INSTITUTO LATINO-AMERICANO DE CIÊNCIAS
DA VIDA E DA NATUREZA (ILACVN)

MATEMÁTICA, GRAU LICENCIATURA

OTIMIZAÇÃO:
CONDIÇÕES DE OTIMALIDADE DE KARUSH-KUHN-TUCKER

LARISSA MORAES DA SILVA

Foz do Iguaçu
2023

INSTITUTO LATINO-AMERICANO DE CIÊNCIAS
DA VIDA E DA NATUREZA (ILACVN)

MATEMÁTICA, GRAU LICENCIATURA

OTIMIZAÇÃO:
CONDIÇÕES DE OTIMALIDADE DE KARUSH-KUHN-TUCKER

LARISSA MORAES DA SILVA

Trabalho de Conclusão de Curso apresentado ao Instituto Latino-Americano de Ciências da Vida e da Natureza da Universidade Federal da Integração Latino-Americana, como requisito parcial à obtenção do título de Licenciada em Matemática.

Orientador(a): Profª. Dra. Adriana Flores de Almeida.
Coorientador(a): Profª. Dra. Camila Isoton

Foz do Iguaçu
2023

LARISSA MORAES DA SILVA

OTIMIZAÇÃO:
CONDIÇÕES DE OTIMALIDADE DE KARUSH-KUHN-TUCKER

Trabalho de Conclusão de Curso apresentado ao Instituto Latino-Americano de Ciências da Vida e da Natureza da Universidade Federal da Integração Latino-Americana, como requisito parcial à obtenção do título de Licenciatura em Matemática.

BANCA EXAMINADORA

gov.br Documento assinado digitalmente
ADRIANA FLORES DE ALMEIDA
Data: 27/06/2023 09:58:35-0300
Verifique em https://validar.iti.gov.br

Orientadora: Profª. Dra. Adriana Flores de Almeida
UNILA

gov.br Documento assinado digitalmente
CAMILA ISOTON
Data: 27/06/2023 10:34:32-0300
Verifique em https://validar.iti.gov.br

Coorientadora: Profª. Dra. Camila Isoton
UFGD

gov.br Documento assinado digitalmente
ELVIS MANUEL RODRIGUEZ TORREALBA
Data: 27/06/2023 14:58:13-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Elvis Manuel Rodriguez Torrealba
UNILA

Foz do Iguaçu, 07 de Junho de 2023 .

TERMO DE SUBMISSÃO DE TRABALHOS ACADÊMICOS

Nome completo do autor(a): Larissa Moraes da Silva.

Curso: Matemática, grau Licenciatura.

Tipo de Documento

(.....) graduação	(.....) artigo
(.....) especialização	(X) trabalho de conclusão de curso
(.....) mestrado	(.....) monografia
(.....) doutorado	(.....) dissertação
	(.....) tese
	(.....) CD/DVD – obras audiovisuais
	(.....) ________________________________

Título do trabalho acadêmico: Otimização: Condições de Otimalidade de Karush-Kuhn-Tucker.

Nome do orientador(a): Adriana Flores de Almeida.

Data da Defesa: 07/06/2023.

Licença não-exclusiva de Distribuição

O referido autor(a):

a) Declara que o documento entregue é seu trabalho original, e que o detém o direito de conceder os direitos contidos nesta licença. Declara também que a entrega do documento não infringe, tanto quanto lhe é possível saber, os direitos de qualquer outra pessoa ou entidade.

b) Se o documento entregue contém material do qual não detém os direitos de autor, declara que obteve autorização do detentor dos direitos de autor para conceder à UNILA – Universidade Federal da Integração Latino-Americana os direitos requeridos por esta licença, e que esse material cujos direitos são de terceiros está claramente identificado e reconhecido no texto ou conteúdo do documento entregue.

Se o documento entregue é baseado em trabalho financiado ou apoiado por outra instituição que não a Universidade Federal da Integração Latino-Americana, declara que cumpriu quaisquer obrigações exigidas pelo respectivo contrato ou acordo.

Na qualidade de titular dos direitos do conteúdo supracitado, o autor autoriza a Biblioteca Latino-Americana – BIUNILA a disponibilizar a obra, gratuitamente e de acordo com a licença pública *Creative Commons* **Licença 3.0 Unported.**

Foz do Iguaçu, 07 de Junho de 2023.

gov.br
Documento assinado digitalmente
LARISSA MORAES DA SILVA
Data: 26/06/2023 22:07:40-0300
Verifique em https://validar.iti.gov.br

____

Assinatura do Responsável

Dedico este trabalho a minha querida mãe, por todo apoio que teve comigo ao longo desta minha trajetória.

AGRADECIMENTOS

Gostaria de agradecer primeiramente a Deus por todas as bençãos concebida em minha vida.

Ás minhas professoras Adriana Flores e Camila Isoton, pela orientação deste trabalho. Obrigada pela paciência, dedicação, confiança, amizade e disponibilidade que tiveram comigo.

Agradeço à minha querida e amada mãe Claudete e ao meu irmão Luann, que sempre estiveram ao meu lado em todos os momentos, me dando excelentes conselhos, amor, incentivo e compreensão. É graças a vocês que eu concluo esta etapa da minha vida.

Aos professores da banca examinadora Adriana, Camila, Elvis e Patricia, pelas sugestões dadas e melhorias deste trabalho de conclusão de curso.

Aos meus colegas de curso, por compartilharem comigo tantos momentos de descobertas e aprendizado e por todo o companheirismo ao longo deste percurso. Em especial Isabelli, Ana Leticia, Luciane e Matheus.

E, finalmente, mas não menos importante, sou muito grata por todos aqueles professores que fizeram parte do meu processo de formação.

Nunca considere os estudos como uma obrigação, e sim como uma oportunidade para entrar no belo mundo dos sábios.

Albert Einstein

RESUMO

Neste trabalho apresentou-se um estudo sobre as condições de otimalidade de primeira e segunda ordem para problemas de otimização sem restrições e ,posteriormente, condições de primeira ordem para problemas com restrições de igualdade e desigualdade, cujo o principal foco foi a demonstração do Teorema de Karush-Kuhn-Tucker (KKT) via teoria de cones. Finalmente, realizou-se um estudo sobre as condições de qualificação mais clássicas existente na literatura.

Palavras-chaves: Programação não linear, condições de otimalidade, teoria de cones, condições de qualificação.

RESUMEN

En este trabajo presentamos un estudio sobre las condiciones de optimalidad de primer y segundo orden para problemas de optimización sin restricciones y ,seguidamente, studiamos condiciones de primer orden para problemas con restricciones de igualdad y desigualdad, el enfoque principal de este trabajo es la demostración del Teorema Karush-Kuhn-Tucker (KKT) vía de la teoría de conos. Finalmente, se realizó un estudio sobre la calificación de restricciones más clásicas de la literatura.

Palabras clave: Programación no lineal, condiciones de optimalidad, teoría de conos, calificación de restricciones.

ABSTRACT

In this work we present a study on the first and second order optimality conditions for unconstrained optimization problems and later on the first order conditions for problems with equality and inequality constraints, whose main focus was the demonstration of the Karush-Kuhn-Tucker (KKT) Theorem by cone theory. Also that, carried out a study on the most classic qualification conditions existing in the literature.

Keywords: Nonlinear programming, optimality conditions, cone theory, qualification conditions.

LISTA DE SÍMBOLOS

x^k	Uma sequência em $\mathbb{R}^n$ é uma aplicação $k \in \mathbb{N} \to x^k \in \mathbb{R}^n$, definida no conjunto $\mathbb{N} = \{0, 1, 2, 3, ...\}$.
$x \in \mathbb{R}^n$	O ponto $x \in \mathbb{R}^n$ é o limite da sequência x^k quando, $\forall \epsilon > 0$ arbitrário, o conjunto $\mathbb{N}_1 = \{k \in \mathbb{N} : \|x^k - x\| \geq \epsilon\}$ é finito.
$x^k \to x$	A sequência x^k converge para x.
	Dado um ponto $x \in \mathbb{R}^n$ e um conjunto $S \subset \mathbb{R}^n$ temos que:
$B(x, \delta)$	A bola aberta de raio δ e centro x é o conjunto $B(x, \delta) = \{a \in \mathbb{R}^n : \|a - x\| < r\}$, com $r > 0$ um número real positivo.
$V(x)$	O conjunto V é um vizinhança do ponto x, quando $x \in intV$.
$int(S)$	O ponto $x \in S$ é um ponto interior de S se existir $\delta > 0$ tal que $B(x, \delta) \subset S$.
$int(S)$	O conjunto de todos os pontos interiores de S.
$S = int(S)$	Um conjunto S é aberto quando todos os seus pontos são interiores.
FrS	O conjunto dos pontos fronteiras de S é chamado de fronteira de S.
$FrS \subset S$	Um conjunto S é fechado quando contém sua fronteira.
$S \subset \mathbb{R}^n$	Um conjunto S é limitado se existe uma bola aberta que o contém.
$S \subset \mathbb{R}^n$	O conjunto S é compacto se for fechado e limitado
v	Valor de ínfimo do conjunto.
$\|x\|$	A norma euclidiana de vetor x do $\mathbb{R}^n$ é dada por $\|x\| = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}$.
	Seja $A \in \mathbb{R}^{n \times m}$, uma matriz:
A^T	Matriz transposta de A.

H_f A matriz hessiana é formada função de várias variáveis $f(x_1, x_2, ..., x_n)$, e tem entradas dadas por derivadas parciais de segunda ordem de uma matriz, isto é,

$$H_f(x_1, x_2, ..., x_n) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & & & \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

$\|A\| = max\{|\lambda_1|, |\lambda_2|, ..., |\lambda_n|\}$ Norma euclidiana do máximo de uma matriz simétrica, coincide com o maior valor absoluto de seus autovalores $\lambda_1, \lambda_2, ..., \lambda_n$.

$A > 0$ Uma matriz é definida positiva, quando $x^T Ax > 0$, $\forall x \in \mathbb{R}^n \backslash \{0\}$.

$A \geq 0$ Uma matriz é semidefinida positiva, se $x^T Ax \geq 0$, $\forall x \in \mathbb{R}^n$.

$proj_S(x)$ Projeção de x sobre um conjunto $S \in \mathbb{R}^n$.

SUMÁRIO

1 INTRODUÇÃO

Diariamente encontramos situações que nos levam a tomada de decisões. Esse tipo de problema surge por exemplo, nas áreas da Engenharia, Economia, Administração, etc. Os modelos de otimização e suas técnicas visam minimizar (ou maximizar) uma função (nem sempre linear), denominada função objetivo, de várias variáveis escolhidas dentro de alguma região admissível (factível), que é definida pelas restrições impostas às variáveis do problema. Esses problemas que buscam calcular o ponto ótimo (de mínimo ou máximo) de alguma função objetivo, sujeita ou não a restrições, são chamados problemas de programação matemática.

Nosso foco neste trabalho é tratar das condições de otimalidade, inicialmente dos problemas irrestritos e, na sequência, buscando atingir aplicações mais realistas. Abordaremos problemas de Programação Matemática restritos ás condições de igualdade e desigualdade. Neste contexto, buscamos estudar as condições de otimalidade necessárias estabelecidas por Karush-Kuhn-Tucker (KKT) em [7, 8], e outras bibliografias básicas sobre o tema, tais como, [3, 4, 5, 6, 11].

As condições de KKT estabelecem condições necessárias e desempenham um papel fundamental para encontrar um ponto x^* candidato a minimizador local. A fim de obtermos condições suficientes para determinar se um ponto é minimizador, é comum adicionarmos hipóteses às restrições, conhecidas como Condições de Qualificação (CQ).

Estudamos neste trabalho de conclusão de curso, as seguintes condições de qualificações: Condição de Slater, independência linear e Mangasarian-Fromovitz.

Estruturação dos Capítulos

No Capítulo 2 detalhamos as condições de otimalidade para o problema irrestrito. No Capítulo 3 estudamos sobre alguns resultados clásicos de análise convexa de extrema importância na teoria de otimização. No Capítulo 4 realizamos uma interpretação geométrica e algébrica do Lema de Farkas, usando o conceito de cone. No Capítulo 5 estudamos o problema de otimização restrito, nele, abordamos as direções viáveis e de descida, fazendo uma explanação do cone viável linearizado e o cone tangente, mostrando as relações entre eles. Além disso, demonstramos as condições de otimalidade de Karush-Kuhn-Tucker. No Capítulo 6 finalmente, estudamos as condições de qualificação já citadas nesta introdução, com o interesse de obtermos condições fracas e fáceis de serem verificadas, completando as condições de KKT definidas no Capítulo 5. Ao final, apresentamos algumas conclusões.

2 OTIMIZAÇÃO IRRESTRITA

Neste capítulo, serão apresentados algumas definições básicas sobre os problemas de otimização, tais como, resultados de existência de soluções e as condições de otimalidade para problemas irrestritos.

Consideremos o seguinte problema:

$$\begin{aligned} &Minimizar \;\; f(x) \\ &Sujeito \;\; a \;\; x \in \Omega. \end{aligned} \tag{P_Ω}$$

Em que f é chamada de função objetivo e $\Omega \subset \mathbb{R}^n$ é um conjunto viável, os pontos de Ω são chamados pontos viáveis ou factíveis.

Ressaltamos que, minimizar uma função $f(x)$ é o mesmo que maximizar a função $-f(x)$, devido a esta equivalência, optamos neste trabalho em tratar somente dos problemas de minimização.

Deste modo, nos concentramos no estudo de dois conceitos de soluções para o problema (P_Ω).

Definição 2.1 *Dizemos que um ponto $x^* \in \Omega$ é*

- Minimizador Local de f em Ω, se existe uma vizinhança V do ponto x^* tal que $f(x^*) \leq f(x)$, para todo $x \in \Omega \cap V$. $\quad (1.1)$

- Minimizador Global de f em Ω, se $f(x^*) \leq f(x)$, $\forall x \in \Omega$. $\quad (1.2)$

Pela definição, todo minimizador global também é minimizador local, mas não reciprocamente. Se para todo $x \neq x^*$ a desigualdade (1.1) ou (1.2) é estrita, o ponto x^* será chamado minimizador estrito (local ou global, respectivamente). Além disso, dizemos que o valor ótimo do problema (P_Ω), se define com $v = inf_{x \in \Omega} f(x)$.

O seguinte teorema é um resultado fundamental para problemas de otimização.

Teorema 2.1 ***(Weierstrass)*** *Sejam $f : \mathbb{R}^n \to \mathbb{R}$ contínua e $\Omega \subset \mathbb{R}^n$ compacto não vazio. Então existe minimizador global de f em Ω.*

Para demonstração, veja [9].

Vejamos agora que é possível garantir a existência de minimizador global em $\mathbb{R}^n$, sem supor compacidade. Contudo, estabelecemos uma hipótese a mais sobre a função objetivo.

Definição 2.2 *Seja $f : \mathbb{R}^n \to \mathbb{R}$, dizemos f que é coerciva quando $\lim\limits_{\|x\|\to\infty} f(x) = \infty$, isto para todo $M > 0$, temos $r > 0$ tal que $f(x) > M$ sempre que $\|x\| > r$. Em outras palavras, em um conjunto $A \subset \mathbb{R}^n$ que não seja limitado, temos que os valores de $f(x), x \in A$, também não o serão.*

Teorema 2.2 *Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função contínua e coerciva. Então, f tem um minimizador global.*

Para prova, veja por exemplo [11].

2.1 Condições de Otimalidade para Problemas Irrestritos

A princípio vamos tratar de métodos para minimizar funções diferenciáveis no conjunto $\mathbb{R}^n$. As condições necessárias são caracterizadas por um minimizador, enquanto as condições suficientes, quando satisfeitas, garantem que o ponto de interesse é um minimizador local.

Definição 2.3 *Dizemos que $d \in \mathbb{R}^n$ é uma direção factível em relação ao conjunto Ω no ponto $x^* \in \Omega$, quando existe $\varepsilon > 0$ tal que*

$$x^* + td \in \Omega, \forall t \in [0, \varepsilon].$$

Portanto, denota-se por $F(x)$ como sendo o conjunto de todas as direções factíveis em relação ao conjunto Ω no ponto x^*.

Teorema 2.3 ***(Condição necessária de 1ª ordem)*** *Seja $f : \mathbb{R}^n \to \mathbb{R}$ diferenciável no ponto $x^* \in \mathbb{R}^n$. Se x^* é um minimizador local de f, então*

$$\nabla f(x^*) = 0. \tag{2.1}$$

Prova: Considere $d \in \mathbb{R}^n \backslash \{0\}$ arbitrário. Dado que x^* é minimizador local, existe uma bola $B(x^*, \delta)$, de modo que

$$f(x^*) \leq f(x^* + td), \ \ \forall t \in (0, \delta). \tag{2.2}$$

Pela expansão de Taylor,

$$f(x^* + td) = f(x^*) + t\nabla f(x^*)^T d + r(t) \text{ com } \lim_{t \to 0} \frac{r(t)}{t} = 0.$$

Dividindo a expressão por $t > 0$, e utilizando 2.2, obtemos

$$0 \leq \nabla f(x^*)^T d + \frac{r(t)}{t}.$$

Passando o limite quando $t \to 0^+$, temos $\nabla f(x^*)^T d \geq 0$. Como d é arbitrário, escolhendo $d = -\nabla f(x^*)$, assim temos que

$$-\|\nabla f(x^*)\|^2 = -\nabla f(x^*)^T \nabla f(x^*) = \nabla f(x^*)^T d \leq 0,$$

o que é um absurdo! Portanto, $\nabla f(x^*) = 0$. □

Definição 2.4 *Um ponto $x^* \in \mathbb{R}^n$ que cumpre a condição (2.1) é dito ponto crítico ou estacionário da função f.*

Teorema 2.4 ***(Condição necessária de 2ª ordem)*** *Seja $f : \mathbb{R}^n \to \mathbb{R}$ duas vezes diferenciável no ponto $x^* \in \mathbb{R}^n$. Se x^* é um minimizador local de f, então a matriz Hessiana de f no ponto x^* é semidefinida positiva, isto é,*

$$d^T \nabla^2 f(x^*) d \geq 0, \ \ \forall d \in \mathbb{R}^n. \tag{2.3}$$

Prova. Como $d \in \mathbb{R}^n$, por Taylor,

$$f(x^* + td) = f(x^*) + t\nabla f(x^*)^T d + \frac{t^2}{2} d^T \nabla^2 f(x^*) d + r(t),$$

com $\lim_{t \to 0} \frac{r(t)}{t^2} = 0$. Por hipótese x^* é minimizador local, então o Teorema 2.4 garante que $\nabla f(x^*) = 0$. Portanto, para t suficientemente pequeno,

$$0 \leq f(x^* + td) - f(x^*) = \frac{t^2}{2} d^T \nabla^2 f(x^*) d + r(t).$$

Dividindo por t^2 e passando o limite quando $t \to 0$, temos $d^T \nabla^2 f(x^*) d \geq 0$. □

Teorema 2.5 ***(Condição suficiente de 2ª ordem)*** *Se x^* é um ponto estacionário da função f e $\nabla^2 f(x^*)$ é definida positiva, então x^* é minimizador local estrito de f.*

Prova. Seja λ o menor autovalor de $\nabla^2 f(x^*)$. Por hipótese temos que a matriz é definida positiva, logo $\lambda > 0$. Assumindo, $d^T \nabla^2 f(x^*) d \geq \lambda \|d\|^2$, $\forall d \in \mathbb{R}^n$. Por Taylor e também usando o fato de x^* ser estacionário, temos

$$f(x^* + d) = f(x^*) + \frac{1}{2} d^T \nabla^2 f(x^*) d + r(d) \geq f(x^*) + \frac{1}{2} \lambda \|d\|^2 + r(d),$$

onde $\lim\limits_{d \to 0} \frac{r(d)}{\|d\|^2} = 0$. Reescrevendo

$$\frac{f(x^* + d) - f(x^*)}{\|d\|^2} \geq \frac{\lambda}{2} + \frac{r(d)}{\|d\|^2}.$$

Como $\lim\limits_{d \to 0} \left(\frac{\lambda}{2} + \frac{r(d)}{\|d\|^2} \right) > 0$, então existe $\delta > 0$ tal que $\frac{\lambda}{2} + \frac{r(d)}{\|d\|^2} > 0$, $\forall d \in B(0, \delta) \backslash \{0\}$, donde segue que $f(x^* + d) - f(x^*) > 0$, $\forall d \in B(0, \delta) \backslash \{0\}$ ou equivalentemente,

$$f(x^*) < f(x), \quad \forall x \in B(x^*, \delta) \backslash \{x^*\}.$$

$\square$

3 CONVEXIDADE

A convexidade é um tópico muito importante na teoria de otimização. Veremos que, sob hipótese de convexidade na função objetivo, qualquer mínimo local torna-se um minimizador global e, além disso, temos que todo ponto estacionário é uma solução do problema.

Definição 3.1 *Um conjunto $A \subset \mathbb{R}^n$ é chamado de convexo se $\forall x, y \in A$ e $\alpha \in [0,1]$, tem-se que $\alpha x + (1-\alpha)y$ está completamente contido em A.*

O ponto $\alpha x + (1-\alpha)y$, onde $\alpha \in [0,1]$, chama-se combinação convexa de x e y (com parâmetro α).

Em outras palavras, podemos dizer que um conjunto é convexo se cada par de pontos dentro do conjunto, cada ponto no segmento de reta que une o par também está dentro do conjunto.

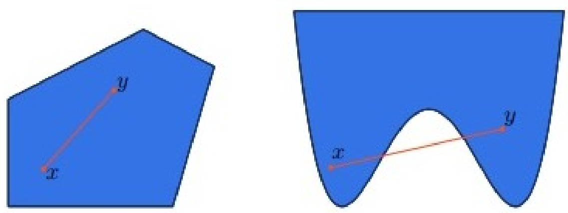


Figura 3.1: Conjuntos convexo e não convexo. Fonte: [11]

Definição 3.2 *Seja um conjunto convexo, $A \subset \mathbb{R}^n$. Dizemos que a função $f : \mathbb{R}^n \to \mathbb{R}$ é convexa em A, quando,*

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y),\ \forall x, y \in A \text{ e } \alpha \in [0,1].$$

O Teorema a seguir é muito útil para determinar quando uma função diferenciável é convexa ou não. Porém, antes vamos enunciar a seguinte definição.

Definição 3.3 *Dizemos que a função $f(x,y)$ é diferenciável no ponto (x_0, y_0) se as derivadas parciais $\dfrac{\partial f}{\partial x}(x_0,y_0)$ e $\dfrac{\partial f}{\partial y}(x_0,y_0)$ existem e se*

$$\lim_{(x,y)\to(x_0,y_0)} \frac{f(x,y) - f(x_0,y_0) - \dfrac{\partial f}{\partial x}(x_0,y_0)(x-x_0) - \dfrac{\partial f}{\partial y}(x_0,y_0)(y-y_0)}{\|(x,y)-(x_0,y_0)\|} = 0.$$

Dizemos que f é diferenciável num conjunto $A \subset D_f$, se f for diferenciável em todos os pontos de A.

Teorema 3.1 *Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função diferenciável e $A \subset \mathbb{R}^n$ um conjunto convexo. A função f é convexa, se, e somente se,*

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \; \forall x, y \in A.$$

Prova: Seja f convexa. Para quaisquer $x, y \in A$ e $\alpha \in (0, 1]$, definindo $d = y - x$, temos $x + \alpha d \in A$ e

$$f(x + \alpha d) = f((1 - \alpha)x + \alpha y) \leq (1 - \alpha)f(x) + \alpha f(y).$$

Assim,

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

Dividindo ambos os lados da desigualdade acima por $\alpha > 0$ e passando o limite quando $\alpha \to 0^+$, temos

$$f(y) - f(x) \geq \lim_{\alpha \to 0^+} \frac{f(x + \alpha d) - f(x)}{\alpha} = \nabla f(x)^T d = \nabla f(x)^T (y - x).$$

Resta provar a recíproca. Veja que,

$$f(x) \geq f(x + \alpha d) - \alpha \nabla f(x + \alpha d)^T d, \tag{3.1}$$

$$f(y) \geq f(x + \alpha d) + (1 - \alpha) \nabla f(x + \alpha d)^T d. \tag{3.2}$$

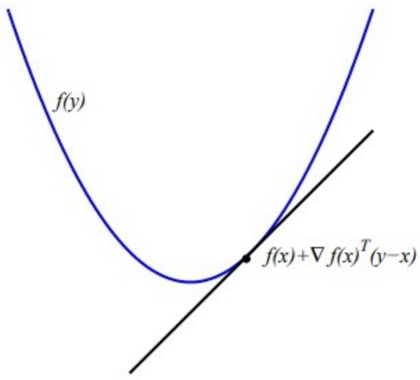
Multiplicando (3.1) por $(1 - \alpha)$ e (3.2) por α e somando obtemos,

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

Completando a demonstração. $\square$

Em particular, este teorema nos diz que o valor da função diferenciável convexa está sempre acima da sua aproximação linear.

Figura 3.2: Aproximação linear de f. Fonte: [4].



Na sequência verificaremos alguns resultados, os quais, vão nos auxiliar na demonstração do clássico Lema de Farkas. Este lema é fundamental na obtenção das condições de Karush-Kuhn-Tucker em problemas com restrições.

Lema 3.1 *Sejam* $u, v \in \mathbb{R}^n$ *com* $u \neq v$. *Se* $\|u\| = \|v\| = r$, *então* $\|(1-t)u + tv\| < r$, $\forall t \in (0, 1)$.

Prova: Seja $t \in (0, 1)$. Pela desigualdade triangular, temos que

$$\|(1-t)u + tv\| \leq (1-t)\|u\| + t\|v\| = r.$$

Vamos supor por absurdo, que $\|(1-t)u + tv\| = r$, então

$$(1-t)^2 u^T u + 2t(1-t)u^T v + t^2 v^T v = \|(1-t)u + tv\|^2 = 0.$$

Por hipótese temos $u^T u = v^T v = r^2$, o que implica, $u^T v = r^2$. Assim,

$$\|u - v\|^2 = u^T u - 2u^T v + v^T v = 0.$$

Logo $u = v$, o que é uma contradição. Conclui-se portanto que $\|(1-t)u + tv\|_2 < r$.

$\square$

Consideremos agora um conjunto $S \subset \mathbb{R}^n$, um ponto $x \in \mathbb{R}^n$ e o seguinte problema:

minimizar $\|s - x\|$

sujeito a $s \in S$.

Note que este problema pode não ter solução e quando tem, não conseguimos garantir a unicidade. Se assumirmos que $\varepsilon > 0$ e S uma bola aberta $B(a, \varepsilon) \subset \mathbb{R}^n$ e $x \notin S$, então concluímos que nenhum elemento $s \in S$ minimiza $\|s - x\|$. Além disso, se caso S for um conjunto fechado, a solução sempre existirá porém pode ser que não seja única. Uma projeção de x sobre S, denotada por $proj_S(x)$, busca encontrar um ponto de S que mais se aproxima de x.

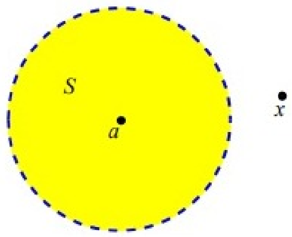


Figura 3.3: Nenhum elemento $s \in S$ minimiza $\|s - x\|$. Fonte: [4].

Na sequência, mostraremos que a unicidade da projeção de x sobre S, ocorre quando S é um conjunto fechado e convexo.

Lema 3.2 *Seja $S \subset \mathbb{R}^n$ um conjunto fechado e não vazio. Dado $z \in \mathbb{R}^n$, existe $\hat{z} \in S$ tal que*

$$\|z - \hat{z}\| \leq \|x - z\|, \forall x \in S.$$

Para maiores detalhes, veja [11].

Lema 3.3 *Seja $S \subset \mathbb{R}^n$ um conjunto convexo, fechado e não vazio. Dado $z \in \mathbb{R}^n$, existe um único $\hat{z} = proj_S(z) \in S$ tal que*

$$\|z - \hat{z}\| \leq \|x - z\|, \forall x \in S.$$

Prova: A existência é garantida pelo Lema 3.2. Vejamos agora a unicidade, para isto, admita que existam $\hat{z} \neq \tilde{z}$ em S tais que, para todo $x \in S$.

$$\|z - \hat{z}\| \leq \|z - x\| \quad \text{e} \quad \|z - \tilde{z}\| \leq \|z - x\|. \tag{3.3}$$

Assim, tome $x = \tilde{z}$ na primeira desigualdade e $x = \hat{z}$ na segunda, temos

$$\|z - \hat{z}\| = \|z - \tilde{z}\|.$$

Porém, sendo $z - \hat{z} \neq z - \tilde{z}$, pelo Lema 3.1, com $r = \|z - \hat{z}\| = \|z - \tilde{z}\|$ e para $t \in (0, 1)$,

$$\|(1-t)(z - \hat{z}) + t(z - \tilde{z})\| < r.$$

Por hipótese S é convexo, temos $\bar{z}$, tomando $t = \dfrac{1}{2}$ e $\bar{z} = \dfrac{1}{2}(\hat{z} + \tilde{z})$, obtemos

$$\|z - \bar{z}\| = \|z - \frac{1}{2}(\hat{z} + \tilde{z})\| = \|\frac{1}{2}(z - \hat{z}) + \frac{1}{2}(z - \tilde{z})\| < \alpha.$$

Portanto, contradiz (3.3). □

Teorema 3.2 ***(Teorema da Projeção)*** *Considere um conjunto convexo fechado $S \subset \mathbb{R}^n$, $z \in \mathbb{R}^n$ e $\hat{z} = proj_S(z)$, então*

$$(x - \hat{z})^T (z - \hat{z}) \leq 0, \forall x \in S.$$

Prova: Dado $t \in (0, 1]$ arbitrário e um ponto de S como sendo $\hat{z} + t(x - \hat{z})$. Por definição de $\hat{z}$, temos que $\|z - \hat{z}\| \leq \|\hat{z} + t(x - \hat{z}) - z\|$.

Assim,

$$\|z - \hat{z}\|^2 \leq \|(\hat{z} - z) + t(x - \hat{z})\|^2.$$

Note que

$$\|(\hat{z} - z) + t(x - \hat{z})\|^2 = \|z - \hat{z}\|^2 + 2t(\hat{z} - z)^T(x - \hat{z}) + t^2\|x - \hat{z}\|^2.$$

Logo,

$$\|z - \hat{z}\|^2 \leq \|z - \hat{z}\|^2 + 2t(\hat{z} - z)^T(x - \hat{z}) + t^2\|x - \hat{z}\|^2.$$

Como $t > 0$, temos que $(x - \hat{z})^T(z - \hat{z}) \leq t\|x - \hat{z}\|^2$ e passando o limite $t \to 0$, conclui-se que

$$(x - \hat{z})^T(z - \hat{z}) \leq 0.$$

□

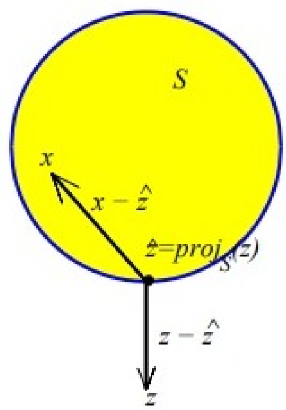


Figura 3.4: Ilustração do Teorema da Projeção. Fonte: [4].

4 CONES

Neste capítulo, focaremos o nosso estudo em alguns tópicos gerais da teoria de cones, pois são de extrema importância para obtermos as condições de otimalidade de Karush-Kuhn-Tucker e também na obtenção do Lema de Farkas, que será estabelecido em duas versões: geométrico e algébrico.

Iniciamos com as seguintes definições:

Definição 4.1 ***(Cone)*** *Seja $K \subset \mathbb{R}^n$ um subconjunto não vazio, é dito ser um cone quando,* $\forall t \geq 0$ *e* $d \in K$*, tem-se* $td \in K$*.*

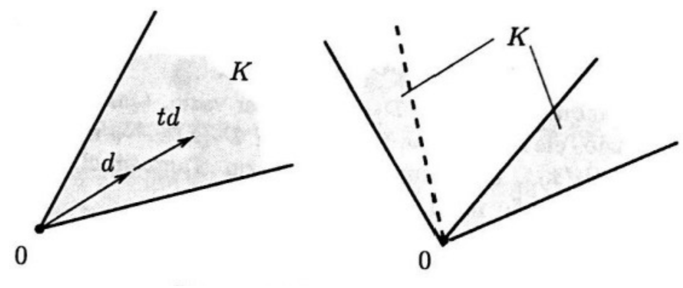


Figura 4.1: Exemplos de cones. Fonte: [6].

Definição 4.2 ***(Cone Polar)*** *Dado um conjunto $S \subset \mathbb{R}^n$, definimos o polar de S por*

$$P(S) = \{p \in \mathbb{R}^n | p^T x \leq 0, \forall x \in S\}.$$

Note que o polar de um conjunto S é um cone formado por todos os vetores que formam ângulo reto ou obtuso com elementos arbitrários de S.

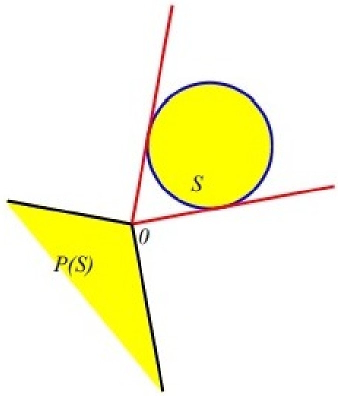


Figura 4.2: Polar de um conjunto S. Fonte: [4].

Lema 4.1 *Dado $S \subset \mathbb{R}^n$, $P(S)$ é cone, convexo e fechado.*

Prova: Por definição, para $t \geq 0$ e $d \in P(S)$ temos $(td)^T x = t(d^T x) \leq 0, \forall x \in S$. Assim, $td \in P(S)$, logo $P(S)$ é um cone. Vamos verificar a convexidade, dado $u, v \in P(S)$ e $t \in [0, 1]$. Para qualquer $x \in S$ temos que S

$$((1-t)u + tv)^T x = (1-t)u^T x + tv^T x \leq 0.$$

Com isso $(1-t)u + tv \in P(S)$ prova que $P(S)$ é convexo.
Agora basta mostrar que $P(S)$ é fechado. Para isso, considere uma sequência $(d^k \subset P(S)$ com $d^k \to d$. Dado $x \in S$, temos $(d^k)^T x \leq 0$, logo $d^T x \leq 0$. Conclui-se que $d \in P(S)$. ☐

Lema 4.2 *Sejam $A, B \subseteq \mathbb{R}^n$. Se $A \subseteq B$, então $P(B) \subseteq P(A)$.*

Prova: Dado $y \in P(B)$ arbitrário, então $y^T x \leq 0, \forall x \in B$. Como $A \subseteq B$, segue que $y^T z \leq 0, \forall z \in A$, assim, $y \in P(A)$. ☐

Teorema 4.1 *Considere $C \subset \mathbb{R}^n$, um cone convexo fechado, $z \in \mathbb{R}^n$ e $\bar{z} = proj_c(z)$. Então, $z - \bar{z} \in \{\bar{z}\}^{\perp} \cap P(C)$.*

Para demonstração, veja [11]

Lema 4.3 *Se $S \subset \mathbb{R}^n$, então $S \subset P(P(S))$.*

Prova. Admitindo $x \in S$ e $C = P(S)$. Tome $d \in C$, então $x^T d \leq 0$. Portanto, $x \in P(C) = P(P(S))$. ☐

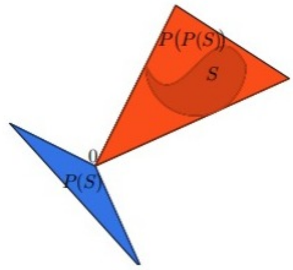


Figura 4.3: O polar do polar de S. Fonte: [11].

Note que, aplicando o polar duas vezes nem sempre nos fornece o conjunto original. Neste caso, tais situações que impedem a igualdade entre o duplo cone e o conjunto é: não ser cone, não ser convexo ou não ser fechado.

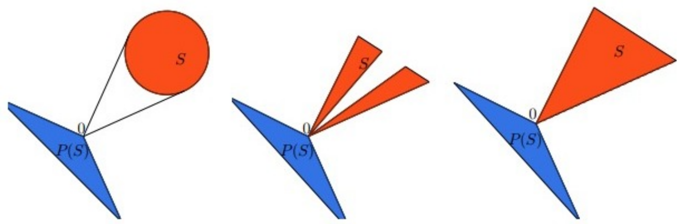


Figura 4.3: situações onde não satisfaz $S = P(P(S))$. Fonte: [11].

Nessa perspectiva, veremos na próxima seção que, com algumas hipóteses adicionais sobre o conjunto S, o Lema de Farkas garante a igualdade $P(P(S)) = S$.

4.1 Teoremas de Alternativa

A seguir avaliaremos dois sistemas de equações e/ou inequações lineares que se relacionam no seguinte sentido: somente um dos dois sistemas tem solução (nunca, ambos). Em virtude, chamamos esses teoremas de Teoremas de Alternativa. Nesta seção, nos atentamos apenas em demonstrar o Lema de Farkas.Lema de Farkas. Sem esquecer que existem vários outros Teoremas de Alternativa na literatura, tais como, Motzkin, Key, Tucker, Gale I, Gale II e de Stiemke. Para maiores detalhes, veja [2 - 4].

Sobretudo, estamos interessados em determinar condições para que a igualdade entre o polar do polar de S e o conjunto S, $P(P(S)) = S$, seja satisfeita.

Como resultado, apresentamos duas versões do Lema de Farkas que garantem essa igualdade.

Lema 4.4 ***(Farkas Geométrico)*** *Seja C um cone convexo fechado não vazio, então $P(P(C)) = C$.*

Prova. Pelo Lema 4.3, segue que $C \subset P(P(C))$. Assim, resta mostrar que $P(P(C)) \subset C$. Considere então $z \in P(P(C))$ e $\hat{z} = proj_C(z)$. Desse modo, iremos provar que $z = \hat{z}$.
Pelo Teorema da Projeção 3.2,

$$(z - \hat{z})^T (x - \hat{z}) \leq 0, \forall x \in C.$$

Por hipótese, C é um cone e é fechado, então $x = 0$ e $x = 2\hat{z}$ são elementos de C. Temos $z \in P(P(C))$, logo $-\hat{z}^T (z - \hat{z}) \leq 0$ e $\hat{z}^T (z - \hat{z})^T \leq 0$. Então,

$$\hat{z}^T (z - \hat{z}) = 0. \tag{4.1}$$

Por outro lado,

$$(z - \hat{z})^T x - (z - \hat{z})^T \hat{z} = (z - \hat{z})^T - (x - \hat{z}) \leq 0. \tag{4.2}$$

Comparando (4.1) e (4.2), conclui-se que

$$(z - \hat{z})^T x \leq 0$$

para todo $x \in C$. E por definição de cone polar, temos que $(z - \hat{z}) \in P(C)$. Como $z \in P(P(C))$, vale $(z - \hat{z})^T z \leq 0$. Usando o Teorema 4.1, obtemos

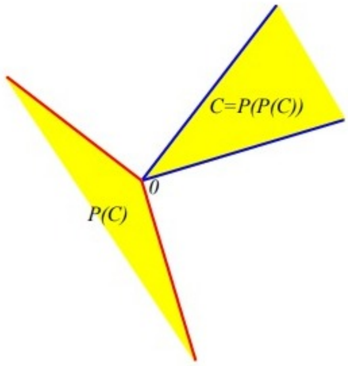
$$\|z - \hat{z}\|^2 = (z - \hat{z})^T z - (z - \hat{z})^T \hat{z} = (z - \hat{z})^T z \leq 0.$$

Assim $z = \hat{z} \in C$. $\square$

Figura 4.3: Ilustração do Lema 4.4. Fonte: [4].

Antes de vermos a outra versão do Lema 4.4. É necessário enunciarmos uma proposição muito importante no que se refere ao cone gerado por um conjunto finito de vetores cujo a prova pode ser encontrada em [4].

Definição 4.3 *Dizemos que um cone $C \subset \mathbb{R}^n$ tem um número finito de geradores ou é finitamente gerado, quando existe $a_i \in \mathbb{R}^n$, $i = 1, .., p$, tais que*



$$C = \{x \in \mathbb{R}^n | x = \sum_{i=1}^{p} \mu_i a_i,\ \mu_i \geq 0,\ i = 1, ...p\}.$$

Proposição 4.1 Sejam $n, m \in \mathbb{N}$ e $A \in \mathbb{R}^{n\times m}$ e $B \in \mathbb{R}^{n\times m}$. Se

i. $C = \{Ay \mid y \in \mathbb{R}^m, y \geq 0\}$, então C é um cone fechado.

ii. $C = \{B^T y; y \geq 0\}$, então C é um cone convexo e fechado.

Lema 4.5 ***(Farkas Algébrico)*** *Considere $A \in \mathbb{R}^{m\times n}$ e $c \in \mathbb{R}^n$. Então, exatamente um dos dois sistemas abaixo tem solução*

$(S1)$ $\qquad Ax \leq 0$ e $c^T x \geq 0,$

ou

$(S2)$ $\qquad A^T y = c$ e $y \geq 0.$

Prova. Considere $C = \{A^T y; y \geq 0\}$ e suponha que $c \in C$, i.e. que o sistema $(S2)$ tem solução, então $c = A^T y$ com $y \geq 0$. Considere $x \in \mathbb{R}^n$ tal que $Ax \leq 0$, assim temos $c^T x = y^T Ax \leq 0$, logo $(S1)$ não tem solução.
Agora vamos supor que $c \notin C$, ou seja, que $(S2)$ não tenha solução. Pelo Lema 4.4 e pela Proposição 4.1, temos que $C = P(P(C))$, logo existe um $\bar{x} \in P(C)$ tal que $c^T \bar{x} > 0$. De modo que,

$$\bar{x}^T A^T y \leq 0.$$

Assim,

$$A\bar{x} \leq 0 \qquad \forall y \leq 0.$$

Portanto, $(S1)$ tem solução. $\square$

De maneira equivalente, pela Definição 4.3 é possível estabelecer uma relação entre as duas versões do Lema de Farkas, quando C é um cone gerado e fechado. Para maiores detalhes, veja [6].

5 OTIMIZAÇÃO COM RESTRIÇÕES

Neste capítulo, nosso principal objetivo é discutir as condições de otimalidade para o problema de programação não linear sujeito a restrições de igualdade e desigualdade que consiste em

$$\begin{aligned} &Minimizar \;\; f(x) \\ &Sujeito \;\; a \;\;\; h(x) = 0 \\ &\qquad\qquad\quad\;\; g(x) \leq 0, \end{aligned} \tag{P}$$

onde $f : \mathbb{R}^n \rightarrow \mathbb{R}, \;\; h : \mathbb{R}^n \rightarrow \mathbb{R}^m \;\; e \;\; g : \mathbb{R}^n \rightarrow \mathbb{R}^p$, são funções continuamente diferenciáveis. Consideramos o conjunto viável (ou factivel) do Problema P como sendo o seguinte conjunto

$$\Omega = \{x \in \mathbb{R}^n | h(x) = 0, g(x) \leq 0\}.$$

Definição 5.1 *Seja $\bar{x} \in \Omega$ o conjunto dos pontos viáveis. Uma restrição de desigualdade $g_j(x) \leq 0$ é dita ativa num ponto $\bar{x}$, se $g_j(\bar{x}) = 0$ e inativa em $\bar{x}$, se $g_j(\bar{x}) < 0$.*

Denotamos por $I(\bar{x})$ o conjunto de índices das restrições de desigualdade ativas em um ponto viável $\bar{x}$, isto é,

$$I(\bar{x}) = \{j = 1, \cdots, p | g_j(\bar{x}) = 0\}.$$

5.1 Direções Viáveis e de Descida

Definição 5.2 *Considere $\bar{x} \in \Omega$ e $d \in \mathbb{R}^n$, dizemos que d é uma direção viável em relação ao conjunto Ω no ponto $\bar{x}$, quando existe $\delta > 0$ tal que*

$$\bar{x} + td \in \Omega, \forall t \in [0, \delta]$$

Denotamos por $V(\bar{x})$ o conjunto de todas as direções viáveis em relação ao conjunto Ω no ponto $\bar{x} \in \Omega$.

Definição 5.3 *Um vetor $d \in \mathbb{R}^n$ é chamada de direção de descida da função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ no ponto $\bar{x} \in \mathbb{R}^n$, se existe $\delta > 0$ tal que*

$$f(\bar{x} + td) < f(\bar{x}), \forall t \in (0, \delta].$$

Denotamos por $F(\bar{x})$ o conjunto de todas as direções de descida da função f no ponto $\bar{x}$. De modo geral, quando obtemos pequenos passos do ponto $\bar{x}$ ao longo de $d \in F(\bar{x})$ nos fornecem pontos que possuem menor valor na função objetivo f.

Teorema 5.1 *Se $\nabla f(\bar{x})^T d < 0$, então d é uma direção de descida para f, a partir de $\bar{x}$.*

Prova: Sendo $\nabla f(\bar{x})^T d = \dfrac{\partial f}{\partial d}(\bar{x}) = \lim\limits_{t \to 0} \dfrac{f(\bar{x} + td) - f(\bar{x})}{t}$. Por hipótese e pela preservação do sinal, existe $\delta > 0$ tal que

$$\frac{f(\bar{x} + td) - f(\bar{x})}{t} < 0, \forall t \in (-\delta, \delta), t \neq 0.$$

Assim, $f(\bar{x} + td) < f(\bar{x})$, $\forall t \in (0, \delta)$, completando a demonstração. $\square$

Observe que a Figura 5.1 ilustra um caso em que d é uma direção de descida e que além disso, forma um ângulo obtuso com $\nabla f(\bar{x})$, isto é, $\nabla f(\bar{x})^T d < 0$.

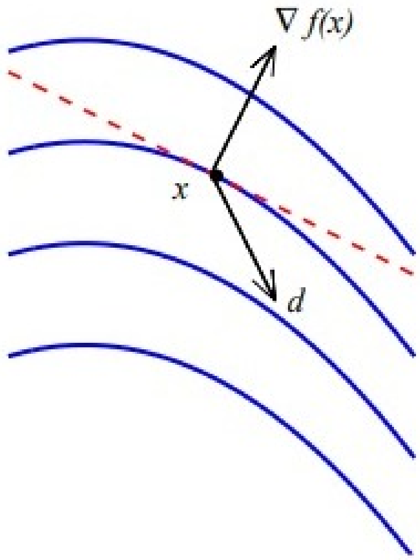


Figura 5.1: Ilustração do Teorema 5.1. Fonte: [4].

Na sequência, enunciaremos aproximações do conjunto viável em termos dos cones viável linearizado e tangente, bem como suas relações.

5.2 Cone Viável Linearizado

Nesta seção apresentaremos uma linearização do conjunto viável em torno de um ponto. Indicaremos por $\nabla h_i(\bar{x})$ a transposta da jacobiana de h no ponto $\bar{x}$.

Definição 5.4 *Dado $\bar{x} \in \Omega$, definimos o cone viável linearizado de Ω em torno de $\bar{x}$, denotado por $D(\bar{x})$, como*

$$D(\bar{x}) = \{d \in \mathbb{R}^n | \nabla h_i(\bar{x})^T d = 0, \, i = 1, \cdots, m \quad \text{e} \quad \nabla g_j(\bar{x})^T d \leq 0, \forall j \in I(\bar{x})\} \; .$$

Na sequência, outro cone referente ao problema (P) é o cone gerado pelos gradientes das restrições em $\bar{x} \in \Omega$ definido por

$$G(\bar{x}) = \left\{ \sum_{i=1}^{m} \lambda_i \nabla h_i(\bar{x}) + \sum_{j \in I(\bar{x})} \mu_j \nabla g_j(\bar{x}) | \mu_j \geq 0, \forall j \in I(\bar{x}) \right\}. \tag{5.1}$$

Lema 5.1 *Seja $G(\bar{x})$ definido pela Equação (5.1). Então $G(\bar{x})$ é um cone convexo e fechado.*

Para maiores detalhes, veja [4].

Como $G(\bar{x})$ é um cone convexo e fechado segue, pelo Lema 4.4, que $P(P(G(\bar{x}))) = G(\bar{x})$. Sendo assim, podemos estabelecer uma relação entre os cones $D(\bar{x})$ e $G(\bar{x})$.

Lema 5.2 *Seja $\bar{x} \in \Omega$, temos que $D(\bar{x}) = P(G(\bar{x}))$.*

Dado $d \in D(\bar{x})$ e $s \in G(\bar{x})$. Temos que

$$d^T s = \sum_{i=1}^{m} \lambda_i d^T \nabla h_i(\bar{x}) + \sum_{j \in I(\bar{x})} \mu_j d^T \nabla g_j(\bar{x}). \tag{5.2}$$

Como $d \in D(\bar{x})$, segue que $d^T \nabla h_i(\bar{x}) = 0 \; \forall i = 1, ..., m$ e $d^T \nabla g_j(\bar{x}) \leq 0 \; \forall j \in I(\bar{x})$. Assim, $\mu_j \geq 0$ resulta de (5.2) que $d^T s \leq 0 \; \forall s \in G(\bar{x})$. Por definição de cone polar, temos que $d \in P(G(\bar{x}))$, isto é, $d^T s \leq 0, \, \forall \in G(\bar{x})$.
Por outro lado, visto que $\nabla h_i(\bar{x})$ e $-\nabla h_i(\bar{x})$ são elementos de $G(\bar{x})$, $\forall i = 1, ..., m$, donde segue que $d^T \nabla h_i(\bar{x}) \leq 0$ e $d^T(-\nabla h_i(\bar{x})) \leq 0$. Portanto, $d^T \nabla h_i(\bar{x}) = 0$, $\forall i = 1, .., m$. Além disso, como $\nabla g_j(\bar{x}) \in G(\bar{x})$, $\forall j \in I(\bar{x})$, temos $d^T \nabla g_j(\bar{x}) \leq 0$, o que completa a demonstração. $\square$

Com efeito aos Lemas 4.4 e 5.1, torna-se possível reformular o Lema 5.2, da seguinte maneira

$$P(D(\bar{x})) = G(\bar{x}). \tag{5.3}$$

Dessa forma, estabelecemos uma ferramenta fundamental para a demonstração das condições de KKT.

5.3 Cone Tangente

Dedicamos a nossa atenção em um cone extremamente importante, o qual caracteriza uma melhor aproximação do conjunto viável em torno de um determinado ponto. Usaremos este cone para obtermos uma relação entre as condições de otimalidade e as condições de qualificação.

Definição 5.5 *Uma direção $d \in \mathbb{R}^n$ é tangente a $\Omega \subset \mathbb{R}^n$ a partir de $\bar{x} \in \Omega$, se e somente se, existe uma sequência de pontos viáveis $(x^k) \in \Omega$ tal que $x^k \to \bar{x}$ e*

$$\frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to \frac{d}{\|d\|}. \tag{5.4}$$

Procede diretamente da definição que se d é tangente, o mesmo vale para td, qualquer que seja $t \geq 0$. Sendo assim, o conjunto formado pelos vetores tangente a Ω em $\bar{x}$ é um cone, chamado de cone tangente a Ω no ponto $\bar{x}$ e denotado por $T(\bar{x})$.

A seguir mostraremos um exemplo

Exemplo 5.1 *Considere as funções g_i , g_j : $\mathbb{R}^n \to \mathbb{R}$ dadas por $g_i(x) = x_1^2 - 2x_1 - x_2$ e $g_j(x) = x_1^2 - 2x_1 + x_2$. Determine o cone $T(\bar{x})$, associado ao conjunto viável $\Omega = \{x \in \mathbb{R}^2$ $\mid g(x) \leq 0\}$ em torno do ponto $\bar{x} = 0$.*

Sejam $x^k = \begin{pmatrix} s_k \\ t_k \end{pmatrix}$ um sequência de pontos de Ω e $d = \begin{pmatrix} s_k \\ t_k \end{pmatrix} \in \mathbb{R}^2$ tais que $x^k \to \bar{x}$ e

$$\frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to \frac{d}{\|d\|}. \tag{5.5}$$

Prova. Provaremos que $-2d_1 \leq d_2 \leq 2d_1$. Sabemos que $x^k \in \Omega$, logo $s_k^2 - 2s_k \leq t_k \leq 2s_k - s_k^2$. Assim,

$$\frac{s_k^2 - 2s_k}{\sqrt{s_k^2 + t_k^2}} \leq \frac{t_k}{\sqrt{s_k^2 + t_k^2}} \leq \frac{2s_k - s_k^2}{\sqrt{s_k^2 + t_k^2}}. \tag{5.6}$$

De (5.5), concluímos que $s_k \to 0$,

$$\frac{s_k}{\sqrt{s_k^2 + t_k^2}} \to \frac{d_1}{\|d\|} \quad \text{e} \quad \frac{t_k}{\sqrt{s_k^2 + t_k^2}} \to \frac{d_2}{\|d\|}.$$

Passando o limite na relação (5.6), temos

$$\frac{-2d_1}{\|d\|} \leq \frac{d_2}{\|d\|} \leq \frac{2d_1}{\|d\|}.$$

Segue então

$$T(\bar{x}) \subset \{d \in \mathbb{R}^2 | -2d_1 \leq d_2 \leq 2d_1\}.$$

Agora vamos provar a inclusão contrária, tome $d = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ e considere

$$s_k = \frac{1}{k} \quad , \quad t_k = 2s_k^2 - s_k^2 \quad \text{e} \quad x^k = \begin{pmatrix} s_k \\ t_k \end{pmatrix}.$$

Logo, $x^k \to \bar{x}$,

$$\frac{s_k}{\sqrt{s_k^2 + t_k^2}} = \frac{1}{\sqrt{1 + (2 - s_k)^2}} \to \frac{1}{\sqrt{5}} \quad \text{e} \quad \frac{t_k}{\sqrt{s_k^2 + t_k^2}} \to \frac{2}{\sqrt{5}}.$$

Portanto, $\dfrac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to \dfrac{d}{\|d\|}$. Por outro lado, considere $d = \begin{pmatrix} 1 \\ \gamma \end{pmatrix}$, com $\gamma \in [0, 2)$. Para todo $k \in \mathbb{N}$, suficientemente grande, temos $\gamma < 2 - \dfrac{1}{k}$, implicando em $y^k = \dfrac{1}{k}\begin{pmatrix} 1 \\ \gamma \end{pmatrix} \in \Omega$. Além disso,

$$y^k \to \bar{x} \quad \text{e} \quad \frac{y^k - \bar{x}}{\|y^k - \bar{x}\|} \to \frac{d}{\|d\|}.$$

Como $T(\bar{x})$ é um cone, então todo vetor $d \in \mathbb{R}^2$ tal que $0 \leq d_2 \leq 2d_1$ é tangente. Segue o caso $-2d_1 \leq d_2 \leq 0$ analogamente. Assim, obtemos

$$T(\bar{x}) = \{d \in \mathbb{R}^2 | -2d_1 \leq d_2 \leq 2d_1\}.$$

Completando a provar. □

5.4 Relação entre o Cone Viável Linearizado e o Cone Tangente

O lema a seguir vai nos mostrar que qualquer direção d pertencente ao cone tangente, também pertence ao cone viável linearizado.

Lema 5.3 *Seja $\bar{x} \in \Omega$. Então $T(\bar{x}) \subset D(\bar{x})$.*

Prova. Considerando $d \in T(\bar{x})$, $d \neq 0$. Pela de Definição 5.5 existe uma sequência viável (x^k) tal que $x^k \neq \bar{x}$, $x^k \to \bar{x}$ e $\dfrac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to \dfrac{d}{\|d\|}$.

A expansão de Taylor em torno de $\bar{x}$ fornece

$$0 = h(x^k) = h(\bar{x}) + \nabla h(\bar{x})^T (x^k - \bar{x}) + o(\|x^k - \bar{x}\|).$$

Como $x^k \neq \bar{x}$, podemos dividir esta expressão por $\|x^k - \bar{x}\|$ e sabendo que $h(\bar{x}) = 0$ temos:

$$\nabla h(x^k)^T \frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} + \frac{o(\|x^k - \bar{x}\|)}{\|x^k - \bar{x}\|} = 0.$$

Passando o limite quando $k \to \infty$, obtemos

$$\nabla h(\bar{x})^T \frac{d}{\|d\|} = 0. \tag{5.7}$$

Para $j \in I(\bar{x})$, temos

$$g_j(\bar{x}) + \nabla g_j(\bar{x})^T (x^k - \bar{x}) + o(\|x^k - \bar{x}\|) \leq 0.$$

De modo análogo obtemos

$$\nabla g_j(\bar{x})^T \frac{d}{\|d\|} = 0. \tag{5.8}$$

Portanto, $d \in D(\bar{x})$. $\square$

5.5 Condições de Otimalidade de Karush Kuhn Tucker

Nesta seção, vamos estudar as condições necessárias de otimalidade para o problema (P). Enunciaremos o importante e clássico Teorema de KKT, visto que ele nos permite escrever o gradiente da função objetivo por meio de uma combinação linear dos gradientes das restrições de igualdade e desigualdades ativas, de modo que o papel da polaridade dos cones é quem garante que o multiplicador associado ao gradiente da f seja não nulo.

Lema 5.4 *Considere $x^* \in \Omega$. Se x^* é um minimizador local do problema (P), então $\nabla f(x^*)^T d \geq 0$, $\forall d \in T(x^*)$.*

Prova. Dado $d \in T(x^*)$, $d \neq 0$. Então existe uma sequência $(x^k) \subset \Omega$, de modo que, $x^k \to x^*$ e $\dfrac{x^k - \bar{x}}{\|x^k - \bar{x}\|} \to \dfrac{d}{\|d\|}$. Por outro lado, temos

$$0 \leq f(x^k) - f(x^*) = \nabla f(x^*)^T (x^k - x^*) + o(\|x^k - x^*\|),$$

para todo k suficientemente grande. Dividindo por $\|x^k - x^*\|$ e passando o limite obtemos $\nabla f(x^*)^T d \geq 0$. □

O Lema 5.4 é pouco prático, no sentido de que não podemos usá-lo para calcular os possíveis minimizadores. Neste sentido, o teorema que se segue, nos garante esta possibilidade. Este é o resultado central de nosso estudo.

Teorema 5.2 ***(Condições de KKT)*** *Considere $x^* \in \Omega$. Se x^* é um minimizador local do problema (P). Se $P(T(x^*)) = P(D(x^*))$, então existem vetores λ^* em $\mathbb{R}^m$ e μ^* em $\mathbb{R}^p$ tais que*

$$-\nabla f(x^*) = \sum_{i=1} \lambda_i^* \nabla h_i(x^*) + \sum_{j \in I} \mu_j^* \nabla g_j(x^*),$$

$$\mu_j^* \geq 0, \;\; j \in I,$$

$$\mu_j^* g_j(x^*) = 0, \;\; j \in I.$$

Prova. Seja x^* um minimizador local do Problema (P). Segue pelo Lema 5.4, que

$$-\nabla f(x^*)^T d \leq 0, \forall d \in T(x^*).$$

Assim, usando a hipótese e a relação (5.3), obtemos

$$-\nabla f(x^*) \in P(T(x^*)) = P(D(x^*)) = G(x^*).$$

Portanto, existem vetores λ e μ, tais que $\mu \geq 0$ e

$$-\nabla f(x^*) = \sum_{i=1} \lambda_i \nabla h_i(x^*) + \sum_{j \in I(x^*)} \mu_j \nabla g_j(x^*).$$

Definindo $\mu_j^* = \begin{cases} \mu_j, & \forall j \in I(x^*) \\ 0, & \text{caso contrário} \end{cases}$

e $\lambda^* = \lambda$, completando a prova. $\square$

Ressaltamos que o teorema anterior, possui uma hipótese sobre os cones $T(x^*)$ e $D(x^*)$. Esta hipótese é na verdade uma condição de qualificação, isto é, uma condição necessária e suficiente para que se obtenha KKT. Entretanto, pode ser muito difícil de se obter os cones $T(x^*)$ e $D(x^*)$ e ainda verificar se a condição $P(T(x^*)) = P(D(x^*))$ é satisfeita. Portanto, veremos no próximo capítulo outras condições de qualificações que também implicam em $P(T(x^*)) = P(D(x^*))$ e mais simples de serem verificadas.

Exemplo 5.2 *Considere $f : \mathbb{R}^2 \to \mathbb{R}$, $g : \mathbb{R}^2 \to \mathbb{R}^2$ definidas no problema abaixo e $x^* = (1,0)^T$ um minimizador local.*

$$\begin{aligned} Minimizar \;\; & f(x) = -x_1 \\ Sujeito \;\; a \;\; & g_1(x) = x_2 - (1-x_1)^3 \leq 0 \\ & g_2(x) = -x_2 \leq 0 \end{aligned}$$

Note que

$$\nabla f(x^*) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \qquad \nabla g_1(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{e} \quad \nabla g_2(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Assim a condição

$$\nabla f(x^*) + \mu_1 \nabla g_1(x^*) + \mu_2 \nabla g_2(x^*) = 0,$$

$$\begin{pmatrix} -1 \\ 0 \end{pmatrix} + \mu_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \mu_2 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$\begin{cases} -1 + 0\mu_1 + 0\mu_2 = 0 & \Rightarrow \quad -1 = 0 \quad \text{absurdo!} \\ 0 + \mu_1 - \mu_2 = 0 & \Rightarrow \quad \mu_1 = \mu_2. \end{cases}$$

Logo não existe, pois os vetores são linearmente dependente, isto é,

$$\begin{pmatrix} 0 \\ -1 \end{pmatrix} = -1 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Portanto as condições KKT não se verificam. $\square$

A Figura 5.2 ilustra a região viável.

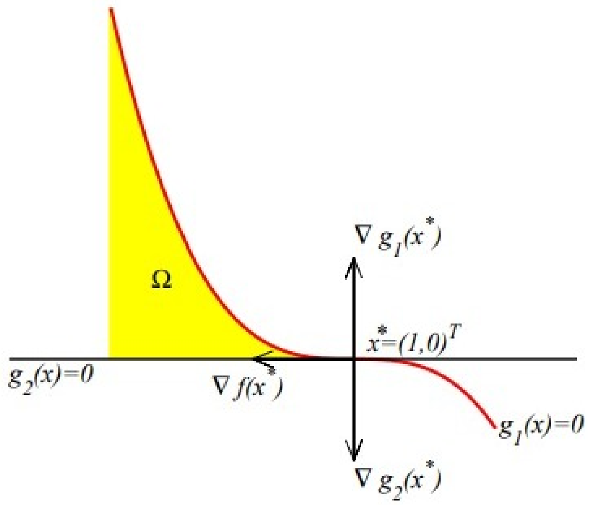


Figura 5.2. Fonte: [4].

6 CONDIÇÕES DE QUALIFICAÇÕES

Uma condição de qualificação pode ser entendida como uma propriedade que garante que as condições de KKT sejam válidas num minimizador local. As condições de qualificações são ferramentas úteis na análise de convergência de métodos de otimização.

Discutiremos neste capítulo diferentes condições de qualificações, as quais podem ser encontradas por exemplo em, [1, 2, 4, 6, 11]. Observemos que, quanto mais fraca as condições de qualificação mais forte será a condição de otimalidade associada. Neste sentido é que apresentaremos a condições de qualificação que se seguem. Contudo, vamos apresentar uma definição precisa de condição de qualificação.

Definição 6.1 *Dizemos que as restrições $h(x) = 0$ e $g(x) \leq 0$ cumprem uma condições de qualificação em $x^* \in \Omega$ quando, dada qualquer função diferenciável f, que tenha mínimo em x^*, relativamente a Ω, sejam satisfeitas as condições de otimalidade de KKT.*

A princípio, abordamos uma situação particular em que as restrições do problema são lineares. Sendo assim, as condições de KKT são válidas em todos pontos de mínimo do problema.

6.1 Problemas com Restrições Lineares

Em programação linear o principal interesse é minimizar uma função f onde as restrições são formadas por igualdades e desigualdades lineares. Desta maneira, considere o seguinte problema:

$$
\begin{aligned}
Minimizar \;\; & f(x) \\
Sujeito \; a \;\; & Mx = c \\
& Ax \leq b
\end{aligned}
\qquad (PL)
$$

onde $A \in \mathbb{R}^{t \times n}$, $M \in \mathbb{R}^{u \times n}$, $b \in \mathbb{R}^t$ e $c \in \mathbb{R}^n$.

Seguidamente, veremos no próximo teorema que, se o conjunto viável é formado apenas por restrições lineares, então as condições de otimalidade de KKT se verificam em um minimizador.

Teorema 6.1 *Se x^* é um minimizador local do problema (PL), então x^* satisfaz as condições de KKT.*

Prova. Fazendo o uso do Lema 5.3 e do Teorema 5.2, vamos provar que $D(x^*) \subset T(x^*)$. Dado $d \in D(x^*)$, temos $Ad = 0$ e $Md \leq 0$. Assim, se $d = 0$, obtemos trivialmente $d \in T(x^*)$. Por outro lado, caso $d \neq 0$, defina $x^k = x^* + \frac{1}{k}d$. Então $x^k \to x^*$,

$$Mx^k = c, Ax^k \leq b \quad \text{e} \quad \frac{x^k - x^*}{\|x^k - x^*\|} = \frac{d}{\|d\|}.$$

Portanto, segue que $d \in T(x^*)$. $\square$

6.2 Condição de Qualificação de Slater

Em muitos problemas de otimização, o conjunto viável é um conjunto convexo. É comum na literatura que um minimizador local torna-se um minimizador global, isto significa que as condições KKT é uma condição necessária e suficiente. Deste modo, definiremos a seguir a condição de qualificação de Slater.

Definição 6.2 *Dizemos que a condição de qualificação de Slater é satisfeita quando h é linear, g é convexa e existe $\tilde{x} \in \Omega$ tal que*

$$h(\tilde{x}) = 0 \quad \text{e} \quad g(\tilde{x}) < 0.$$

É fácil ver que esta condição de qualificação requer a existência de um ponto no interior relativo do conjunto viável.

6.3 Condição de Qualificação de Independência Linear - LICQ

Ressaltamos que uma condição de qualificação bem conhecida na literatura é a independência linear dos gradientes das restrições de igualdade e de desigualdade ativas.

Definição 6.3 *Dizemos que a condição de qualificação de independência linear é satisfeita em $\bar{x}$ quando o conjunto formado pelos gradientes das restrições de igualdades e das restrições de desigualdades ativas é linearmente independente, isto é,*

$$\{\nabla h_i(\bar{x})\}_{i\in\{1,\ldots,m\}} \cup \{\nabla g_j(\bar{x})\}_{j\in I(\bar{x})} \quad \text{é} \quad LI. \tag{6.1}$$

Veja que esta condição é mais simples de ser verificada do que daquela hipótese do Teorema 5.2, envolvendo cones. Para tornar mais convincente, retomamos as restrições estabelecidas ao Exemplo 5.1, onde apenas a determinação do cone tangente $T(\bar{x})$ ja foi trabalhosa.

Exemplo 6.1 *Considere duas restrições de desigualdades definidas por* $g_1, g_2 : \mathbb{R}^n \to \mathbb{R}$*, onde* $g_1(x) = x_1^2 - 2x_1 - x_2$ *e* $g_2(x) = x_1^2 - 2x_1 + x_2$*. Verifique que o ponto* $\bar{x} = 0$ *cumpre LICQ.*

Como as restrições são ativas em $\bar{x}$, então basta mostrar que os vetores $\nabla g_1(\bar{x}) = \begin{pmatrix} -2 \\ -1 \end{pmatrix}$ e $\nabla g_2(\bar{x}) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ são linearmente independentes:

$$\mu_1 \begin{pmatrix} -2 \\ -1 \end{pmatrix} + \mu_2 \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{cases} -2\mu_1 - 2\mu_2 = 0 & (II) \quad -2\mu_1 - 2\mu_1 = 0 \Rightarrow \mu_1 = 0 \\ -\mu_1 + \mu_2 = 0 & (I) \qquad\qquad\quad \mu_2 = \mu_1. \end{cases}$$

Logo, $\mu_1 = \mu_2 = 0$ e portanto os vetores são linearmente independente. $\square$

Contudo, esta condição apresenta uma desvantagem, pois exige muito das restrições. Por outro lado, existem minimizadores que podem satisfazer as condições de KKT sem que satisfaça LICQ.

Exemplo 6.2 *Considere o seguinte problema*

$$\begin{aligned} Minimizar \;\; & f(x) = x_2 \\ Sujeito \;\; a \;\; & g_1(x) = -x_1^2 + x_2 \leq 0 \\ & g_2(x) = -x_2 \leq 0 \end{aligned}$$

O ponto $x^* = (0,0)^T$ é o minimizador deste problema, logo cumpre as condições de KKT mas não satisfaz LICQ. De fato, as restrições são ativas em x^* e os vetores

$$\nabla g_1(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \nabla g_2(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

$$\mu_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \mu_2 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{cases} 0\mu_1 + 0\mu_2 = 0 & \Rightarrow \quad\; 0 = 0 \\ \mu_1 - \mu_2 = 0 & \Rightarrow \quad \mu_1 = \mu_2. \end{cases}$$

Logo, os vetores são linearmente dependentes pois

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -1 \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Além disso, $-\nabla f(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \nabla g_2(x^*)$, ou seja, vale KKT. □

6.4 Condição de Qualificação de Mangasarian-Fromovitz - MFCQ

Se a condição de Slate exige um ponto no interior relativo do conjunto viável. Já a condição de Mangasarian e Fromovitz ordena que o conjunto viável linearizado tenha interior relativo não vazio.

Definição 6.4 *Dizemos que a condição de qualificação de Mangasarian-Fromovitz (MFCQ) é satisfeita em $\bar{x}$ quando os gradientes das restrições de igualdade são linearmente independentes e existir um vetor $d \in \mathbb{R}^n$ tal que:*

$$\nabla h(\bar{x})^T d = 0 \quad \text{e} \quad \nabla g_j(\bar{x})^T d < 0, \forall j \in I(\bar{x}).$$

Teorema 6.2 *Se $\bar{x} \in \Omega$ satisfaz LICQ, então $\bar{x}$ satisfaz MFCQ.*

Prova. Suponha $i = \{1, .., m\}$ e $I(\bar{x}) = \{m+1, ..., m+q\}$. Considere a matriz

$$M = (\nabla c_1(\bar{x}) ... \nabla c_m(\bar{x}) \nabla c_{m+1}(\bar{x}) ... \nabla c_{m+q}(\bar{x})),$$

e $b \in \mathbb{R}^{m+q}$ dado $b_i = 0$, $\forall i = 1, ..., m$ e $b_i = -1$, $\forall i = m+1, ..., m+q$. Como as colunas de M são linearmente independentes, o sistema $M^T d = b$ é possível, já que a matriz de coeficientes tem posto linha completo e portanto igual ao posto da matriz ampliada. Logo, d é uma solução do sistema, e assim, obtemos

$$\nabla h_i(\bar{x})^T d = 0 \quad \text{e} \quad \nabla g_j(\bar{x})^T d = -1 < 0, \forall j \in I(\bar{x}).$$

Portanto, MFCQ é satisfeita. □

Em seguida, provaremos que MFCQ implica $T(\bar{x}) = D(\bar{x})$, para isto, recorremos a dois resultados auxiliares. Cuja a prova pode ser encontrada em [11].

Lema 6.1 *Seja $\bar{x}, d \in \mathbb{R}^n$ tais que $c_\varepsilon(\bar{x}) = 0$ e $\nabla c_i(\bar{x})^T d = 0$, $\forall i \in \varepsilon$. Suponha que os gradientes $\nabla c_i(\bar{x})$, $i \in \varepsilon$, são linearmente independentes. Então, existe uma curva diferenciável $\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^{n\times m}$ tal que $c_\varepsilon(\gamma(t)) = 0$, $\forall t \in (-\varepsilon, \varepsilon)$, $\gamma(0) = \bar{x}$ e $\gamma'(0) = d$.*

Lema 6.2 *Seja $\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ uma curva diferenciável tal que $c_\varepsilon(\gamma(t)) = 0$, $\forall t \in (-\varepsilon, \varepsilon)$. Se $\gamma(0) = \bar{x}$ e $\gamma'(0) = d \neq 0$, então existe uma sequência (x^k) tal que $c_\varepsilon(x^k) = 0$, $x^k \to \bar{x}$ e*

$$\frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} = \frac{d}{\|d\|}.$$

Teorema 6.3 *Se $\bar{x}$ satisfaz MFCQ, então $T(\bar{x}) = D(\bar{x})$.*

Prova. Seja $d \in D(\bar{x})$ uma direção arbitrária e $\bar{d}$ um vetor que cumpre MFCQ. Dado $t \in (0, 1)$, defina

$$\hat{d} = (1 - t)d + t\bar{d}.$$

Vamos mostrar que $\hat{d} \in T(\bar{x})$. Como $d, \hat{d} \in D(\bar{x})$, temos $\nabla c_i(\bar{x})^T \hat{d} = 0$, $\forall i \in \varepsilon$. Pelo Lema 6.1, existe uma curva diferenciável $\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ tal que $c_\varepsilon(\gamma(t)) = 0$, $\forall t \in (-\varepsilon, \varepsilon)$, $\gamma(0) = \bar{x}$ e $\gamma'(0) = \hat{d}$. Aplicando o Lema 6.2, concluímos que existe uma sequência (x^k) tal que $c_\varepsilon(x^k) = 0, x^k \to \bar{x}$ e

$$\frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} = \frac{\hat{d}}{\|\hat{d}\|}.$$

Para concluir que $\hat{d} \in T(\bar{x})$ resta mostrar que $c_I(x^k) \leq 0$, $\forall k$ suficientemente grande. Se $i \in I \backslash I(\bar{x})$, então $c_i(\bar{x}) < 0$ e, pela continuidade de $c_i(x^k) \leq 0$, $\forall k$ suficientemente grande. Por outro lado, se $i \in I(\bar{x})$, temos que $\nabla c_i(\bar{x})^T d \leq 0 e \nabla c_i(\bar{x})^T d < 0$. Portanto, $\nabla c_i(\bar{x})^T \hat{d} < 0$. Pela diferenciabilidade de c_i, segue que

$$c_i(x^k) = c_i(\bar{x}) + \nabla c_i(\bar{x})^T (x^k - \bar{x}) + o(\|x^k - \bar{x}).$$

Assim,

$$\frac{c_i(x^k)}{\|x^k - \bar{x}\|} = \nabla c_i(\bar{x})^T \frac{x^k - \bar{x}}{\|x^k - \bar{x}\|} + \frac{o\|x^k - \bar{x}\|}{\|x^k - \bar{x}\|} \to \nabla c_i(\bar{x})^T \frac{\hat{d}}{\|\hat{d}\|} < 0,$$

o que implica $c_i(x^k) < 0$, para todo k grande. Logo $\hat{d} \in T(\bar{x})$. Como $T(\bar{x})$ é fechado, temos que $d \in T(\bar{x})$. □

Com efeito, os Teoremas 6.2 e 6.3 nos permitem concluir que tanto LICQ quanto MFCQ são condições de qualificação. Além disso, através da hipótese adicionada, isto é, pela igualdade entre o Cone Tangente e o Cone Viável Linearizado, têm-se que o Teorema 5.2 nos garante a existência de multiplicadores de Lagrange associados a um minimizador.

As relações entre as condições de qualificação demonstrada nesta dissertação são ilustrada no diagrama abaixo.

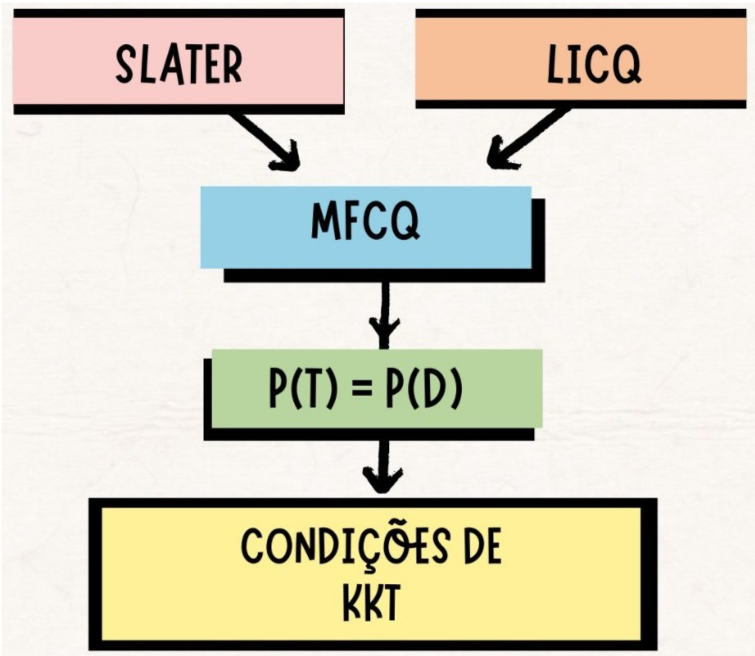


Figura 6.1. Condições de Qualificação Fonte: Autora.

7 CONSIDERAÇÕES FINAIS

Neste trabalho de conclusão de curso, apresentamos uma prova das condições necessárias de otimalidade de Karush-Kuhn-Tucker (KKT) fundamentada em uma abordagem de cones. Desse modo, vimos que o cone tangente concede informações mais interessantes sobre a estrutura local do conjunto viável do que o conjunto das direções viáveis. Além disso, salientamos as relações existentes entre o cone viável linearizado e o cone tangente.

No entanto, a demonstração de KKT utiliza fortemente o famoso Lema de Farkas. Sendo assim, verificamos o tal lema sob duas versões: a versão geométrica, via cones, e a versão algébrica e ainda apresentamos a equivalência entre elas quando o cone em questão é um cone finitamente gerado.

Por fim, vimos que as condições de KKT são verdadeiras se o minimizador satisfaz uma condição de qualificação. Em síntese, relacionamos o cone tangente e o cone viável linearizado para obtermos uma condição de qualificação. Com efeito, as condições KKT se verificam, quando a igualdade entre o polar desses cones ocorre num minimizador. Embora essa condição seja fraca, ela é difícil de ser verificada. Contudo, estudamos algumas condições de qualificações mais clássicas da literatura, com a intenção de se obter condições fracas e fáceis de serem verificadas.

REFERÊNCIAS

[1] H. D. S. C. M. BAZARAA, Mokhtar S.; SHERALI. ***Nonlinear programming: theory and algorithms***. John Wiley Sons, 2013.

[2] D. P. BERTSEKAS. **Nonlinear programming**. *Journal of the Operational Research Society*, 1997.

[3] L. Contese. **Introduccion a La Optimizacion con Restricciones**. *Universidad de Chile-Departamento de Matematicas y Ciencias de la Computacion, Casilla*.

[4] R. G. Eustáquio. ***Condições de otimalidade e de qualificação para problemas de programação não linear***. PhD thesis, Master's thesis, Universidade Federal do Paraná, Curitiba, PR, 2007.

[5] C. Isoton. ***Condições necessárias e suficientes de otimalidade para problemas com um e com vários objetivos***. PhD thesis, Dissertação de Mestrado. Universidade Federal do Paraná, 2013.

[6] A. Izmailov and M. Solodov. **Otimiza c ao: Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade. 2a edição**. *IMPA, Rio de Janeiro*, 2009.

[7] H. W. KUHN. ***Nonlinear Programming: a historical view.*** In: Traces and emergence of nonlinear programming., Springer Basel, 2013.

[8] H. W. KUHN and A. W. TUCKER. ***Nonlinear Programming***. Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley CA, 1951.

[9] E. L. LIMA. ***Análise no Espaço Rn***. IMPA, Rio de Janeiro, 2 edition, 2016.

[10] O. L. Mangasarian. ***Nonlinear Programming***. SIAM, Rio de Janeiro, 2 edition, 1994.

[11] A. A. Ribeiro and E. W. Karas. **Otimização Contínua: aspectos teóricos e computacionais**. *Cengage Learning, Sao Paulo*, 2013.