $\alpha^k \geq \frac{(c_2-1)}{L} \frac{\nabla f_0(x^k)^T d^k}{\|d_k\|^2}$
- Substituindo em (10): $f_0(x^{k+1}) \leq f_0(x^k) - c_1 \frac{1-c_2}{L} \frac{(\nabla f_0(x^k)^T d^k)^2}{\|d^k\|^2} =$ $f_0(x^k) - c\cos^2(\theta^k)\|\nabla f_0(x^k)\|^2$
- Somando para os primeiros k termos: $f_0(x^{k+1}) - f_0(x^0) \leq -c\sum_{j=1}^k \cos^2(\theta^j)\|\nabla f_0(x^j)\|^2$
- Como f_0 é limitada inferiormente, tomando o limite quando $k \to \infty$, $\sum_{j=1}^k \cos^2(\theta^j)\|\nabla f_0(x^j)\|^2 < \infty$.

Método globalmente convergente significa

método que gera uma subsequência convergente para ponto estacionário, $\lim_{k\to\infty}\|\nabla f_0(x^k)\| = 0$, independentemente de x^0.

Veja que:

$$\sum_{k\geq 0}\cos^2(\theta^k)\|\nabla f_0(x^k)\|^2 < \infty \to \lim_{k\to\infty}\cos^2(\theta^k)\|\nabla f_0(x^k)\|^2 = 0$$

Logo, se $\theta^k << \frac{\pi}{2}$, $\cos^2(\theta^k) >> 0$, necessariamente $\lim_{k\to\infty}\nabla f_0(x^k) = 0$.

Ou seja, se as direções forem "suficientemente não ortogonais" a $\nabla f_0(x^k)$, temos convegência global para ponto estacionário.

Algoritmos de busca unidirecional que satisfazem $\lim_{k\to\infty}\nabla f_0(x^k) = 0$ são chamados de globalmente convergentes.

- Gradiente puro: $\cos(\theta_k) = 1$, então, globalmente convergente, se escolha do passo garante (12).
- Método de descida qualquer, $\nabla f_0(x^k)^T d < 0$: Desde que $\cos(\theta_k) > 0$, a condição (12) vale e é globalmente convergente.
- Método de descida (toda iteração) + reinicialização no gradiente satisfazendo (10)+(11), isto é, efetua-se um passo na direção de $\nabla f_0(x^k)$ a cada p iterações: globalmente convergente para um ponto estacionário.
- Métodos baseados em defleção do gradiente, $d^k = -D^k \nabla f_0(x^k)$ onde $D^k \succ 0$: satisfazendo (10) e (11) e $\kappa(D^k) := \|D^k\| \|(D^k)^{-1}\| \leq M$ (D^k bem condicionada para todo k), que $\cos(\theta^k) \geq 1/M$ e (12) vale. Globalmente convergente.

Definição

Se a sequência nos reais $\{r^k\}$ converge para r^*, assumindo $r^k \neq 0, \forall k$, a ordem de convergência é o supremo dos reais não negativos p para o qual

$$\lim_{k\to\infty} \frac{|r^{k+1} - r^*|}{|r^k - r^*|^p} = \beta < \infty.$$

1. Se a ordem é $p = 1$ e a taxa $\beta < 1$ dizemos que a convergência é linear (ou geométrica, significa dizer que $\{r^k\}$ converge pelo menos tão rápido quanto $\delta\beta^k$ para alguma constante δ).
2. Quando $p > 1$ ou quando $p = 1, \beta = 0$, temos convergência super-linear.
3. Dado p, quanto menor β maior a velocidade de convergência. β depende do algoritmo e também do problema.

$r^k = a^k$ para $a \in (0, 1)$

- $\lim_{k\to\infty} a^k = 0$ (sequência converge)
- $\lim_{k\to\infty} \frac{|a^{k+1}-0|}{|a^k-0|} = a$ (ordem de convergência é pelo menos 1)
- $\lim_{k\to\infty} \frac{a^{k+1}}{(a^k)^p} = a^{-(p-1)k+1} = \infty$ para $p > 1$.

Ordem de convergência é 1, a taxa de convergência é a
Convergência linear, $a < 1$.

$r^k = \frac{1}{k}$

- $\lim_{k\to\infty} \frac{1}{k} = 0$ (sequência converge)
- $\lim_{k\to\infty} \frac{k}{k+1} = 1$
- $\lim_{k\to\infty} \frac{\frac{1}{k+1}}{\frac{1}{k^p}} = \lim_{k\to\infty} \frac{k^p}{k+1} = \infty$ para $p > 1$
  (ordem de convergência não é superior a 1)

Ordem de convergência é 1, mas a taxa de convergência é 1, logo não há convergência linear

$r^k = (\frac{1}{k})^k$

- $\lim_{k\to\infty}(\frac{1}{k})^k = 0$ (sequência converge)
- $\lim_{k\to\infty} \frac{k^k}{(k+1)^{k+1}} = 0$ (ordem de convergência é pelo menos 1 e neste caso é superlinear, pois a taxa $\beta = 0$).
- para $p > 1$, $\lim_{k\to\infty} \frac{k^{pk}}{(k+1)^{k+1}} = \infty$ (ordem de convergência não excede 1)

Ordem de convergência é 1, convergência superlinear

Sequência gerada por $r^{k+1} = \frac{r^k+1}{2}$

1. Observe que $r^k = \frac{r^0+2^k-1}{2^k}$.
2. Logo, $\{r^k\} \to 1$ para qualquer r^0 inicial.
3. $p = 1$: $\lim_{k\to\infty} \frac{\frac{r^0+2^{k+1}-1}{2^{k+1}}-1}{\frac{r^0+2^k-1}{2^k}-1} = \frac{1}{2}$ (ordem 1, convergência pelo menos linear)
4. $p > 1$: $\lim_{k\to\infty} \frac{|r^{k+1}-1|}{|r^k-1|^p} = \lim_{k\to\infty} \frac{2^{kp}}{2^{k+1}} = \infty$

$\Rightarrow$ Assim sendo, temos convergência apenas linear.

Sequência gerada por $r^{k+1} = 1 + \frac{r^k - 1}{2^k}$

- $\{r^k\} \to 1$ para qualquer r^0 inicial.
- $\frac{r^{k+1}-1}{r^k-1} = \frac{1}{2^k} \to 0, k \to \infty$

$\Rightarrow$ Assim sendo, temos convergência super-linear.

- A análise de convergência dos métodos é feita avaliando-se o comportamento do método nas vizinhanças do ponto ótimo local, não singular.
- Não faremos, como no caso de outras áreas da Programação Matemática (Programação Inteira por exemplo), uma contagem de operações elementares para alcançar o ponto ótimo local.
- Estamos interessados no comportamento assintótico da função $e(x^k)$ (definida a seguir), isto é, estamos interessados na cauda da curva.

- Nos restringimos às sequências $\{x^k\}$ que convergem para um único ponto limite x^*.
- A taxa de convergência é avaliada usando uma *função de erro* $e : \mathbb{R}^n \rightarrow \mathbb{R}$ satisfazendo as seguintes propriedades:
  - $e(x) \geq 0, \quad \forall x \in \mathbb{R}^n$
  - $e(x^*) = 0$
- Exemplos de função erro:
  - Distância Euclideana: $e(x) = \|x - x^*\|$
  - Desvio da função objetivo ótima $e(x) = |f(x) - f(x^*)|$

Suponha que um método do tipo gradiente seja aplicado a uma função $f_0 \in C^2$ e que uma sequência $\{x^k\}$ convergente para um mínimo local não singular x^* seja obtida. Pelas condições necessárias de otimalidade em x^*, temos que $\nabla f_0(x^*) = 0$ e $\nabla^2 f_0(x^*) \succ 0$ (x^* é não singular !).

- Por Taylor temos
  $f_0(x) = f_0(x^*) + \frac{1}{2}(x - x^*)^T \nabla^2 f_0(x^*)(x - x^*) + o(\|x - x^*\|^2)$
- Para x suficientemente próximo de x^*,
  $f_0(x) \approx f_0(x^*) + \frac{1}{2}(x - x^*)^T \nabla^2 f_0(x^*)(x - x^*)$

Assim sendo, mesmo quando f não é quadrática, **o comportamento do método na vizinhança do ponto estacionário não singular depende de** $\nabla^2 f_0(x^*)$.

Hipóteses que não tornam o problema menos geral

- $x^* = 0$ e $f_0(x^*) = 0$
- É claro que isto nem sempre é verdade em todo problema quadrático. Podemos entretanto fazer uma mudança de variáveis $y = x - x^*$ e somar/subtrair uma constante de f_0.
- Logo, podemos usar as funções erro $e(x) = \|x - x^*\|$ e $e(x) = |f_0(x) - f_0(x^*)|$ na análise de convergência nas vizinhanças do ótimo.
- Então o problema toma a forma:

$$\min \ f_0(x) = x^T Qx : \ \ x \in \mathbb{R}^n$$

onde $Q \in \mathcal{S}^n_{++}$.

$$
\begin{aligned}
f_0(x) &= \frac{1}{2}x^T Q x \\
\nabla f_0(x) &= Qx \\
\nabla^2 f_0(x) &= Q
\end{aligned}
$$

O método do Gradiente Puro opera segundo a iteração:

$$
\begin{aligned}
x^{k+1} &= x^k - \alpha^k \nabla f_0(x^k) \\
&= (I - \alpha^k Q)x^k \\
\Rightarrow \quad \|x^{k+1}\|_2^2 &= (x^k)^T (I - \alpha^k Q)^2 x^k
\end{aligned}
$$

Recordando

- Sabemos que para qualquer $x \in \mathbb{R}^n$:

$$x^T(I - \alpha^k Q)^2 x \leq \lambda_{max}((I - \alpha^k Q)^2)\|x\|_2^2$$

onde $\lambda_{max}((I - \alpha^k Q)^2)$ é o maior autovalor de $(I - \alpha^k Q)^2$.

- Se $\{\lambda_i : i = 1, \ldots, n\}$ são os autovalores de Q, então os autovalores de $(I - \alpha^k Q)^2$ são: $\{(1 - \alpha^k \lambda_i)^2 : i = 1, \ldots, n\}$.
- Então o máximo autovalor de $(I - \alpha^k Q)^2$ é dado por:

$$\lambda_{max}((1 - \alpha^k m)^2) = \max\{(1 - \alpha^k m)^2, (1 - \alpha^k M)^2\},$$

onde m, M denotam o menor e o maior autovalor de Q, respectivamente.

Recordando

- Logo:

$$\|x^{k+1}\|_2^2 = (x^k)^T(I-\alpha^k Q)^2(x^k) \leq \max\{(1-\alpha^k m)^2, (1-\alpha^k M)^2\}\|x^k\|^2$$

$$\frac{\|x^{k+1}\|}{\|x^k\|} \leq \max\{|1-\alpha^k m|, |1-\alpha^k M|\}, \text{ assumindo } x^k \neq 0$$

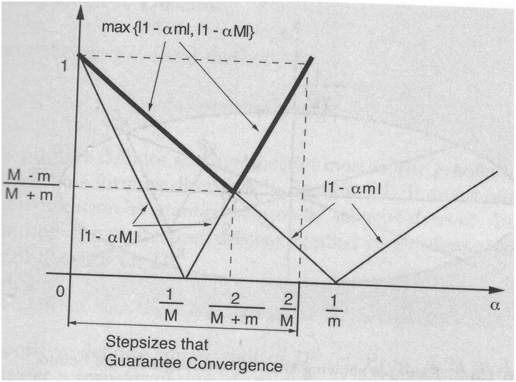
$$\frac{\|x^{k+1}\|}{\|x^k\|} \leq \max\{|(1-\alpha^k m)|, |(1-\alpha^k M)|\}, \quad x^k \neq 0$$

1. Se $|1-\alpha^k m| \geq |1-\alpha^k M|$ ($|1-\alpha^k m| \leq |1-\alpha^k M|$) a desigualdade é satisfeita na igualdade sempre que x^k for proporcional ao autovetor associado a m (associado a M)
2. $\alpha^k = \frac{2}{M+m}$ minimiza o lado direito da desigualdade.
3. Para $\alpha^k = \frac{2}{M+m}$ a melhor taxa permite escrever

$$\frac{\|x^{k+1}\|}{\|x^k\|} \leq \frac{M-m}{M+m}.$$

   Esta desigualdade será justa se x^0 for proporcional a um destes autovetores.
4. Se $m <<< M$, $\frac{M-m}{M+m} \approx 1$.

Análise de Convergência: taxa pode ser muito baixa !



- Para $Q \in \mathcal{S}^n$, o parâmetro $\kappa_2(Q) = \|Q\|_2\|Q^{-1}\|_2 = \frac{M}{m}$ é chamado de número de condição de Q.
- Quando $\kappa_2(Q)$ é elevado, o problema é chamado de mal-condicionado.
- Observe que $\lim_{\frac{M}{m}\to\infty} \frac{M-m}{M+m} = 1$ e a velocidade de convergência é muito baixa.
- Valores elevados de $\kappa_2(Q)$ são característicos de problemas com curvas de nível, elipsóides, alongadas. Nestes casos, o método do gradiente apresenta uma baixa taxa de convergência, e na medida em que o ponto x^k se aproxima do ponto estacionário, os gradientes tornam-se praticamente paralelos.

As propriedades de convergência do Método do Gradiente (conv. linear) são as mesmas para qualquer função objetivo não linear. Se $k_2(\nabla^2 f_0(x^*))$ é grande, o problema é malcondicionado.

- Pode-se provar que se α^k for obtido via busca unidirecional exata, vale a expressão:

$$\frac{|f_0(x^{k+1})|}{|f_0(x^k)|} \leq \left(\frac{M-m}{M+m}\right)^2$$

- Pode-se também mostrar que a desigualdade acima é satisfeita na igualdade para todo k, para um x^0 convenientemente escolhido.

Recordando

Para a matriz simétrica $A \succeq 0$:

- Existe uma matriz simétrica Q tal que $Q^2 = A$ e é designada $Q = A^{\frac{1}{2}}$ a raiz quadrada simétrica de A
- $A^{\frac{1}{2}}$ admite inversa $A^{-\frac{1}{2}}$ se e somente se $A \succ 0$.
- Valem as seguintes propriedades $A^{-\frac{1}{2}}A^{-\frac{1}{2}} = A^{-1}$ e $AA^{\frac{1}{2}} = A^{\frac{1}{2}}A$

Os métodos que operam sob a iteração típica

$$x^{k+1} = x^k - \alpha^k D^k \nabla f_0(x^k)$$

para $D^k \in \mathcal{S}^n_{++}$, podem ser vistos como métodos que implementam uma mudança de escala do problema, ou seja, esta iteração típica corresponde ao método do Gradiente puro em um outro sistema de coordenadas, que depende de D^k.

- Defina $S = (D^k)^{\frac{1}{2}}$ e $x = Sy$
- O problema $f_0(x) : x \in \mathbb{R}^n$ pode ser escrito em termos de y

$$\min f_0(x) = \min f_0(Sy) = \min h(y)$$

- O método do gradiente para resolver $\min h(y) : \ y \in \mathbb{R}^n$ gera

$$y^{k+1} = y^k - \alpha^k \nabla h(y^k)$$

que, multiplicado por S, fornece: $Sy^{k+1} = Sy^k - \alpha^k S \nabla h(y^k)$

Substituindo $x^k = Sy^k, \nabla h(y^k) = S \nabla f(x^k), S^2 = D^k$:

$$x^{k+1} = x^k - \alpha^k D^k \nabla f(x^k).$$

- Ou seja: fazer a operação defletida $x^{k+1} = x^k - \alpha^k D^k \nabla f(x^k)$ consiste em fazer uma busca do tipo gradiente pura no espaço y.

Aplicando os resultados de convergência para o caso quadrático em y

$$f_0(x) = \frac{1}{2}x^T Qx \to h(y) = f_0(Sy) = \frac{1}{2}y^T SQSy$$
$$y^{k+1} = y^k - \alpha^k \nabla h(y^k) = y^k - \alpha^k SQSy^k$$

$$\frac{\|y^{k+1}\|}{\|y^k\|} \leq \max\{|(1-\alpha^k m^k)|, |(1-\alpha^k M^k)|\}, \quad y^k \neq 0$$

onde m^k, M^k são o menor e o maior autovalor de $SQS = (D^k)^{\frac{1}{2}} Q (D^k)^{\frac{1}{2}}$.

- Observe então que se D^k for uma matriz próxima a Q^{-1} temos

$$\nabla h^2(y^k) = (D^k)^{\frac{1}{2}} Q (D^k)^{\frac{1}{2}} \approx I.$$

  e então $m^k \approx M^k \approx 1$ e o problema torna-se muito bem condicionado. Então $S = Q^{-\frac{1}{2}}$ corrige perfeitamente a escala.
- Se o problema não for quadrático, nas vizinhaças do ponto ótimo local, por analogia, o ideal é que $D^k \approx (\nabla^2 f(x^k))^{-1}$.

A idéia central do método de Newton é, localmente, aproximar a função objetivo por uma função quadrática, minimizando-a e tomando seu minimizador como novo ponto.

- Se o ponto x^k é próximo do mínimo local x^* ou se a função a ser minimizada $f_0(x)$ pode ser bem aproximada por uma função quadrática em x^k temos que:

$$\begin{aligned} f_0(x) &= f_0(x^k) + \nabla f_0(x^k)^T(x - x^k) + \frac{1}{2}(x - x^k)^T \nabla^2 f_0(x^k)(x - x^k) \\ &+ o(\|x - x^k\|^2) \\ &\approx f_0(x^k) + \nabla f_0(x^k)^T(x - x^k) + \frac{1}{2}(x - x^k)^T \nabla^2 f_0(x^k)(x - x^k) \end{aligned}$$

cujo mínimo ocorre quando $\nabla f_0(x^k) + \nabla^2 f_0(x^k)(x - x^k) = 0$.

Iteração típica

$$x^{k+1} = x^k - (\nabla^2 f_0(x^k))^{-1} \nabla f_0(x^k)$$

- Para que o método seja definido, $\nabla^2 f_0(x^k)$ deve admitir inversa.
- Caso $\nabla^2 f_0(x^k) \succ 0$, $d^k = -(\nabla^2 f_0(x^k))^{-1} \nabla f_0(x^k)$ é uma direção de descida.

Observe para qualquer $D \succ 0$, $d = -D\nabla f_0(x^k)$ satisfaz:
$d^T \nabla f_0(x^k) = \frac{df_0(x^k + \alpha d)}{d\alpha} = -\nabla f_0(x^k)^T D \nabla f_0(x^k) < 0.$

Desejamos resolver o sistema não linear $\nabla f_0(x) = 0$ para obter um ponto estacionário de f_0.

Usando a aproximação de primeira ordem para $\nabla f_0(x^k)$ nas vizinhanças de x^k temos:

$$\nabla f_0(x^k + d^k) = \nabla f_0(x^k) + \nabla^2 f_0(x^k) d^k + o(\|d^k\|)$$
$$\nabla f_0(x^k + d^k) \approx \nabla f_0(x^k) + \nabla^2 f_0(x^k) d^k$$

Então impondo a condição desejada $\nabla f_0(x^k + d^k) = 0$, temos:

$$d^k = -(\nabla^2 f_0(x^k))^{-1} \nabla f_0(x^k)$$

$$x^{k+1} = x^k + d_k$$

Supondo que $\nabla^2 f_0(x^2)^{-1} \succ 0$ existe.....

- Definindo $S = (D^k)^{\frac{1}{2}} = ((\nabla^2 f_0(x^k))^{-1})^{\frac{1}{2}}$, $x = Sy$, $\min f_0(x) = \min f_0(Sy) = \min h(y)$
- Fazer uma iteração

$$y^{k+1} = y^k - \alpha^k \nabla h(y^k)$$

- equivale a fazer uma iteração

$$x^{k+1} = x^k - \alpha^k (\nabla^2 f_0(x^k))^{-1} \nabla f_0(x^k)$$

Método de Newton pode ser visto como o método realizado na direção de $d^k = -\nabla h(y^k)$ no espaço $y = S^{-1}x$ com passo $\alpha^k = 1, k = 0, 1, \ldots$

- Não é globalmente convergente.
- Convergência quadrática para x^*, caso x^0 seja suficientemente próximo de x^*, ponto estacionário.
- Costuma ser implementado com busca unidirecional, caso o passo de Newton não promova redução de f_0.

$$x^{k+1} = x^k - \alpha^k (\nabla^2 f_0(x^k))^{-1} \nabla f_0(x^k)$$

- Visando redução do custo computacional, algumas estratégias utilizam a mesma matriz $\nabla^2 f_0(x^p)$ para algumas iterações $p, p+1, \ldots, z$. Entretanto, não há unanimidade em relação à efetividade desta abordagem.

A busca unidirecional na direção de Newton $(\pm\nabla^2 f_0(x^k)^{-1}\nabla f_0(x^k))$ pode falhar.

- Considere $f_0(x) = x_1^4 + x_1x_2 + (1+x_2)^2$ e o ponto $x^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.
- $\nabla f_0(x) = \begin{pmatrix} 4x_1^3 + x_2 \\ 2(1+x_2) + x_1 \end{pmatrix} \rightarrow \nabla f_0(0) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$
- $\nabla^2 f_0(x) = \begin{pmatrix} 12x_1^2 & 1 \\ 1 & 2 \end{pmatrix} \rightarrow \nabla^2 f_0(0) = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$ é indefinida.
- A direção de Newton é: $d^0 = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$
- Introduzindo a busca unidirecional $x^1 = x^0 \pm \alpha d^0$ não resolve pois d^0 só altera x_1 e para $x_2 = 0$, o mínimo de $f(x^0 \pm \alpha d^0)$ ocorre em x^0. Ou seja, $\alpha^* = 0$ e o método pára pois $\nabla f_0(x^0) d^0 = 0$.

$\nabla^2 f_0(x^k) \not\succ 0$

- A direção de Newton $d^k = -\nabla^2 f_0(x^k)^{-1} \nabla f_0(x^k)$ pode não ser de descida.
- Perturba-se a matriz Hessiana por meio de $E^k \succ 0$ de forma a obter $B^k = E^k + \nabla^2 f_0(x^k)^{-1} \succ 0$ e utiliza-se a direção modificada $d_m^k = -(B^k)^{-1} \nabla f_0(x^k)$, que é de descida.
- Esta modificação gera um método globalmente convergente, mas a convergência quadrática é perdida se o ponto estacionário x^* é singular. Só é possível garantir convergência linear, neste caso.

Deseja-se que a alteração não seja muito grande para preservar a informação de segunda ordem tanto quanto se possa.

- Modificação na fatoração de Cholesky de $\nabla^2 f_0(x^k)$ *on-the-fly*, corrigindo o termo na diagonal da fatoração $A = LDL^T$.
- Se dispomos de $\nabla^2 f_0(x^k) = Q \Lambda Q^T$, substitui-se um autovalor λ_i negativo por $\delta \geq \sqrt{u}$ onde u é a precisão da máquina.

Método de Newton Puro

Teorema

Vamos assumir que (1) $f_0 \in C^2$, (2) $\nabla^2 f_0(x)$ satisfaça uma condição Lipschitz $\|\nabla^2 f_0(x) - \nabla^2 f_0(y)\| \leq \lambda \|x - y\|$ nas vizinhanças de um minimizador local x^* de f_0. Então se x^k é suficientemente próximo a x^* para algum k e se $\nabla^2 f_0(x^*) \succ 0$, então o Método de Newton é bem definido para todo k e converge em segunda ordem (possui convergência quadrática).

Prova

- Vamos definir a função erro em k como $e^k = x^k - x^*$.
- Pela continuidade de f_0 temos que para qualquer vetor $e \in \mathbb{R}^n$

$$\nabla f_0(x^k + e) = \nabla f_0(x^k) + \nabla^2 f_0(x^k)e + O(\|e\|^2)$$

- Em particular para $e = -e^k = x^* - x^k$ temos:

$$0 = \nabla f_0(x^*) = \nabla f_0(x^k) - \nabla^2 f_0(x^k)e^k + O(\|e^k\|^2) \qquad (13)$$

Seja x^k um ponto em uma vizinhança de x^* onde $(\nabla^2 f_0(x^k))^{-1}$ exista e seja limitada superiormente. Esta vizinhança em torno de x^* existe pela continuidade de $\nabla^2 f_0(x)$, já que assumimos $\nabla^2 f_0(x^*) \succ 0$.

Prova - continua

- Multiplicando (13) por $(\nabla^2 f_0(x^k))^{-1}$ temos:

$$
\begin{aligned}
0 &= (\nabla^2 f_0(x^k))^{-1}\nabla f_0(x^k) - e^k + (\nabla^2 f_0(x^k))^{-1} O(\|e^k\|^2) \\
0 &= -d^k - e^k + O(\|e^k\|^2) \\
0 &= -d^k - x^k + x^* + O(\|e^k\|^2) \\
0 &= -e^{k+1} + O(\|e^k\|^2)
\end{aligned}
$$

$$e^{k+1} = O(\|e^k\|^2)$$

Pela definição da notação $O(\cdot)$, existe uma constante $c > 0$ tal que

$$\|e^{k+1}\| \leq c\|e^k\|^2$$

Prova - continua

$$\|e^{k+1}\| \leq c\|e^k\|^2$$

- Se x^k for suficientemente próximo de x^*, tal que $\|e^k\| \leq \frac{\alpha}{c}$ onde $0 < \alpha < 1$ então temos que

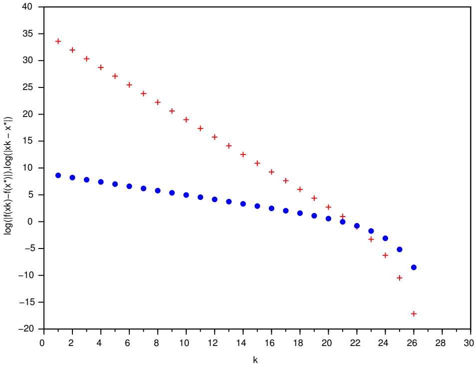
$$\|e^{k+1}\| \leq c\|e^k\|\frac{\alpha}{c} = \alpha\|e^k\|$$

- Consequentemente, x^{k+1} também está próximo o suficiente de x^*, $\nabla^2 f_0(x^{k+1}) \succ 0$....
- Por indução, as iterações estão bem definidas e $\|e^k\| \to 0$ a uma ordem quadrática.

$$f_0(x) = x_1^4 + x_1x_2 + (1 + x_2)^2$$

$x^0 = \begin{pmatrix} 6653.811 \\ 6283.9179 \end{pmatrix}$, curva vermelha $|f(x^k) - f(x^*)|$, azul $\|x^k - x^*\|_2$

Fase linear x fase quadrática



Mesmo quando $\nabla f_0^2(x) \succ 0$ para todo $x \in \mathbb{R}^n$, o passo de Newton ($\alpha^k = 1$) pode promover um deslocamento muito grande, fazendo com que:

$$f_0(x^k - (\nabla^2 f_0(x^k))^{-1} \nabla f_0(x^k)) > f_0(x^k).$$

- Nestes casos, a busca unidirecional com backtracking é introduzida e o comportamento do método é similar à figura ilustrada anteriormente, com duas fases distintas:
  - Damped (o passo de Newton raramente é aceito, pois é grande demais)
  - Quadrática (o passo de Newton puro é aceito, pois a aproximação quadrática já é muito boa).

- Foram motivados pelo desejo de acelerar o Método do Gradiente, sem incorrer nos custos computacionais do Método de Newton e na necessidade de informação de segunda ordem.

- Foram propostos originalmente para resolver o problema quadrático

$$\min f_0(x) = \frac{1}{2}x^T Qx - b^T x : x \in \mathbb{R}^n$$

onde $Q \in \mathcal{S}^n_{++}$.

- Resolvem as condições necessárias e suficientes de otimalidade deste problema: $\nabla f_0(x) = 0 \to Qx = b$.

- Podem ser utilizados para resolver $Ax = b, A \notin \mathcal{S}^n_{++}$ (mas A^{-1} existe) após a transformação $A^T Ax = A^T b$.

- Resolvem o problema quadrático em no máximo n iterações, mas são considerados métodos iterativos.
- Para um problema quadrático de grande escala com bem menos de n iterações obtém-se razoável aproximação para $x : Qx = b$.
- São bastante empregados para resolver problemas de programação não linear irrestritos mais gerais. Nestes casos, perdem a terminação finita observada no caso quadrático.
- A implementação de uma de suas variantes mais conhecidas (Método do Gradiente Conjugado) é apenas um pouco mais complexa que a implementação de um método do gradiente puro.

Definição

Dada uma matriz $Q \in \mathcal{S}^n$, dizemos que os vetores $d^1, d^2, \ldots, d^n$ são Q-conjugados se

$$(d^i)^T Q d^j = 0, \forall i \neq j$$

Alguns textos definem que $d^1, d^2, \ldots, d^n$ são Q-conjugadas se forem linearmente independentes e $(d^i)^T Q d^j = 0, \forall i \neq j$.

Conjugacidade implica em independência linear se $Q \in \mathcal{S}^n_{++}$

Se as direções $d^1, d^2, \ldots, d^n$ são Q conjugadas e $Q \in \mathcal{S}^n_{++}$, então $d^1, d^2, \ldots, d^n$ são linearmente independentes

(assuma o oposto, escreva uma direção como combinação linear das demais, multiplique por $(d^k)^T Q$, obtendo uma contradição, já que $(d^k)^T Q d^k > 0$)

Considerando o problema quadrático estritamente convexo:

$$\min f_0(x) = \frac{1}{2}x^T Qx - b^T x$$
$$x \in \mathbb{R}^n$$

e as direções $d^0, d^1, \ldots, d^{n-1}$, Q conjugadas para $Q \succ 0$. Como são linearmente independentes, a solução x^* de $Qx = b$ pode ser escrita da seguinte forma:

$$x^* = \sum_{i=0}^{n-1} \alpha_i d^i$$
$$(d^j)^T Qx^* = \sum_{i=0}^{n-1} \alpha_i (d^j)^T Qd^i$$
$$\rightarrow \alpha_j = \frac{(d^j)^T Qx^*}{(d^j)^T Qd^j} = \frac{(d^j)^T b}{(d^j)^T Qd^j} \tag{14}$$

Então podemos escrever

$$x^* = \sum_{i=0}^{n-1} \left(\frac{(d^i)^T b}{(d^i)^T Q d^i} \right) d^i$$

1. Tomando o produto interno com uma direção convenientemente escolhida ($(d^j)^T Q$) apenas um termo não nulo nos permitiu calcular o coeficiente de cada d^i na combinação linear.
2. Ao usar a $Q-$ortogonalidade (e não a I-ortogonalidade) fomos capazes de escrever a expressão acima sem usar o vetor x^* desconhecido, já que Qx^* foi substituído por b.

$$x^* = \sum_{i=0}^{n-1} \left(\frac{(d^i)^T b}{(d^i)^T Q d^i} \right) d^i$$

A ideia central do método

A expressão acima mostra que, disponíveis n direções $Q-$ conjugadas, o problema quadrático pode ser resolvido sequencialmente, adicionando-se uma direção conjugada por vez, com o respectivo cálculo de seu passo α.

Exemplo

$$\min f_0(x) = 4x_1^2 + 4x_2^2 - 4x_1x_2 - 12x_2 = \frac{1}{2}(8x_1^2 + 8x_2^2 - 8x_1x_2) - 12x_2$$

$$\nabla f_0(x) = \begin{pmatrix} 8x_1 + 4x_2 \\ 8x_2 + 4x_1 - 12 \end{pmatrix}, Q = \begin{pmatrix} 8 & -4 \\ -4 & 8 \end{pmatrix}, b = \begin{pmatrix} 0 \\ 12 \end{pmatrix}$$

- Primeira direção (arbitrariamente escolhida): $d^0 = (1 \quad 0)^T$.
- Segunda direção:

  $$(d^1)^T Q d^2 = 0 \rightarrow (1 \quad 0) \begin{pmatrix} 8 & -4 \\ -4 & 8 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 0 \rightarrow 8a = 4b.$$

  Escolhemos $d^2 = (1 \quad 2)^T$.
- Ponto inicial arbitrário: $x^0 = (-\frac{1}{2} \quad 1)^T$.

Primeira busca unidirecional

$$f_0(x^0 + \alpha d^0) = \frac{1}{2}\left(8(-\frac{1}{2} + \alpha)^2 + 8(1 + 0\alpha)^2 - 8(-\frac{1}{2} + \alpha)(1 + 0\alpha)\right)$$
$$- 12(1 + 0\alpha)$$
$$= \frac{1}{2}(8\alpha^2 - 16\alpha + 14) - 12$$

Resolvendo de forma exata para α:
$\alpha^0 = 1$ e $x^1 = (-\frac{1}{2} \quad 1) + (1 \quad 0)^T = (\frac{1}{2} \quad 1)^T$.

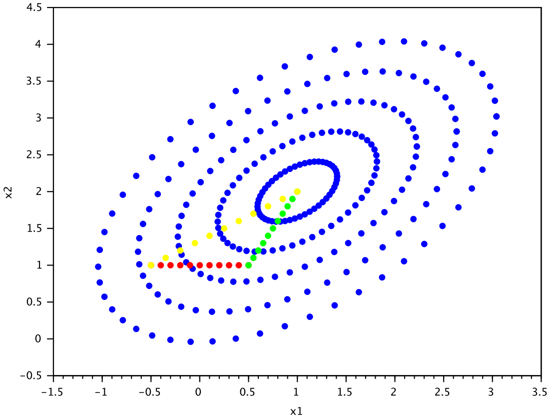
Segunda busca unidirecional

$$f_0(x^1 + \alpha d^1) = \frac{1}{2}\left(8(\frac{1}{2} + \alpha)^2 + 8(1 + 2\alpha)^2 - 8(\frac{1}{2} + \alpha)(1 + 2\alpha)\right)$$
$$- 12(\frac{1}{2} + 1\alpha)(1 + 2\alpha)$$
$$= \frac{1}{2}\left(24\alpha^2 + 24\alpha + 6\right) - 12 - 24\alpha$$

Resolvendo de forma exata para α:
$\alpha^1 = \frac{1}{2}$ e $x^2 = (\frac{1}{2} \quad 1)^T + \frac{1}{2}(1 \quad 2)^T = (1 \quad 2)^T$.

Solução exata obtida em 2 iterações.

```
-->Q,b
 Q  =
    8.  - 4.
  - 4.    8.
 b  =
     0.
    12.
 -->(inv(Q)*b)'
 ans  =
     1.    2.
```



Direção de Newton, Direções conjugadas: d^0 e d^1.

$Q \in \mathcal{S}^n_{++}$

Mediante uma transformação de coordenadas, a otimização sobre direções Q-conjugadas pode ser entendida como a minimização ao longo de direções ortogonais.

1. Definindo $y = Q^{\frac{1}{2}}x$, $\min_{x \in \mathbb{R}^n} \frac{1}{2}x^T Qx = \min_{y \in \mathbb{R}^n} \frac{1}{2}y^T y$.
2. Se $\{w^0, w^1, \ldots, w^{n-1}\}$ são n direções ortogonais, a iteração típica $y^{k+1} = y^k + \alpha^k w^k$, com $\alpha^k = \arg\min_\alpha \|y^k + \alpha w^k\|_2^2$ encontra o mínimo $y^* = 0$ em no máximo n iterações.
3. Voltando para o espaço x, multiplicando $y^{k+1} = y^k + \alpha^k w^k$ por $Q^{-\frac{1}{2}}$, temos $x^{k+1} = x^k + \alpha^k d^k$ onde $d^k = Q^{-\frac{1}{2}} w^k$.
4. A ortogonalidade de w^i, w^j implica em $Q-$conjugacidade de $\{d^0, d^1, \ldots, d^{n-1}\}$, mediante a transformação acima:

$$0 = (w^i)^T w^j = (d^i)^T Q^{\frac{1}{2}} Q^{\frac{1}{2}} d^j = (d^i)^T Q d^j.$$

Vamos fazer o desenvolvimento do método para

$$\min f_0(x) = \frac{1}{2}x^T Qx - b^T x$$
$$x \in \mathbb{R}^n$$

considerando para isto, as direções $d^0, d^1, \dots, d^{n-1}$, Q conjugadas (ainda não discutimos como são obtidas).

Iteração típica

$$x^{k+1} = x^k + \alpha d^k, k = 0, 1, \dots, n-1,$$

onde $x^0 \in \mathbb{R}^n$ é um ponto inicial arbitrário e α^k é obtido de forma exata, isto é:

$$\alpha^k = \arg\min_\alpha f_0(x^k + \alpha d^k).$$

- Impondo as condições necessárias e suficientes para que o ponto $x^k + \alpha d^k$ minimize a função estritamente convexa de uma variável α, $f_0(x^k + \alpha d^k)$ ao longo da linha, temos:

$$
\begin{aligned}
0 &= \frac{df_0(x^k + \alpha d^k)}{d\alpha} \\
&= (d^k)^T \nabla f_0(x^k + \alpha d^k) \\
&= (d^k)^T (Q(x^k + \alpha d^k) - b))
\end{aligned}
$$

- Isolando α, o valor ótimo é dado por:

$$\alpha = \frac{(d^k)^T (b - Qx^k)}{(d^k)^T Q d^k}$$

observe que se $x^0 = 0$ temos que $\alpha^0 = \frac{(d^0)^T b}{(d^0)^T Q d^0}$ e há equivalência plena entre a expressão acima e (14).

Um resultado teórico muito importante

Proposição

Considere o problema de otimização quadrático irrestrito, definido por uma matriz $Q \in \mathcal{S}^n_{++}$ e um método de direções $Q-$conjugadas $d^0, \ldots, d^{n-1}$, operando sobre a iteração típica:

$$x^{k+1} = x^k + \alpha d^k, \text{ onde } \alpha = \frac{(d^k)^T(b - Qx^k)}{(d^k)^T Q d^k} \tag{15}$$

isto é, α minimiza de forma exata $f_0(x^k + \alpha d^k)$. Então:

1. $x^{k+1} = \arg\min\{f_0(x) : \ x \in \mathcal{M}^k\}$, onde $\mathcal{M}^k$ é o conjunto afim

$$\mathcal{M}^k = \{x \in \mathbb{R}^n : x = x^0 + \operatorname{span}\{d^0, d^1, \ldots, d^k\}\}.$$

2. Em particular, x^n minimiza f_0 sobre $\mathcal{M}^{n-1} = \mathbb{R}^n$.

Prova

Parte 1

- *Uma vez que α^i é obtido por minimização exata ao longo de d^i:*

$$0 = \left.\frac{df_0(x^i + \alpha d^i)}{d\alpha}\right|_{\alpha=\alpha^i} = \nabla f_0(x^{i+1})^T d^i, \quad \forall i = 0, 1, \ldots, n-1$$

Prova

(continua) Por outro lado, para $i = 0, \ldots, k-1$ temos:

$$\begin{aligned}
\nabla f_0(x^{k+1})^T d^i &= (Qx^{k+1} - b)^T d^i \\
&= (x^{i+1} + \sum_{j=i+1}^{k} \alpha^j d^j)^T Q d^i - b^T d^i \\
&= (x^{i+1})^T Q d^i - b^T d^i \\
&= \nabla f_0(x^{i+1})^T d^i \\
&= 0
\end{aligned}$$

No desenvolvimento acima, usamos os seguites argumentos, em ordem:

- *a $Q-$conjugacidade das direções d^j, d^i para $j = i+1, \ldots, k$.*
- *o resultado obtido na primeira parte.*

Então $(\nabla f_0(x^{k+1}))^T d^i = 0, \quad \forall i = 0, \ldots, k \rightarrow \nabla f_0(x^{k+1}) \perp \mathcal{M}^k$.

Observações

- O método funciona de acordo com $x^{k+1} = x^0 + \sum_{j=0}^{k} \alpha^j d^j$ para $k = 0, \ldots, n-1$, onde $\alpha^j = \frac{(d^j)^T(b - Qx^j)}{(d^j)^T Q d^j}, \forall j = 0, \ldots, n-1$.
- $x^{k+1} = \arg\min f_0(x), \quad x \in \mathcal{M}^k$, equivale a $\nabla f_0(x^{k+1}) \perp \mathcal{M}^k$, dada a convexidade estrita de f_0 sobre M^k. Isto é:

$$\left.\frac{\partial f_0(x^0 + \sum_{j=0}^{k} \gamma^j d^j)}{\partial \gamma^i}\right|_{\gamma^j = \alpha^j,\ j=0,\ldots,k} = 0 \text{ para todo } i \le k$$

Logo, para um dado $i \le k$:

$$\left.\frac{\partial f_0(x^0 + \sum_{j=0}^{k} \gamma^j d^j)}{\partial \gamma^i}\right|_{\gamma^j = \alpha^j, j=0,\ldots,k} = \nabla f_0(x^0 + \sum_{j=0}^{k} \alpha^j d^j)^T d^i = 0$$

Definindo

$$E(x) = \frac{1}{2}(x - x^*)^T Q(x - x^*),$$

temos

$$E(x) = f_0(x) + \frac{1}{2}(x^*)^T Qx^*.$$

- Por um lado, $E(x)$ mede o quão próximo x encontra-se da solução ótima x^* em uma distância generalizada, definida por Q.
- Se $Q = I$, a distância é a distância Euclideana. Se $Q \neq I$, $Q \succ 0$, $E(x)$ é uma distância generalizada.
- Por outro lado, dado que $\frac{1}{2}(x^*)^T Qx^*$ independe de x, minimizar $E(x)$ equivale a minimizar $f_0(x)$, ou seja, podemos interpretar $E(x)$ como a função a ser minimizada pelo procedimento.
- As taxas de convergência do método podem ser medidas em relação a $E(x)$.

- Ainda não mencionamos como gerar as direções conjugadas.
- Uma maneira eficiente (mas pouco estável) de obter estas direções consiste em adaptar o Procedimento de Gram-Schmidt para gerar uma base ortonormal para um subespaço linear.

Dado um conjunto de direções $\xi^0, \ldots, \xi^k$ linearmente independentes, vamos gerar direções $d^0, \ldots, d^k$ $Q-$conjugadas iterativamente, uma direção por vez, a partir de cada direção linearmente independente, utilizando a ideia que

$$\text{span}\{\xi^0, \ldots, \xi^k\} = \text{span}\{d^0, \ldots, d^k\}$$

e impondo a $Q-$conjugacidade em cada etapa.

1. Fazemos $\xi^0 = d^0$
2. (Indução) Suponha que para $i < k$ tenhamos selecionado direções $d^0, \ldots, d^i$, $Q-$conjugadas
3. Então calculamos uma nova direção conjugada escrevendo

$$d^{i+1} = \xi^{i+1} + \sum_{m=0}^{i} c^{i+1,m} d^m \qquad (16)$$

   e escolhendo os coeficientes $c^{i+1,m}$ de forma a garantir que $(d^j)^T Q d^{i+1} = 0, j = 0, \ldots, i$.
4. Para cada $j = 0, \ldots, i$, mutiplicamos (16) por $(d^j)^T Q$ e impomos $(d^j)^T Q d^{i+1} = 0$
5. Obtemos os coeficientes $c^{i+1,j} = -\frac{(d^j)^T Q \xi^{i+1}}{(d^j)^T Q d^j}, \;\; j = 0, \ldots, i$

- Direções linearmente independentes (a ser provado) empregadas em Gram-Schmidt:

$$\xi^0 = -\nabla f(x^0), \ldots, \xi^{n-1} = -\nabla f_0(x^{n-1})$$

- Iteração típica: $x^{k+1} = x^k + \alpha^k d^k$, onde α^k é obtido via minimização exata, sendo dado por:

$$\alpha^k = -\frac{(d^k)^T \nabla f_0(x^k)}{(d^k)^T Q d^k}$$

- E as direções conjugadas são dadas por:

$$d^0 = -\nabla f(x^0), \quad d^k = -\nabla f_0(x^k) + \sum_{j=0}^{k-1} \left(\frac{(d^j)^T Q \nabla f_0(x^k)}{(d^j)^T Q d^j} \right) d^j$$

poderão ser simplificadas, levando a um método mais eficiente, graças à propriedade de otimização sobre $\mathcal{M}^k$.

- Provar que as direções $\xi^k = -\nabla f_0(x^k)$ usadas no método são linearmente independentes.

- Usar a propriedade de otimização sobre $\mathcal{M}^k$ para simplificar o cálculo da direção, a cada iteração.

Independência linear dos gradientes nos pontos gerados pelo Método do Gradiente Conjugado

Proposição

Assuma que $\{x^0, \ldots, x^{n-1}\}$ *seja a sequência de* n *vetores gerados pelo Método do Gradiente Conjugado. Então temos que:*

1. $\nabla f_0(x^{k+1}) \perp span\{d^0, \ldots, d^i\} : i = 0, \ldots, k$ *(já provado)*
2. $\nabla(f_0(x^{k+1}))^T \nabla f_0(x^i) = 0, \;\; i = 0, \ldots, k$ *e portanto os gradientes nos pontos gerados pelo método são linearmente independentes.*

Independência linear dos gradientes nos pontos gerados pelo Método do Gradiente Conjugado

Definimos $g^k = \nabla f_0(x^k) = Qx^k - b$. A prova é por indução.

Prova

- *(Caso base) Vamos assumir que* $g^0 \neq 0$*, caso contrário o método terminou. Assim sendo,* g^0 *é li.*
- *(Hip. Indução) Assumindo que após* k *passos o método não terminou, isto é,* $g^k \neq 0$*, temos que* $g^0, \ldots, g^{k-1}$ *são li.*
- *O que ocorre então com* $(g^k)^T g^j : 0 \leq j \leq k-1$ *?*
  - *Então temos* $g^k \neq 0$*, cc o método terminou.*
  - *Recorde que:* $0 = \nabla f_0(x^k)^T d^0 = \cdots = \nabla f_0(x^k)^T d^{k-1}$*, isto é,* $g^k = \nabla f_0(x^k)$ *é ortogonal às direções conjugadas.*
  - *Logo* $g^k \perp span\{d^0, \ldots, d^{k-1}\}$ *e uma vez que, por contrução, no Método de Gram-Schmidt* $span\{g^0, \ldots, g^{k-1}\} = span\{d^0, \ldots, d^{k-1}\}$*, temos que* $g^k \perp span\{g^0, \ldots, g^{k-1}\}$*.*

Proposição

Sobre o método do Gradiente Conjugado, pode-se dizer:

- *As suas direções podem ser calculadas da seguinte forma:*

$$d^0 = -g^0, \dots$$

$$d^k = -g^k + \beta^k d^{k-1}, k = 1, \dots, n-1,$$

  onde $\beta^k = \frac{(g^k)^T g^k}{(g^{k-1})^T g^{k-1}}$. *Veja que apenas um coeficiente é necessário para definir a direção.*

- *O método termina com uma solução ótima do Problema Quadrático, em no máximo n passos.*

Prova

Parte 1 - vamos primeiro mostrar que o método (sem a simplificação do cálculo da direção) converge em no máximo n iterações, isto é, que em no máximo n iterações teremos $g^k = 0$.

Obviamente $g^k \neq 0, k < n$, caso contrário nada temos a provar.

Então, vamos assumir que $g^n \neq 0$ e gerar uma contradição.

- *Pela independência linear de $\{g^0, \ldots, g^{n-1}\}$ temos que $\mathbb{R}^n = span\{g^0, \ldots, g^{n-1}\}$.*
- *Pelo resultado anterior, teríamos que: $g^n \perp span\{g^0, \ldots, g^{n-1}\} = \mathbb{R}^n$.*
- *Como não podemos gerar mais de n direções li, temos que $g^n = 0$.*

Prova

- *Observe que para todo* $j : g^j \neq 0$ *temos:* $g^{j+1} - g^j = Q(x^{j+1} - x^j) = \alpha^j Q d^j$, *onde* $\alpha^j \neq 0$, *uma vez que* $g^{j+1} \perp span\{g^0, \ldots, g^j\}$ *(se* $\alpha^j = 0$, *teríamos* $g^{j+1} = g^j = 0$*).*
- *Então temos:*

$$(g^i)^T Q d^j = \frac{1}{\alpha^j}(g^i)^T(g^{j+1} - g^j) = \begin{cases} 0 & se \quad j = 0, \ldots, i-2 \\ \frac{1}{\alpha^j}(g^i)^T g^i & se \quad j = i-1 \end{cases}$$

$$(d^j)^T Q d^j = \frac{1}{\alpha^j}(d^j)^T(g^{j+1} - g^j)$$

- *Substutindo na expressão de* d^k *de Gram-Schmidt temos* $d^k = -g^k + \beta^k d^{k-1}$, *onde* $\beta^k = \frac{(g^k)^T g^k}{(d^{k-1})^T (g^k - g^{k-1})}$.
- *Usando* $d^{k-1} = -g^{k-1} + \beta^{k-1} d^{k-2}$, *a ortogonalidade entre* g^k *e* g^{k-1} *e entre* d^{k-2} *e* $(g^k - g^{k-1})$, *o denominador pode ser escrito como desejamos.*

Método do Gradiente Conjugado

Iteração típica: $x^{k+1} = x^k + \alpha^k d^k$ para $k = 0, 1, \ldots, n-1$

- $\alpha^k = -\frac{(g^k)^T d_k}{(d^k)^T Q d^k}$,
- $d^0 = -g^0$, $d^k = -g^k + \beta^k d^{k-1}$
- $\beta^k = \frac{(g^k)^T g^k}{(g^{k-1})^T g^{k-1}}$
- $g^k = \nabla f_0(x^k) = Qx^k$ - b.

Proposição

Se o método não termina na iteração k, as iterações $k = 0, 1, \ldots, n-1$ do Gradiente Conjugado satisfazem:

1. $\text{span}\{d^0, d^1, \ldots, d^k\} = \text{span}\{g^0, g^1, \ldots, g^k\}$ (por construção).
2. $\text{span}\{g^0, g^1, \ldots, g^k\} = \text{span}\{g^0, Qg^0, \ldots, Q^k g^0\}$
3. $\text{span}\{d^0, d^1, \ldots, d^k\} = \text{span}\{g^0, Qg^0, \ldots, Q^k g^0\}$

Prova: por indução

- Caso base $k = 0$ é verificado. (HI) Vamos assumir que (1)-(3) valem para k. Para $k+1$ temos:

$$
\begin{aligned}
d^{k+1} &= -g^{k+1} + \beta^{k+1} d^k \\
x^{k+1} &= x^k + \alpha^k d^k \\
Qx^{k+1} &= Qx^k + \alpha^k Qd^k \\
Qx^{k+1} - b &= Qx^k - b + \alpha^k Qd^k \\
g^{k+1} &= g^k + \alpha^k Qd^k
\end{aligned}
$$

- HI para (2) $\rightarrow g^{k+1} \in \text{span}\{g^0, Qg^0, \dots, Q^{k+1}g^0\}$ e $g^{k+1} \perp \mathcal{M}^k \rightarrow g^{k+1} \notin \text{span}\{g^0, Qg^0, \dots, Q^k g^0\}$
- HI para (3), $\rightarrow Qd^k \in \text{span}\{g^0, Qg^0, \dots, Q^{k+1}g^0\}$.
- HI para (3) e com os dois resultados acima, $d^{k+1} \in \text{span}\{g^0, Qg^0, \dots, Q^{k+1}g^0\}$.

decorrentes de que

- $\{\mathcal{M}^k : k = 0, 1, \ldots, k\}$ são os conjuntos gerados pela translação de x^0 por meio dos subespaços $\text{span}\{g^0, Qg^0, \ldots, Q^k g^0\}$.
- x^{k+1} minimiza $f_0(x)$ sobre $\mathcal{M}^k$.
- Minimizar $E(x)$ equivale a minimizar $f_0(x)$.

...A cada iteração do método, os conjuntos $\mathcal{M}^k$ sobre os quais a minimização é realizada são gerados pela introdução de uma potência adicional de Q por g^0.

$\Rightarrow$ **Isto nos permite pensar em um algoritmo que funcione sobre a seguinte iteração típica....**

Um algoritmo funcionando sobre as iterações

$$x^{k+1} = x^0 + P^k(Q)g^0, \;\; k = 0, 1, 2, \ldots$$

- onde $P^k(Q)$ é um polinômio de grau k na matriz Q.
- Uma escolha dos coeficientes do polinômio $P^k(Q) = \sum_{i=0}^{k} \gamma_i Q^i$ determina a sequência dos pontos x^{k+1} deste método.

Para este método hipotético, teríamos:

$$\begin{aligned} x^{k+1} &= x^0 + P^k(Q)g^0 \\ x^{k+1} - x^* &= x^0 - x^* + P^k(Q)(Qx^0 - b) - P^k(Q)(Qx^* - b) \\ &= x^0 - x^* + P^k(Q)Q(x^0 - x^*) \\ &= (I + QP^k(Q))(x^0 - x^*) \end{aligned}$$

Logo temos

$$
\begin{aligned}
x^{k+1} - x^* &= (I + QP^k(Q))(x^0 - x^*) \\
E(x^{k+1}) &= \frac{1}{2}(x^{k+1} - x^*)^T Q(x^{k+1} - x^*) \\
&= \frac{1}{2}(x^0 - x^*)^T Q(I + QP^k(Q))^2(x^0 - x^*)
\end{aligned}
$$

Questão:

Considerando os polinômios $P^k(Q) = \sum_{i=0}^{k} \gamma^i Q^i$ de grau k em Q, qual polinômio minimiza $E(x^{k+1})$? Ou seja, qual a escolha de coeficientes $\gamma^0, \ldots, \gamma^k$ que minimizam $E(x^{k+1})$?

Expandindo $x^{k+1} = x^0 + P^k(Q)g^0$ em função dos coeficientes do polinômio:

$$x^{k+1} = x^0 + \gamma^0 g^0 + \gamma^1 Q g^0 + \cdots + \gamma^k Q^k g^0$$

e comparando com a expansão

$$x^{k+1} = x^0 + \alpha^0 d^0 + \alpha^1 d^1 \cdots + \alpha^k d^k$$

gerada pelo Método do Gradiente Conjugado (que possui a mesma forma) e diante de:

1. $\text{span}\{d^0, d^1, \ldots, d^k\} = \text{span}\{g^0, Qg^0, \ldots, Q^k g^0\}$
2. $x^{k+1} = \arg\min_{x \in \mathcal{M}^k} f_0(x)$ para qualquer método de Direções Conjugadas

temos o seguinte resultado...

Teorema

Embora a relação entre os α's, β's que calculamos para o Método do Gradiente Conjugado e os γ's dos polinômios ótimos não estejam explícitos, o Método do Gradiente Conjugado está implicitamente calculando este polinômio ótimo, uma vez que a cada iteração está minimizando $E(x^{k+1})$, ou seja:

$$E(x^{k+1}) = \min \left\{ \frac{1}{2}(x^0 - x^*)^T Q(I + Q \sum_{i=0}^{k} \gamma^i Q^i)^2 (x^0 - x^*) : \gamma \in \mathbb{R}^{k+1} \right\}$$

- Sejam $\{(\lambda_i, u^i) : i = 1, \dots, n\}$ os auto-pares (normalizados) de Q.
- Ou seja: $Q = U \Lambda U^T$ $(Q \in \mathcal{S}^n_{++},\ U^T U = U U^T = I)$
- Estes autovetores $\{u^1, \dots, u^n\}$ são linearmente independentes, o que nos permite expandir:

$$
\begin{aligned}
x^0 - x^* &= \sum_{i=1}^{n} \xi_i u^i \\
Q(x^0 - x^*) &= \sum_{i=1}^{n} \xi_i Q u^i \\
&= \sum_{i=1}^{n} \xi_i \lambda_i u^i
\end{aligned}
$$

..que nos permite escrever:

$$E(x^0) = \frac{1}{2}(x^0 - x^*)^T Q(x^0 - x^*) = \frac{1}{2}\sum_{i=1}^{n} \lambda_i \xi_i^2$$

e, dado que:

$$E(x^{k+1}) = \min_{P^k(Q)} \left\{ \frac{1}{2}(x^0 - x^*)^T Q(I + QP^k(Q))^2 (x^0 - x^*) \right\}$$

após alguma manipulação, temos que, para um polinômio de grau k qualquer:

$$\begin{aligned} E(x^{k+1}) &\leq \frac{1}{2}\sum_{i=1}^{n}(1 + \lambda_i P^k(\lambda_i))^2 \lambda_i \xi_i^2 \\ &\leq \max_{\lambda_i}\{(1 + \lambda_i P^k(\lambda_i))^2\} \frac{1}{2}\sum_{i=1}^{n} \lambda_i \xi_i^2 \\ &= \max_{\lambda_i}\{(1 + \lambda_i P^k(\lambda_i))^2\} E(x^0) \end{aligned}$$

Após alguma manipulação

$$
\begin{aligned}
E(x^{k+1}) &= \frac{1}{2}(x^0 - x^*)^T Q(I + QP^k(Q))^2(x^0 - x^*) \\
&= \frac{1}{2}\left(\sum_i \xi_i \lambda_i (u^i)^T\right)(UP(\Lambda)U^T)\left(\sum_i \xi_i u^i\right) \\
&= \frac{1}{2}\left(\sum_i \xi_i \lambda_i (\mathbf{e}^i)^T\right) P(\Lambda)\left(\sum_i \xi_i \mathbf{e}^i\right) \\
&= \frac{1}{2}\sum_i \xi_i^2 \lambda_i (1 + \lambda_i P^k(\lambda_i))^2 \\
&\leq \frac{1}{2}\max_{\lambda_i}\{(1 + \lambda_i P^k(\lambda_i))^2\} E(x^0)
\end{aligned}
$$

No desenvolvimento acima, usamos $(\mathbf{e}^i)^T P(\Lambda)\mathbf{e}^j = 0$ para $i \neq j$ e $(\mathbf{e}^i)^T P(\Lambda)\mathbf{e}^i = (1 + \lambda_i P^k(\lambda_i))^2$.

$$E(x^{k+1}) \leq \max_{\lambda_i}\{(1 + \lambda_i P^k(\lambda_i))^2\} E(x^0)$$

O que acontece se Q possuir apenas $m < n$ autovalores distintos ?

- Seria possível algum método que operasse sobre a iteração típica

$$x^{k+1} = x^k + P^k(Q) g^0,$$

  por exemplo o Método do Gradiente Conjugado que escolhe $P^k(Q)$ de forma ótima, tirar proveito disto ?

- *Tirar proveito* significa com poucas (digamos, m) iterações, reduzir muito o erro, idealmente reduzi-lo a zero, resolvendo o problema quadrático em menos de n iterações.

$$E(x^{k+1}) \leq \max_{\lambda_i}\{(1+\lambda_i P^k(\lambda_i))^2\}E(x^0)$$

1. Se o polinômio de grau m (em λ)

$$q(\lambda) = 1 + \lambda P^{m-1}(\lambda)$$

satisfizer $q(\lambda_i) = 0$ para $i = 1, \ldots, m$, então

$$E(x^m) \leq \max_{\lambda_i}\{(1+\lambda_i P^{m-1}(\lambda_i))^2\}E(x^0) = 0.$$

2. Ou seja, escolhendo convenientemente $P^{m-1}(Q)$ (equivalentemente, $P^{m-1}(\lambda)$), é possível obter $E(x^m) = 0$ para $m < n$.

O Método do Gradiente Conjugado, por fazer escolhas ótimas de $P^k(Q) : k = 0, 1, \ldots, n-1$, resolve o problema em não mais de $m < n$ iterações.

Métodos com reinicialização

Diante do último resultado, é natural considerarmos classes de procedimentos que realizam $m+1$ iterações do Método do Gradiente Conjugado e, em seguida, implementam uma reinicialização, isto é, ao invés de calcular $d^{m+1} = -g^{m+1} + \beta^{m+1} d^m$ faz-se $d^0 = g^{m+1}$ e o método começa um novo ciclo de adicionais m iterações do Método do Gradiente Conjugado.

Casos particulares de m:

- $m = 0$: Método do Gradiente Puro
- $m = n - 1$: Método do Gradiente Conjugado completo
- $0 < m < n - 1$: Método do Gradiente Conjugado parcial.

Notação empregada

- Função objetivo quadrática estritamente convexa
- x^k: denota o ponto obtido após a aplicação de k ciclos de $m+1$ iterações do Gradiente Conjugado (ou seja, não indexamos as iterações do Gradiente Conjugado, mas sim os ciclos de $m+1$ iterações).
- Então o método funciona sob a seguinte iteração típica:

$$x^{k+1} = x^k + P^{k,m}(Q)g^k$$

  onde $P^{k,m}(Q)$ é o polinômio de grau m gerado ao longo do $k-$ésimo ciclo de $m+1$ iterações do Gradiente Conjugado.
- Podemos selecionar $P^{k,m}(Q)$ de forma a minimizar $E(x^{k+1}) = \frac{1}{2}(x^{k+1} - x^*)^T Q(x^{k+1} - x^*)$, mas face aos resultados anteriores, ao invés de determinar os coeficientes explícitos de $P^{k,m}(Q)$, implementamos os $m+1$ passos do $k-$ésimo ciclo do Gradiente Conjugado.

Problema a resolver

$$\min f_0(x) = \frac{1}{2} x^T Q x - b^T x$$
$$c^T x = 0$$

é aproximado por uma sequência de problemas irrestritos, penalizados por $\mu > 0, \mu \to \infty$.

Problema irrestrito penalizado

$$\min \frac{1}{2} x^T Q x - b^T x + \mu (c^T x)^2$$

Observações:

1. Valores elevados de $\mu > 0$ tendem a evitar a violação da restrição $c^T x = 0$.
2. A parte quadrática do problema penalizado é $\frac{1}{2}x^T(Q + \mu cc^T)x$ e, portanto, é importante considerar o espectro de $(Q + \mu cc^T)$ para avaliação das propriedades de convergência do método empregado para resolver o problema irrestrito.
3. Se $\lambda_i(Q) \in [a, A]$, veremos que quando $\mu \to \infty$, um autovalor da matriz $(Q + \mu cc^T)$ tende a $+\infty$ enquanto os $n - 1$ demais permanecem limitados ao intervalo $[a, A]$.
4. Se o método do gradiente for empregado, a taxa de convergência será governada pela razão $\frac{\mu - a}{\mu + a} \to 1$.
5. Se o método do Gradiente Conjugado parcial for empregado com $m < n - 1$, o impacto deste autovalor μ não seria percebido na taxa de convergência (próximo Teorema).

Teorema (Luenberger, p. 275, Sec. 9.5, Bertsekas, p.144, Sec. 1.6)

Assuma que $Q \in \mathcal{S}^n_{++}$ possui $n - m$ autovalores no intervalo $[a, b]$ $(a > 0)$ e os demais m superiores a b.
Então, o Método do Gradiente Parcial, reinicializado a cada $m + 1$ iterações, satisfaz

$$E(x^{k+1}) \leq \left(\frac{b-a}{b+a}\right)^2 E(x^k),$$

onde x^{k+1} é obtido a partir de x^k realizando-se $m + 1$ passos completos do Gradiente Conjugado e então reinicializando.

Prova.

- Vamos assumir que os m autovalores de Q que excedem b são $\lambda_1, \ldots \lambda_m$ enquanto que os $n - m$ demais pertencem a $[a, b]$.
- Partindo de x^k e realizando $m + 1$ iterações do Gradiente Conjugado temos:

$$E(x^{k+1}) \leq \max_{\lambda_i : i=1,\ldots,n} \left\{ (1 + \lambda_i P^{k,m}(\lambda_i))^2 \right\} E(x^k)$$

- Observe que isto é válido para qualquer polinômio $P^{k,m}(\lambda)$ e não apenas para o polinômio ótimo implicitamente empregado pelo Método do Gradiente Conjugado.
- Vamos escolher $P^{k,m}$ tal que $q(\lambda) = 1 + \lambda P^{k,m}$ seja de grau $m + 1$ e possua as raizes:
  - $\lambda = \frac{a+b}{2}$
  - $\lambda = \lambda_i : i = 1, \ldots, m$.

Prova.

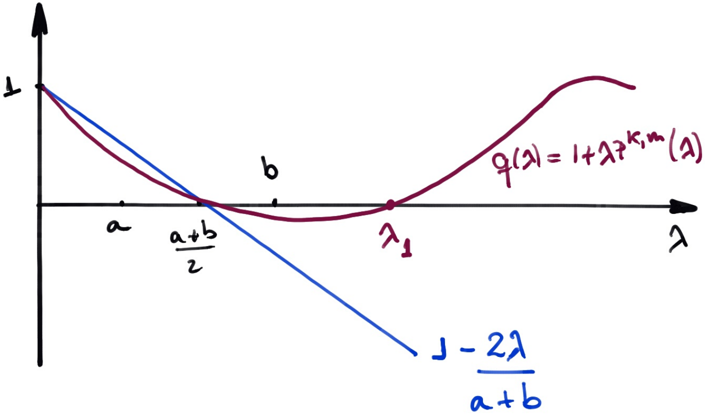
- Por exemplo, escolhemos tal que:

$$1 + \lambda P^{k,m}(\lambda) = \frac{2}{(a+b)\lambda_1 \cdots \lambda_m} \left(\frac{a+b}{2} - \lambda \right) (\lambda_1 - \lambda) \cdots (\lambda_m - \lambda)$$

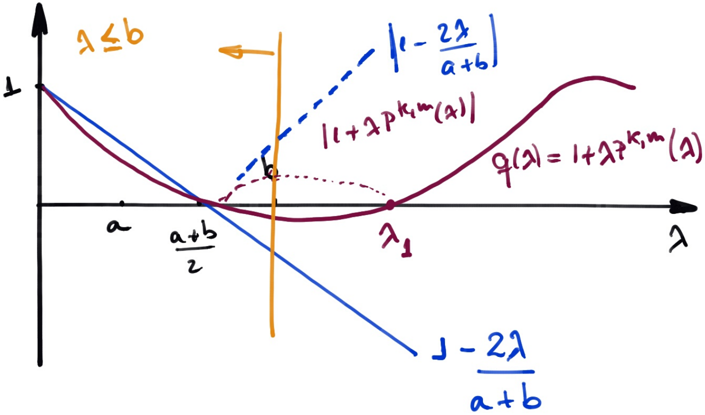
- Como $1 + \lambda_i P^{m,k}(\lambda_i) = 0, i = 1, \ldots, m$ temos que

$$E(x^{k+1}) \leq \max_{\lambda_i \in [a,b]} \left\{ (1 + \lambda_i P^{k,m}(\lambda_i))^2 \right\} E(x^k)$$

- Vamos mostrar que no intervalo $[a, b]$, vale a desigualdade $|1 + \lambda P^{k,m}(\lambda)| \leq |1 - \frac{2\lambda}{a+b}|$
- e, portanto segue o resultado uma vez que o máximo de $\left|1 - \frac{2\lambda}{a+b}\right|^2$ é $\left(\frac{a-b}{a+b}\right)^2$ e ocorre para $\lambda \in \{a, b\}$.


1
b
a
a+b
2
λ₁
λ
q(λ) = 1+λp^{k,m}(λ)
1 - 2λ
a+b

Construção da prova



Prova: para $\lambda \in [0, \frac{a+b}{2}]$, vale $q(\lambda) \leq 1 - \frac{2\lambda}{a+b}$

1. $q(\lambda) = 1 + \lambda P^{k,m}(\lambda)$ possui $m+1$ raizes reais $\Rightarrow$ $q'(\lambda)$ possui m raizes reais intercaladas entre duas raizes de $q(\lambda)$.
2. Por igual motivo, $q''(\lambda)$ possui $m-1$ raizes reais, intercaladas entre as raizes de $q'(\lambda)$.
3. Como não há raiz de $q(\lambda)$ para $\lambda \in (-\infty, \frac{a+b}{2})$, $q''(\lambda)$ não muda de sinal para $\lambda < \frac{a+b}{2}$.
4. Não é difícil verificar que $q''(0) > 0$ neste intervalo ($q(\lambda)$ é decrescente, $q'(\lambda)$ é crescente) e portanto, $q(\lambda)$ é convexa para $\lambda < \frac{a+b}{2}$.

   Logo, por convexidade $q(\lambda)$ é sub-estimada pela corda que conecta $(0 \;\; 1)^T$ e $(\frac{a+b}{2} \;\; 0)^T$ e então

$$q(\lambda) \leq 1 - \frac{2\lambda}{a+b} \text{ para } \lambda \in [0, \frac{a+b}{2}]$$

Prova: para $\lambda \in [\frac{a+b}{2}, b]$, vale $q(\lambda) \geq 1 - \frac{2\lambda}{a+b}$

Em função do último resultado, $q'(\frac{a+b}{2}) \geq -\frac{2}{a+b}$ já que pelo teorema do valor médio $q'(\frac{a+b}{2}) = -\frac{2}{a+b}$ para algum $\lambda \leq \frac{a+b}{2}$.

- $\lambda \in [\frac{a+b}{2}, b]$ vale para $q(\lambda) \geq 1 - \frac{2\lambda}{a+b}$ pois não é possível que a curva $q(\lambda)$ cruze a linha $1 - \frac{2\lambda}{a+b}$ e depois o eixo λ, no intervalo $[\frac{a+b}{2}, b]$:

  para que isto ocorresse seriam necessárias pelo menos duas trocas de sinal de $q''(\lambda)$ à esquerda da segunda raiz de $q(\lambda)$, enquanto no máximo uma raiz de $q''(\lambda)$ pode existir à esquerda de λ_1, segunda raiz de $q(\lambda)$.

Para $\lambda = b$ (ou para $\lambda = a$), $\left(1 - \frac{2\lambda}{a+b}\right)^2 = \left(\frac{b-a}{b+a}\right)^2$ e o resultado segue.

$$E(x^{k+1}) \leq \left(\frac{b-a}{b+a} \right)^2 E(x^k)$$

1. O Gradiente Conjugado Parcial não é apenas similar ao Método do Gradiente quanto à simplicidade de implementação, mas também sua taxa de convergência é limitada exatamente pela mesma fórmula.
2. Os maiores autovalores de Q são removidos da expressão da taxa (se $m = 0$, a taxa é a mesma que obtivemos para o Método do Gradiente).
3. Esta vantagem é conseguida ao preço de intercalar m passos do gradiente conjugado a cada iteração do gradiente puro, ou seja, fazendo $m + 1$ passos do Gradiente Conjugado.
4. Ainda assim k iterações do Gradiente Conjugado Parcial podem produzir melhor resultados bastante melhores que $(m + 1)k$ iterações do Gradiente Puro, para problemas mal condicionados.

Abordagens principais para o uso do método

1. Aproximação quadrática:
   - $g^k = \nabla f_0(x^k)$ é utilizado em cada iteração
   - $B^k = \nabla^2 f_0(x^k)$ substitui Q nas expressões do método do Gradiente Conjugado.
   - Calcula-se α^k com estas substituições
   - β^k é calculado sem as simplificações decorrentes do Gradiente Conjugado
2. Busca unidirecional:
   - α^k é obtido via busca unidirecional, exata ou aproximada (evita-se a avaliação de $\nabla^2 f_0(x^k)$).
   - O valor de β^k simplificado é usado, bem como outras variantes de cálculo que são equivalentes à busca unidirecional exata, quando o problema é quadrático (a de Polak-Ribiere sendo bastante empregada)

Repita até a satisfação de algum critério de convergência:

1. Iniciando com x^0, calcule $g^0 = \nabla f_0(x^0)$ e $d^0 = -g^0$.
2. Para todo $k = 0, 1, \ldots, n-1$
   1. Calcule $B^k = \nabla^2 f_0(x^k)$ e faça $x^{k+1} = x^k + \alpha^k d^k$ onde $\alpha^k = -\frac{(g^k)^T d^k}{(d^k)^T B^k d^k}$
   2. Calcule $g^{k+1} = \nabla f_0(x^{k+1})$
   3. Se $k < n-1$, faça $d^{k+1} = -g^{k+1} + \beta^{k+1} d^k$ onde $\beta^{k+1} = \frac{(g^{k+1})^T B^k d^k}{(d^k)^T B^k d^k}$ e, então repita (2-1).
3. Se $k = n-1$, reinicialize, fazendo $x^0 = x^n$, repetindo (1)

Dificuldades:

- Necessidade de avaliação da Hessiana
- Não é necessariamente globalmente convergente.

Repita até a satisfação de algum critério de convergência:

1. Iniciando com x^0, calcule $g^0 = \nabla f_0(x^0)$ e $d^0 = -g^0$.
2. Para todo $k = 0, 1, \ldots, n-1$
   1. $x^{k+1} = x^k + \alpha^k d^k$ onde α^k minimiza $f_0(x^k + \alpha d^k)$
   2. Calcule $g^{k+1} = \nabla f_0(x^{k+1})$
   3. Se $k < n-1$, faça $d^{k+1} = -g^{k+1} + \beta^{k+1} d^k$ onde $\beta^{k+1} = \frac{(g^{k+1})^T g^{k+1}}{(g^k)^T g^k}$ e, então repita a partir do passo (2-1).
3. Se $k = n-1$, reinicialize, fazendo $x^0 = x^n$ e repetindo (1).

- Convergência global, uma vez que ocorre a reinicialização e a busca unidirecional. Direções d^k podem não ser de descida.
- Alternativa para o cálculo dos β's, equivalente caso o problema fosse quadrático (Polak-Ribiere) $\beta^{k+1} = \frac{(g^{k+1})^T (g^{k+1} - g^k)}{(g^k)^T g^k}$

Reinicialização:

- Não há terminação finita para o caso não quadrático
- Nas duas estratégias, recomenda-se a reinicialização: a cada (pelo menos) n passos, é feito
  - $d^0 \leftarrow -\nabla f_0(x^n)$
  - $x^0 \leftarrow x^n$
  - $k \leftarrow 0$
  - repete-se e o processo de gerar direções $d^0, d^1, \ldots, d^{n-1}$ por mais um ciclo, até uma nova reinicialização ou a satisfação de algum critério de convergência.

1. Globalmente convergentes com a implementação da reinicialização + busca unidirecional.
2. Nas vizinhanças do ponto ótimo x^*, $f_0(x)$ é bem aproximada por uma função quadrática. Assumindo que $\nabla^2 f_0(x^*) \succ 0$:
   - Convergência assintótica pelo menos tão boa quanto a do gradiente.
   - Método possui ordem de convergência 2 em relação a cada ciclo de n buscas unidirecionais:

$$\|x_{k+n} - x^*\| \leq c\|x_k - x^*\|,$$

para algum c e $k = 0, n, 2n, \ldots,$.

- Perdemos a propriedade de convergência finita.
- Alternativas para o cálculo de β^k (equivalentes para o caso quadrático)

$$\text{(Fletcher-Reeves)}\ \beta^k = \frac{(g^k)^T g^k}{(g^{k-1})^T g^{k-1}}$$

$$\text{(Polak-Ribiere)}\ \beta^k = \frac{(g^k)^T (g^k - g^{k-1})}{(g^{k-1})^T g^{k-1}}$$

- Mesmo em alguns casos quadráticos, em decorrência de perda de $Q-$conjugacidade (g^k perde ortogonalidade em relação a $\text{span}\{g^0, \ldots, g^{k-1}\}$) associada à erros de arredondamentos, a expressão de β^k de Pola-Ribiere pode funcionar melhor.
- Reinicialização: A cada n passos (pelo menos), fazemos uma reinicialização, isto é, fazemos $d^0 = -\nabla f_0(x^n)$ e recomeçamos o procedimento.

$$x^{k+1} = x^k - \alpha^k H^k \nabla f_0(x^k)$$

- Sabemos que para o problema quadrático $f_0(x) = \frac{1}{2}x^T Qx - b^T x$, $Q \succ 0$ e
$$H^k = \nabla^2 f_0(x^k)^{-1} = Q^{-1},$$
o ótimo é obtido com o passo de Newton ($\alpha^k = 1$) em uma iteração.
- Nas vizinhanças de um ponto estacionário não singular o método de Newton também possui ótimas propriedades de convergência para uma função $f_0(x)$ mais geral.
- A obtenção explícita da Hessiana $\nabla^2 f_0(x^k)$ e a resolução do sistema linear
$$\nabla^2 f_0(x^k) d^k = -\nabla f_0(x^k)$$
podem ser um problema seja pelo custo computacional ou pela necessidade de informação de segunda ordem.

$$x^{k+1} = x^k - \alpha^k (\nabla^2 f_0(x^0))^{-1} \nabla f_0(x^k)$$

- Se a Hessiana $\nabla^2 f_0(x^0)$ varia pouco, o método funciona não muito diferente do Método de Newton.
- Este é exatamente o Método de Newton aplicado para o caso quadrático, com $\alpha^k = 1$, quando a Hessiana é constante.

$$x^{k+1} = x^k - \alpha^k H^k \nabla f_0(x^k) \qquad (17)$$

- Inspirados pelo método de Newton para a resolução de problemas quadráticos, os métodos do tipo Quasi-Newton procuram

  aproximar $\nabla^2 f_0(x^k)^{-1} = Q^{-1}$ por uma matriz H^k

  ao longo de suas iterações, de forma que, idealmente, seja obtida uma direção de descida.
- A aproximação da inversa Hessiana é construída ou atualizada a cada iteração do método (métodos de métrica variável).
- Para um problema quadrático, em um número finito de iterações, obtém-se $H^k = Q^{-1}$, o ótimo x^* (ponto estacionário não singular, $Qx^* = b$) e o método funciona aproximadamente como o Método de Newton.

- A ideia central consiste em coletar informações sobre a taxa de modificação dos gradientes, usando-as para construir uma aproximação da inversa da Hessiana.

$$
\begin{aligned}
\nabla f_0(x+h) &= \nabla f_0(x) + \nabla^2 f_0(x)h + o(\|h\|) \\
\nabla f_0(x+h) - \nabla f_0(x) &= \nabla^2 f_0(x)h + o(\|h\|) \\
\nabla f_0(x+h) - \nabla f_0(x) &\approx \nabla^2 f_0(x)h
\end{aligned}
$$

Se a Hessiana é constante, por exemplo, no caso quadrático:

$$
\begin{aligned}
\nabla f_0(x^{k+1}) - \nabla f_0(x^k) &= Q(x^{k+1} - x^k) \\
g^k &= Qp^k \\
Q^{-1}g^k &= p^k
\end{aligned}
$$

onde $g^k = \nabla f_0(x^{k+1}) - \nabla f_0(x^k)$ e $p^k = x^{k+1} - x^k$

Condição de Quasi-Newton

- Para o caso quadrático, a avaliação do gradiente de f_0 em dois pontos distintos fornece informação sobre Q, Q^{-1}:

$$Q^{-1}g^k = p^k$$

- Em particular, se H^{k+1} fosse constante igual a Q^{-1}, teríamos a condição:

$$H^{k+1}g^i = p^i, \;\; 0 \leq i \leq k, \;\; k = 0, 1, \ldots,$$

Ideia central

Iniciamos com uma aproximação $H^0 \succ 0$ (por exemplo $H^0 = I$) e obtemos H^1 a partir de H^0, e assim por diante, obtemos H^{k+1} a partir de H^k.

Uso da ideia

- Claramente, não se observa $H^0 g^0 = p^0$ pois $\nabla f_0(x^1)$ só é disponível depois que o deslocamento $-\alpha^0 H^0 \nabla f_0(x^0)$ a partir de x^0 é dado.
- Então define-se a matriz H^1 que se deseja obter de forma que se verifique a condição $H^1 g^0 = p^0$.
- Em suma, utiliza-se a invariante dada pela condição de Quasi-Newton para $i = k$ e obtém-se a aproximação H^{k+1} de Q^{-1}.

Condição de Quasi-Newton:

$$H^{k+1}g^i = p^i,\ \ 0 \leq i \leq k,\ \ k = 0, 1, \ldots,$$

- A condição acima é observada no caso quadrático, onde a Hessiana é constante. Esta forma construtiva será explorada para construir as aproximações da inversa Hessiana.
- Se forem disponíveis $\{p^k : k = 0, 1, \ldots, n-1\}$ linearmente independentes e os correspondentes $g^k : k = 0, 1, \ldots, n-1$, o sistema linear obtido quando a última direção p^{n-1} tiver sido obtida (última das buscas unidirecionais), permite recuperar a Hessiana, de forma única:

  $$Q = GP^{-1}$$

  fornece a Hessiana da função objetivo quadrática, onde $G = (g^0 \ \ \ldots \ \ g^{n-1})$ e $P = (p^0 \ \ \ldots \ \ p^{n-1})$.
- Ou seja, após n passos linearmente independentes, obtemos Q.

1. A propriedade satisfeita ao longo das iterações dos algoritmos Quasi-Newton
$$H^{k+1} g^i = p^i, \; 0 \le i \le k$$
em conjunto com o fato de que a Hessiana no caso quadrático é constante e postiva definida, garante que os vetores $p^i : i = 0, 1, \ldots$ gerados são linearmente independentes.
2. Em particular, p^i é autovetor de $H^{k+1} Q$ para $i = 0, \ldots, k$, com respectivo autovalor 1.

$$\begin{aligned} H^{k+1} g^i &= p^i & i &= 0, \ldots, k \\ Q p^i &= g^i & i &= 0, \ldots, k \\ H^{k+1} Q p^i &= H^{k+1} g^i & i &= 0, \ldots, k \end{aligned}$$

Condição Quasi Newton

$$H^{k+1}g^i = p^i,\ \ 0 \leq i \leq k,\ \ k = 0, 1, \ldots, \tag{18}$$

$$H^{k+1} = H^k + E^k$$

Desejavelmente:

- E^k deve ser fácil de ser calculada, envolvendo poucas operações aritméticas.
- Sua avaliação deve ser estável numericamente.
- E^k deve garantir a positividade e simetria de H^{k+1}, gerando direções de descida.

- Há grande flexibilidade para sua atualização.
- A maneira como E^k é calculada define um conjunto de família de métodos do tipo Quasi-Newton.

- Dispomos de $x^k, H^k, \nabla f_0(x^k)$, calculamos d^k, damos um passo e obtemos α^k, x^{k+1} e os vetores g^k, p^k.
- Desejamos que a aproximação da inversa Hessiana satisfaça a

Restrição Secante (Condição Quasi-Newton para $i = k$)

$$H^{k+1} g^k = p^k \tag{19}$$

- Isto é equivalente a impor que B^{k+1}, a aproximação da Hessiana ($B^{k+1} = (H^{k+1})^{-1}$), satisfaça $B^{k+1} p^k = g^k$.

Restrição Secante

$$H^{k+1} g^k = p^k$$

- Para que exista $H^{k+1} \succ 0$ satisfazendo o mapeamento (19), devemos ter $(g^k)^T p^k > 0$ já que $(g^k)^T H^{k+1} g^k = (g^k)^T p^k > 0$ pela positividade de H^{k+1}.
- A condição $(g^k)^T p^k > 0$ é satisfeita caso a busca unidirecional seja exata ou, se a condição de curvatura de Wolfe (11), for atendida. Veja que como $0 < c_2 < 1$ e d^k é de descida, temos que

$$\begin{aligned} \nabla f_0(x^k + \alpha d^k)^T d^k \geq c_2 \nabla f_0(x^k)^T d^k & \quad \rightarrow \\ \nabla f_0(x^k + \alpha d^k)^T d^k > \nabla f_0(x^k)^T d^k & \quad \rightarrow \\ (g^k)^T p^k > 0 & \end{aligned}$$

- A condição $(g^k)^T p^k > 0$ é necessária para existir solução para (19), mas temos excesso de graus de liberdade.

Restrição Secante

$$H^{k+1}g^k = p^k, \;\; k = 0, 1, \ldots,$$

Como Q, Q^{-1} são simétricas, é natural propor algo do tipo:

$$H^{k+1} = H^k + a^k z^k (z^k)^T$$

- Para $a^k z^k (z^k)^T \neq 0$, posto$(a^k z^k (z^k)^T) = 1$ onde $a^k \in \mathbb{R}, z^k \in \mathbb{R}^n$
- Os valores de a^k, z^k são escolhidos impondo-se a condição Quasi-Newton para $i = k$, obtendo-se:

$$H^{k+1}g^k = p^k, \;\; k = 0, 1, \ldots, \tag{20}$$

$$
\begin{aligned}
H^{k+1} &= H^k + a^k z^k (z^k)^T \\
p^k &= H^{k+1} g^k \\
p^k &= H^k g^k + a^k z^k (z^k)^T g^k
\end{aligned}
$$

- Temos então que z^k é proporcional a $p^k - H^k g^k$ ($a^k (z^k)^T g^k \in \mathbb{R}$)

$$z^k a^k (z^k)^T g^k = p^k - H^k g^k,$$

- Fazemos $z^k = p^k - H^k g^k$ e, para garantir, por exemplo que, $a^k (z^k)^T g^k = 1$, impomos $a^k = \frac{1}{(z^k)^T g^k} = \frac{1}{(p^k - H^k g^k)^T g^k}$.

$$H^{k+1} = H^k + \frac{(p^k - H^k g^k)(p^k - H^k g^k)^T}{(p^k - H^k g^k)^T g^k}$$

$$H^{i+1} = H^i + \frac{(p^i - H^i g^i)(p^i - H^i g^i)^T}{(p^i - H^i g^i)^T g^i} \tag{21}$$

1. Por construção, a expressão acima garante que

$$H^{k+1} g^i = p^i, \ \ i = k.$$

2. Falta mostrar que, para o caso no qual a Hessiana (e sua inversa) são constantes, também é garantido que

$$H^{k+1} g^i = p^i, \ \ i < k$$

$$H^{i+1} = H^i + \frac{(p^i - H^i g^i)(p^i - H^i g^i)^T}{(p^i - H^i g^i)^T g^i}$$

Teorema

Suponha que Q seja uma matriz simétrica fixa e $p^0, p^1, \ldots, p^k$ sejam dados. Defina os vetores $g^i = Qp^i$ para $i = 0, 1, 2, \ldots, k$ e considere a sequência gerada de matrizes H^{i+1} geradas por (21), onde H^0 é qualquer matriz simétrica. Então temos

$$p^i = H^{k+1} g^i, \quad \text{para } i \leq k$$

e H^{k+1} converge para Q^{-1} em no máximo n passos.

Prova por indução em i.

(HI) Vamos supor que seja válida para H^k e $i \leq k-1$

Verificando o que se passa para H^{k+1} e $i \leq k$:

- Para $i = k$, a condição de Quasi-Newton (18) foi usada para $k = i$, obtendo-se (21). Então, vale para $i = k$.
- Multiplicando (21) (formulada para $i = k$) por $g^i : i < k$ e definindo $y^k = \frac{(p^k - H^k g^k)}{(p^k - H^k g^k)^T g^k}$, temos:

$$
\begin{aligned}
H^{k+1} &= H^k + y^k (p^k - H^k g^k)^T \\
H^{k+1} g^i &= H^k g^i + y^k ((p^k)^T g^i - (g^k)^T H^k g^i) \\
H^{k+1} g^i &= p^i + y^k ((p^k)^T g^i - (g^k)^T p^i)
\end{aligned}
$$

- Observe que $(g^k)^T p^i = (p^k)^T Q p^i = (p^k)^T g^i$ e então o segundo termo do lado direito da última expressão é nulo, concluindo a prova.

Partindo de $H^0 \in \mathcal{S}^n_{++}$, geramos as aproximações da inversa por:

$$H^{i+1} = H^i + \frac{(p^i - H^i g^i)(p^i - H^i g^i)^T}{(p^i - H^i g^i)^T g^i}, \quad i = 0, 1, \ldots$$

Dificuldades ao usarmos $d^i = -H^i \nabla f_0(x^i)$ em um algoritmo:

- H^{i+1} pode não ser positiva definida se $(p^i - H^i g^i)^T g^i < 0$.
- Ainda que $(p^i - H^i g^i)^T g^i > 0$, estes valores podem ser muito próximos de zero, levando à instabilidade numérica.
- A atualização de posto-1 ilustra bem a ideia dos métodos nesta classe, mas possui muitas limitações para poder ser usada para minimizar funções $f_0(x)$ mais gerais.

$$H^{k+1} = H^k + auu^T + bvv^T$$

Davidon-Fletcher-Powell: Proposto por Davidon (1959) e posteriormente aprimorado por Fletcher e Powell (1963).

- Se p^k e $H^k g^k$ não forem proporcionais, são li e candidatos naturais para gerar correções de posto 2: Fazemos $u = p^k$, $v = H^k g^k$.

$$\begin{aligned} H^{k+1} g^k &= H^k g^k + auu^T g^k + bvv^T g^k \\ p^k &= H^k g^k + u(au^T g^k) + v(bv^T g^k) \end{aligned}$$

- Fixando $au^T g^k = 1$ e $bv^T g^k = -1$, obtemos:

$$H^{k+1} = H^k + \frac{p^k (p^k)^T}{(p^k)^T g^k} - \frac{H^k g^k (g^k)^T H^k}{(g^k)^T H^k g^k} \tag{22}$$

(Inicialização) Dada uma matriz $H^0 \in \mathcal{S}^n_{++}$ e um ponto inicial $x^0 \in \mathbb{R}^n$, $k = 0$

Repita:

(Passo 1) Defina a direção $d^k = -H^k \nabla f_0(x^k)$

(Passo 2) Encontre $\alpha^k = \arg\min_{\alpha \geq 0} f_0(x^k + \alpha d^k)$ e defina:

- $x^{k+1} = x^k + \alpha^k d^k$
- $p^k = \alpha^k d^k$

(Passo 3) Defina $g^k = \nabla f_0(x^{k+1}) - \nabla f_0(x^k)$ e atualize a aproximação da inversa da Hessiana, segundo a atualização de posto 2:

$$H^{k+1} = H^k + \frac{p^k (p^k)^T}{(p^k)^T g^k} - \frac{H^k g^k (g^k)^T H^k}{(g^k)^T H^k g^k}$$

Faça $k \leftarrow k + 1$ e retorne ao passo 1

Teorema

Se $H^k \succ 0$, então a matriz H^{k+1} obtida segundo (22) satisfaz $H^{k+1} \succ 0$.

Prova

$$H^{k+1} = H^k + \frac{p^k (p^k)^T}{(p^k)^T g^k} - \frac{H^k g^k (g^k)^T H^k}{(g^k)^T H^k g^k}$$

$$x^T H^{k+1} x = x^T H^k x + \frac{(x^T p^k)^2}{(p^k)^T g^k} - \frac{(x^T H^k g^k)^2}{(g^k)^T H^k g^k}$$

Definindo $a = (H^k)^{\frac{1}{2}} x, b = (H^k)^{\frac{1}{2}} g^k$ segue que:

$$x^T H^{k+1} x = \frac{(a^T a)(b^T b) - (a^T b)^2}{b^T b} + \frac{(x^T p^k)^2}{(p^k)^T g^k}$$

Prova - continua

$$x^T H^{k+1} x = \frac{(a^T a)(b^T b) - (a^T b)^2}{b^T b} + \frac{(x^T p^k)^2}{(p^k)^T g^k}$$

- Observamos que $(p^k)^T g^k = (p^k)^T(\nabla f_0(x^{k+1}) - \nabla f_0(x^k)) = -(p^k)^T \nabla f_0(x^k)$, já que $(p^k)^T \nabla f_0(x^{k+1}) = 0$ uma vez que α^k foi obtido de forma exata.
- Pela definição de $p^k = \alpha^k d^k = -\alpha^k H^k \nabla f_0(x^k)$ segue que $(p^k)^T g^k = \alpha^k \nabla f_0(x^k)^T H^k \nabla f_0(x^k) > 0$ para qualquer $\alpha^k > 0, \nabla f_0(x^k) \neq 0$. Logo temos:

$$x^T H^{k+1} x = \frac{(a^T a)(b^T b) - (a^T b)^2}{b^T b} + \frac{(x^T p^k)^2}{\alpha^k \nabla f_0(x^k)^T H^k \nabla f_0(x^k)}$$

Prova - continua

$$x^T H^{k+1} x = \frac{(a^T a)(b^T b) - (a^T b)^2}{b^T b} + \frac{(x^T p^k)^2}{\alpha^k \nabla f_0(x^k)^T H^k \nabla f_0(x^k)}$$

- Por Cachy-Schwartz: $\|a\|\|b\| \geq \langle a, b \rangle$, logo os dois termos do lado direito da expressão acima são não negativos.
- Vamos mostrar que os dois nunca se anulam simultaneamente. O primeiro se anula apenas se a e b forem proporcionais, o que implica que $x = \beta g^k$. Neste caso, entretanto, temos:

$$(p^k)^T x = \beta (p^k)^T g^k = \beta \alpha^k \nabla f_0(x^k)^T H^k \nabla f_0(x^k) \neq 0$$

que mostra que $x^T H^{k+1} x > 0$ para qualquer $x \neq 0$.

- Usa o fato de que α^k foi determinado de forma exata e conclui a importante propriedade $(p^k)^T g^k > 0$, que garante $H^{k+1} \succ 0$.

Proposição

Qualquer α^k que garanta $(p^k)^T g^k > 0$ pode ser usado, garantindo que $H^{k+1} \succ 0$.

- O resultado anterior indica que a busca exata pode ser substituída por outra mais fraca, mas realista. Por quê é realista ?
    - Se o problema é quadrático, $(p^k)^T g^k = (p^k)^T Q p^k > 0$, já que $Q \succ 0$.
    - Se não é, basta usar a condição de curvatura de Wolfe (11) em uma busca aproximada de α.
- Apesar destas vantagens, DFP é razoavelmente sensível à busca inexata: há evidência numérica de degradação das taxas convergência de DFP, quando c_2 em (11) cresce.
- Com a popularização das buscas unidirecionais inexatas ao longo da década de 1970, cairam e desuso e são menos utilizados que os métodos Quasi-Newton do tipo BFGS.

Teorema

Seja $f_0(x)$ uma função quadrática definida por matriz Hessiana $Q \succ 0$. Então para o método DFP, vale:

1. $p^i Q p^j = 0, \;\; 0 \leq i < j \leq k$
2. $H^{k+1} Q p^i = p^i, \;\; 0 \leq i \leq k$

Interpretação dos resultados

1. As direções $p^0, \ldots, p^j : j = 0, \ldots, n-1$ são $Q-$conjugadas ou $Q-$ortogonais.
2. Para todo $i = 0, \ldots, k$, p^i é autovetor de $H^{k+1}Q$ com autovalor unitário. Como pela $Q-$conjugacidade estes vetores $p^i : i = 0, \ldots, n-1$ são li, temos que $H^n = Q^{-1}$.

Broyden (1970), Fletcher (1970), Goldfarb (1970) e Shanno (1970)

Aproxima a Hessiana por $\{B^k\}$ a partir de $B^0 \succ 0$. Ao invés de utilizar $H^{k+1}g^k = p^k$, utiliza-se a análoga $B^{k+1}p^k = g^k$ e obtém-se uma expressão análoga à expressão de DFP:

$$B^{k+1} = B^k + \frac{g^k(g^k)^T}{(g^k)^T p^k} - \frac{B^k p^k (p^k)^T B^k}{(p^k)^T B^k p^k} \tag{23}$$

1. Partindo de:

$$B^{k+1} = B^k + \frac{g^k(g^k)^T}{(g^k)^T p^k} - \frac{B^k p^k (p^k)^T B^k}{(p^k)^T B^k p^k}$$

2. Utilizando $(B^{k+1})^{-1} = H^{k+1}$ e a relação de Sherman - Morrison (a inversa de uma matriz corrigida por um matriz de posto 1 é a inversa da matriz original mais um termo de posto 1):

$$(A + ab^T)^{-1} = A^{-1} - \frac{1}{b^T A^{-1} a}(A^{-1} a b^T A^{-1})$$

3. Escrevemos expressão de $H^{k+1} = (B^{k+1})^{-1}$, aplicando a expressão de Sherman - Morrison duas vezes, uma para $a = b = g^k$ e a outra para $a = b = B^k p^k$ como:

$$H^{k+1} = H^k + \left(1 + \frac{(g^k)^T H^k g^k}{(p^k)^T g^k}\right) \frac{p^k (p^k)^T}{(p^k)^T g^k} - \left(\frac{p^k (g^k)^T H^k + H^k g^k (p^k)^T}{(p^k)^T g^k}\right) \tag{24}$$

Combinação afim de DFP e BFGS

$$\begin{aligned} H &= \phi H^{DFP} + (1-\phi)H^{BFGS} \\ &= H^{DFP} + \phi v v^T \end{aligned}$$

para algum vetor v.

- Busca unidirecional *controlada*: $H^{DFP} \succ 0$ se $H^0 \succ 0$
- A escolha de ϕ é irrelevante para problemas quadráticos, diante de busca unidirecional exata (resultado análogo ao Teorema anterior para DFP, garante convergência finita em n passos para qualquer método da família Broyden).
- $\phi \geq 0$ é normalmente adotado para garantir positividade da nova matriz H.
- Permitem modificar H para que $-H\nabla f_0(x)$ sempre seja uma direção de descida (através de um múltiplo conveniente de v).

Efeito da busca unidirecional inexata no problema quadrático

- Se a busca unidirecional for *pobre*, DFP pode funcionar tão mal quanto o método do Gradiente puro.
- As taxas de convergência do BFGS são menos sensíveis à busca inexata que o método DFP.

1. Os métodos Quasi-Newton são equivalentes aos métodos de direções conjugadas para o caso quadrático.
2. Podem ser utilizados de forma contínua, sucessivamente atualizando as aproximações da inversa.
3. Métodos Quasi-Newton parciais: reinicialização a cada $m + 1$ passos. Vantagens:
   - A reinicialização garante convergência global para o caso não quadrático.
   - Pequena complexidade de memória: a inversa Hessiana pode ser recuperada a partir dos $g^i, p^i : i = 0, \ldots, m$.
   - Em torno de uma vizinhança do ponto estacionário, devem se comportar *aproximadamente* como um método de direções conjugadas.