UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Exatas
Departamento de Pós-graduação em Matemática

Alfonso Enrique Oré Rosales

Uma extensão do Lema Local de Lovász e suas aplicações em Programação Inteira

Belo Horizonte

2019

Alfonso Enrique Oré Rosales

Uma extensão do Lema Local de Lovász e suas aplicações em Programação Inteira

Dissertação apresentada ao Programa de Pós-Graduação em Matemática do Instituto de Ciências Exatas (ICEx) da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Mestre em Matemática.

Orientador: Prof. Dr. Bhalchandra Digambar Thatte.
Coorientador: Prof. Dr. Maurício de Lemos Rodrigues Collares Neto.

Belo Horizonte
2019

Oré Rosales, Alfonso Enrique.

O66e Uma extensão do Lema Local de Lovász e suas aplicações em programação inteira [recurso eletrônico] / Alfonso Enrique Oré Rosales. – 2019.
1 recurso online (67 f. il.): pdf.

Orientador: Bhalchandra Digambar Thatte.
Coorientador: Maurício de Lemos Rodrigues Collares Neto.
Dissertação (Mestrado) - Universidade Federal de Minas Gerais, Instituto de Ciências Exatas, Departamento de Matemática.
Referências: f. 66-67.

1. Matemática – Teses. 2. Programação inteira – Teses. 3. Métodos de relaxação (Matemática) – Teses. 4. Algoritmos de aproximação – Teses. 5. Verificação probabilística de modelos I. Thatte, Bhalchandra Digambar.Thatte. II. Collares Neto, Maurício de Lemos Rodrigues. III. Universidade Federal de Minas Gerais, Instituto de Ciências Exatas, Departamento de Matemática. IV. Título.

CDU 51(043)

Ficha catalográfica elaborada pela bibliotecária Irenquer Vismeg Lucas Cruz
CRB 6/819 - Universidade Federal de Minas Gerais – ICEx

Universidade Federal de Minas Gerais
Departamento de Matemática
Programa de Pós-Graduação em Matemática

FOLHA DE APROVAÇÃO

Uma extensão do Lema Local de Lovász e suas aplicações em Programação Inteira

ALFONSO ENRIQUE ORÉ ROSALES

Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores:

Prof. Bhalchandra Digambar Thatte
UFMG

Maurício Collares Neto

Prof. Maurício de Lemos Rodrigues Collares Neto
UFMG

Prof. Aldo Procacci
UFMG

Sokol Ndreca

Prof. Sokol Ndreca
EST-UFMG

Belo Horizonte, 10 de dezembro de 2019.

Av. Antônio Carlos, 6627 – Campus Pampulha - Caixa Postal: 702
CEP-31270-901 - Belo Horizonte – Minas Gerais - Fone (31) 3409-5963
e-mail: pgmat@mat.ufmg.br - home page: http://www.mat.ufmg.br/pgmat

ATA DA TRECENTÉSIMA TRIGÉSIMA SEXTA DEFESA DE DISSERTAÇÃO DO ALUNO ALFONSO ENRIQUE ORÉ ROSALES, REGULARMENTE MATRICULADO NO PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA, DO INSTITUTO DE CIÊNCIAS EXATAS, DA UNIVERSIDADE FEDERAL DE MINAS GERAIS, REALIZADA NO DIA 10 DE DEZEMBRO DE 2019.

Aos dez dias do mês de dezembro de 2019, às 14h00, na sala 3060, reuniram-se os professores abaixo relacionados, formando a Comissão Examinadora homologada pelo Colegiado do Programa de Pós-Graduação em Matemática, para julgar a defesa de dissertação do aluno **Alfonso Enrique Oré Rosales**, intitulada: "*Uma extensão do Lema Local de Lovász e suas aplicações em Programação Inteira*", requisito final para obtenção do Grau de mestre em Matemática. Abrindo a sessão, o Senhor Presidente da Comissão, Prof. Bhalchandra Digambar Thatte, após dar conhecimento aos presentes do teor das normas regulamentares do trabalho final, passou a palavra ao aluno para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores com a respectiva defesa do aluno. Após a defesa, os membros da banca examinadora reuniram-se sem a presença do aluno e do público, para julgamento e expedição do resultado final. Foi atribuída a seguinte indicação: o aluno foi considerado aprovado sem ressalvas e por unanimidade. O resultado final foi comunicado publicamente ao aluno pelo Senhor Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente Ata, que será assinada por todos os membros participantes da banca examinadora. Belo Horizonte, 10 de dezembro de 2019.

PROF. DR. BHALCHANDRA DIGAMBAR THATTE
Orientador (UFMG)

Maurício Collares Neto
PROF. DR. MAURÍCIO DE LEMOS RODRIGUES COLLARES NETO
Coorientador (UFMG)

PROF. DR. ALDO PROCACCI
Examinador (UFMG)

Sokol Ndreca
PROF. DR. SOKOL NDRECA
Examinador (EST-UFMG)

Dedicado a todas as pessoas que tornaram este trabalho possível.

Agradecimentos

A mi familia, a todos y cada uno de ellos que con sus ejemplos, consejos y castigos, me encaminaron en la búsqueda de mis ambiciones; especialmente a mi Madre que me ejemplifico el trabajo duro y la responsabilidad.

A mi Alma Mater... la Universidad Nacional Mayor de San Marcos (UNMSM) que con sus limitaciones logró alimentar mis conocimientos para interpretar el lenguaje de esta fémina llamada Matemática, quedando cautivado de su complejidad. A mis amistades en su conjunto, que siempre están a mi lado, aún cuando yo no soy reciproco con ellos. Especialmente a los de San Marcos y las de mi barrio.

A Brasil en su conjunto como sociedad, que desde el primer momento me hizo sentir como uno de los suyos. A su gente, a la Universidad Federal de Minas Gerais (UFMG) por darme la oportunidad de realizar mis metas. A las amistades que hice en el departamento de matemática, que me ayudaron tanto en lo académico como en lo personal, al ejemplificar que los malos momentos se pueden superar con un 'tudo bem' o al errar en mi escrito decir 'mio portuguez é ...' y el poder utilizar 'tren' para cualquier cosa que he olvidado. Mi gratitud eterna y mi deuda personal con todos y cada uno de ellos y ellas. A los profesores del departamento de matemática de UFMG especialmente al profesor Bhalchandra por creer en mí, al aceptar ser mi orientador y al profesor Mauricio por inyectar la energía necesaria para poder finiquitar el trabajo.

GRACIAS TOTALES.

Não é possível democratizar a educação em um país sem democratizar a sua economia e sem democratizar, portanto, sua superestrutura política.
José Carlos Mariátegui.

Resumo

O Lema Local de Lovász (LLL) apresentado por László Lovász e Paul Erdős é uma ferramenta muito útil para aplicações em métodos probabilísticos. Apresentaremos uma extensão desse lema, enfraquecendo a hipótese de dependência de eventos. Como aplicação consideramos dois tipos de problemas de Programação Inteira: Problemas Minimax e problemas de Cobertura, desenvolvendo uma técnica fundamental, o arredondamento aleatório de relaxações lineares, para obter bons algoritmos de aproximação para esses problemas.

Palavras-chave: lema local de Lovász; algoritmos de aproximação; programação inteira; relaxações lineares.

Abstract

The Lovász Local Lemma (LLL) presented by László Lovász and Paul Erdős is a very useful tool for applications of probabilistic methods. We will present an extension of the lemma, weakening the hypothesis of event dependence. As an application we consider two types of Integer Programming problems: Minimax Problems and Covering Problems, developing a fundamental technique, the random rounding of linear relaxations, to obtain good approximation algorithms for these problems.

Key words: Lovász local lemma; approximation algorithms; integer programming; linear relaxation.

Sumário

Introdução

O Método Probabilístico é um método para provar a existência de estruturas com certas propriedades. Ele consiste em construir um espaço de probabilidade apropriado e mostrar que, com probabilidade positiva, sortearemos uma estrutura com as propriedades desejadas. Em [4], Paul Erdős e László Lovász provaram um lema muito útil para aplicações do método probabilístico, hoje conhecido como Lema Local de Lovász.

Essa dissertação estuda o uso do Lema Local de Lovász para mostrar que certos problemas de programação inteira admitem soluções de valor próximo aos de suas relaxações lineares (isto é, do programa linear obtido removendo-se as restrições de integralidade das variáveis). Além disso, tais técnicas são usadas para obter resultados de aproximação para problemas computacionalmente difíceis.

Na parte preliminar introduzimos conceitos que serão fundamentais para o trabalho, como grafos, hipergrafos, probabilidade condicional, independência de eventos e esperança condicional, e rudimentos de complexidade computacional e algoritmos de aproximação.

No Capítulo 2 vamos estabelecer lemas sobre o comportamento de números binomiais que nos levarão a um teorema que é simultaneamente uma desigualdade de concentração e um teorema de decomposição (Teorema 2.8). Definiremos funções úteis para o uso de tal desigualdade de concentração no contexto de algoritmos de aproximação. Além disso, definiremos eventos monótonos e provaremos a desigualdade FKG.

No Capítulo 3 vamos desenvolver o Lema Local de Lovász no caso geral e no caso simétrico. Vamos então estender tal lema, enfraquecendo a dependência em d, um parâmetro crucial na aplicação do Lema Local (Lema 3.5).

Vamos desenvolver no Capítulo 4 a Programação Minimax Inteira (MIP); utilizaremos a extensão do Lema Local de Lovász para demonstrar o Teorema 4.6 que garante que, dado um MIP, o arrendondamento aleatório produz uma solução que aproxima a solução da relaxação linear com probabilidade positiva. Tal aproximação depende de certos parâmetros do MIP; no Teorema 4.9, cuja prova usa o Teorema 4.6, encontramos outra aproximação que depende de um parâmetro que é frequentemente muito menor que o que aparece no enunciado da primeira aproximação.

Na última parte abordaremos os problemas de cobertura (CIP). Provamos que o esquema de arredondamento aleatório também fornece boas soluções inteiras no Teorema 5.16. Definimos então CIPs com várias funções-custo, os chamados CIP multicritério, e provamos um teorema análogo nesse caso (Teorema 5.19). Finalmente, desenvolvemos uma versão computacionalmente eficiente do resultado de aproximação para CIPs no Teorema 5.26.

Capítulo 1

Preliminares

Neste capítulo introduzimos alguns conceitos básicos que serão necessários nos próximos capítulos. Na primeira seção, definimos grafos e hipergrafos. Na segunda seção definimos conceitos básicos de teoria de probabilidade, tais como probabilidade condicional, independência de eventos e esperança condicional. Na última parte, fazemos uma rápida introdução aos conceitos de complexidade computacional que usaremos nessa dissertação.

As referências principais para esse Capítulo são os livros *An Introduction to the Theory of Probability* de Mukhopadhyay [10], *Combinatorial Optimization* de Schrijver [13], *Introduction to the Theory of Computation* de Sipser [14] e *Uma Introdução Sucinta a Algoritmos de Aproximação* [2].

1.1 Teoria de grafos

Definição 1.1 (Grafos não-direcionados)**.** Um grafo ou grafo não direcionado é um par $G = (V, E)$ onde V é um conjunto finito e E é uma família de pares não-ordenados de elementos de V. Os elementos de V são chamados de vértices, e os elementos de E são chamados de arestas. Usaremos a notação uv para denotar a aresta $\{u, v\}$.

Denotamos o conjunto de vértices de G por $V(G)$ e a família de arestas de G por $E(G)$. Nas estimativas de tempo de execução dos algoritmos, denotamos $n := |V(G)|$ e $m := |E(G)|$.

Definição 1.2 (Grafo direcionado)**.** Um digrafo ou grafo direcionado é um par $D = (V, E)$ em que V é um conjunto finito e E é uma família de pares ordenados de elementos de V. Os elementos de V são chamados de vértices, e os elementos de E são chamados de arestas ou arcos.

Denotamos por $V(D)$ o conjunto de vértices de D e $E(D)$ o conjunto de arestas de D.

Definição 1.3 (Hipergrafo)**.** Um hipergrafo $H = (V, \mathcal{E})$ é um par onde V é um conjunto finito e $\mathcal{E}$ é uma família de subconjuntos de V. Os elementos de V e $\mathcal{E}$ são chamados de vértices e arestas, respectivamente. Se $|F| = k$ para todo $F \in \mathcal{E}$ então o hipergrafo é chamado k-uniforme.

1.2 Conceitos básicos de teoria de probabilidade

Definição 1.4 (Probabilidade condicional)**.** Seja A um evento com $\mathcal{P}(A) > 0$. A fração

$$\frac{\mathcal{P}(A \wedge B)}{\mathcal{P}(A)}$$

é chamada probabilidade condicional de B dado A (ou simplesmente probabilidade de B dado A) e é denotada por $\mathcal{P}(B \mid A)$.

Definição 1.5 (Independência de eventos)**.** Dizemos que os eventos A e B são independentes se

$$\mathcal{P}(A \wedge B) = \mathcal{P}(A)\,\mathcal{P}(B).$$

Definição 1.6 (Esperança)**.** Seja X uma variável aleatória definida em um espaço de probabilidade $(S, \mathcal{A}, \mathcal{P})$. A esperança de X é definida por

$$\mathbf{E}[X] = \int_S X \, \mathrm{d}\mathcal{P}$$

onde a integral é a integral de Lebesgue com relação à medida $\mathcal{P}$.

Se X é uma variável aleatória discreta (assumindo um conjunto enumerável de valores $\{x_1, x_2, \ldots\}$, então

$$\mathbf{E}[X] = \sum_{k=1}^{\infty} x_k \cdot \mathcal{P}(X = x_k).$$

Teorema 1.7. *Tem-se as seguintes propriedades da esperança:*

1. *Se a variável aleatória X é tal que $\mathcal{P}(X = c) = 1$, então*

$$\mathbf{E}[X] = c.$$

2. *Se $\mathbf{E}[X]$ existe, então*

$$\mathbf{E}[aX + b] = a\mathbf{E}[X] + b,$$

onde a, b são constantes arbitrárias.

3. *Se $g(X)$ é uma função Borel de X, então*

$$\Big|\mathbf{E}\big[g(X)\big]\Big| \leq \mathbf{E}\big[|g(X)|\big].$$

4. *Se $g(X) = g_1(X) + g_2(X)$, onde $g_1(X)$ e $g_2(X)$ são funções Borel da variável aleatória X, então*

$$\mathbf{E}\big[g(X)\big] = \mathbf{E}\big[g_1(X)\big] + \mathbf{E}\big[g_2(X)\big],$$

desde que as esperanças existam.

5. *Se $g_1(X) \leq g_2(X)$ para todo valor possível de X, então*

$$\mathbf{E}\big[g_1(X)\big] \leq \mathbf{E}\big[g_2(X)\big].$$

Uma demonstração para este teorema pode ser encontrada em [10] na subseção 6.2.1.

Definição 1.8 (Esperança condicional)**.** Seja X uma variável aleatória definida em $(S, \mathcal{A}, \mathcal{P})$ tal que $\mathbf{E}[X]$ exista, e A um subconjunto mensurável com $\mathcal{P}(A) > 0$. Definimos

$$\mathbf{E}[X \mid A] = \int_S X \; \mathrm{d}\mathcal{P}(\,\cdot \mid A)$$

como esperança condicional de X dado A. Para $A \in \mathcal{A}$ tal que $\mathcal{P}(A) > 0$, vale que

$$\mathcal{P}(B \mid A) = \frac{\mathcal{P}(B \cap A)}{\mathcal{P}(A)}.$$

para todo $B \in \mathcal{A}$. Assim, denotando a função indicadora de um conjunto A por $\chi(A)$, podemos escrever

$$\mathbf{E}[X \mid A] = \frac{1}{\mathcal{P}(A)} \int_S X \, d\mathcal{P} = \frac{\mathbf{E}\left[X \cdot \chi(A)\right]}{\mathcal{P}(A)}$$

desde que $\mathbf{E}\left[X \cdot \chi(A)\right]$ exista; como $|X \cdot \chi(A)| \leq |X|$, então temos que $\mathbf{E}\left[X \cdot \chi(A)\right]$ existe se $\mathbf{E}[X]$ existe. Em particular, se $X = \chi(B)$ com $B \in \mathcal{A}$ então

$$\mathbf{E}\left[\chi(B) \mid A\right] = \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(A)} = \mathcal{P}(B \mid A).$$

Teorema 1.9. *Sejam X, Y variáveis aleatórias independentes e $\mathbf{E}[X], \mathbf{E}[Y] < \infty$. Então*

$$\mathbf{E}[X \mid Y] = \mathbf{E}[X].$$

Para nossas estimativas probabilísticas, usaremos várias vezes a seguinte proposição.

Proposição 1.10 (Desigualdade de Bernoulli)**.** *Seja $y \geq -1$. Então, se $\beta \geq 1$,*

$$(1+y)^\beta \geq 1 + \beta y.$$

Por outro lado, se $0 \leq \beta \leq 1$,

$$(1+y)^\beta \leq 1 + \beta y$$

1.3 Algoritmos e complexidade

Nessa seção apresentaremos uma versão intuitiva dos principais conceitos de complexidade relevantes para o assunto que estudaremos. Uma formalização de tais conceitos dependeria de um significado preciso do conceito de "algoritmo", formalizado através do conceito de Máquina de Turing, e está fora do escopo da nossa dissertação. Mais detalhes podem ser encontrados em [14].

No contexto de algoritmos e complexidade, dizemos que um problema L é um *problema de decisão* se sua resposta é "sim" ou "não"; para uma entrada x, escrevemos $x \in L$ se a resposta para tal entrada é "sim" e $x \notin L$ caso contrário.

Os problemas de decisão podem ser classificados de acordo com sua complexidade dos algoritmos para resolvê-los. Abaixo listamos três importantes classes de complexidade.

Definição 1.11 (P)**.** A classe de complexidade P é a classe dos problemas de decisão L que admitem algoritmos que determinam se $x \in L$ em tempo polinomial no tamanho da entrada x.

O tamanho da entrada é frequentemente medido pelo número de dígitos binários (bits) necessários para representá-la.

Definição 1.12 (Verificador)**.** Um algoritmo verificador para um problema L é um algoritmo M cujas entradas são da forma (x, u), onde x é entrada do problema L e u é uma entrada adicional denominada *certificado*. M deve satisfazer a seguinte propriedade: Dado um x, existe u tal que x tal que M aceita a entrada (x, u) se e somente se $x \in L$.

Definição 1.13 (NP)**.** A classe de complexidade NP é a classe dos problemas L de decisão que admitem um verificador que executa em tempo polinomial com a seguinte propriedade: dado $x \in L$, existe certificado u com tamanho polinomial.

Exemplo 1.14. Determinar se um dado grafo tem um ciclo Hamiltoniano é um problema que está em NP. De fato, é fácil fazer um algoritmo polinomial para verificar se um dado conjunto de n vértices forma um ciclo Hamiltoniano. Se um grafo possui um ciclo Hamiltoniano, os vértices de um tal ciclo (em ordem) formam um certificado aceito por tal algoritmo. Por outro lado, se um grafo não possui ciclo Hamiltoniano, não existe nenhum certificado que "engana" o algoritmo verificador.

Frequentemente, diz-se que os problemas em P admitem algoritmos "eficientes". Um jeito de tornar essa afirmação concreta é o seguinte: Seja um problema em P e uma entrada de tamanho n. Ao dobrarmos o tamanho da entrada, nossa estimativa de tempo de execução é multiplicada por uma constante (que depende do grau do polinômio correspondente ao tempo de execução do algoritmo, mas não de n). Tal propriedade não é satisfeita por algoritmos cujo tempo de execução é exponencial no tamanho da entrada.

Um dos principais problemas em aberto da teoria da computação é o problema P versus NP. Expandindo as definições acima, tal questão consiste em determinar se existe algum problema que não pode ser resolvido em tempo polinomial, mas cuja solução pode ser verificada em tempo polinomial. Um avanço importante na questão P versus NP ocorreu no início dos anos 70 com os trabalhos de Stephen Cook [3] e Leonid Levin [8].

Definição 1.15 (NP-difícil)**.** Um problema L é NP-difícil se, dado qualquer problema L' em NP e qualquer entrada x' para L', existe um modo de convertê-la em tempo polinomial para uma entrada x de L com a propriedade de que $x' \in L'$ se e somente se $x \in L$.

Definição 1.16 (NP-completo)**.** Um problema é NP-completo se é NP-difícil e está em NP.

Num certo sentido, os problemas NP-completos são "os mais difíceis" da classe NP. Cook e Levin mostram que tais problemas existem. Mais especificamente, eles mostraram que o problema SAT é NP-completo. Posteriormente, Karp mostrou reduções entre 21 problemas importantes, provando que todos eles eram NP-completos desenvolvido em [7]; hoje em dia existem milhares de problemas que sabemos ser NP-completos. Por definição, resolver qualquer um desses problemas implicaria numa solução polinomial para todos os problemas em NP.

1.3.1 Algoritmos de aproximação

Além dos problemas de decisão, uma outra categoria importante de problemas para essa dissertação é a dos *problemas de otimização*, para os quais a resposta é uma estrutura que minimiza ou maximiza o valor de uma função objetivo. Dado um problema de otimização, o problema de determinar se existe uma solução de valor no mínimo/no máximo x é um problema de decisão; se soubermos resolver esse problema de decisão para todo x em tempo polinomial, podemos procurar (através de busca binária) o valor ótimo com qualquer precisão que desejarmos.

Infelizmente, muitos problemas de otimização de grande importância teórica (e econômica) são tais que a versão de decisão é NP-completa. Como exemplos, podemos citar o Problema do Caixeiro-Viajante, o Problema da Mochila, o problema da Cobertura Miníma por Conjuntos, o Empacotamento de Objetos em Contêineres, entre outros. Os problemas estudados nos Capítulos 4 e 5 também têm essa propriedade.

Encontrar a solução ótima para tais problemas não é computacionalmente viável a menos que $\mathsf{P} = \mathsf{NP}$. Por isso, é interessante buscar algoritmos que rodem em tempo polinomial e encontrem soluções que são garantidamente não muito piores que a solução ótima; esses são os chamados *algoritmos de aproximação*. No projeto de algoritmos de aproximação, há um balanço delicado entre tempo de execução e qualidade da aproximação retornada.

Um exemplo ilustrativo, extraído de [2], é o problema de escalonamento, enunciado abaixo. Informalmente, temos n tarefas (cada uma com um tempo de execução) e m processadores igualmente rápidos, e queremos distribuí-las de modo que o conjunto total de tarefas seja concluído o mais rápido possível.

Problema 1.17. ESCALONAMENTO(m, n, t): Dados inteiros positivos m, n e tempos $t_i \in \mathbb{Q}_+$, $1 \leq i \leq n$, encontrar uma partição $\{M_1, \ldots, M_m\}$ de $\{1, \ldots, n\}$ que minimize

$$\max_{1 \leq j \leq m} \sum_{i \in M_j} t_j.$$

Determinar a solução ótima para tal problema é computacionalmente difícil. É possível provar, no entanto, que há um algoritmo eficiente que retorna uma resposta que é no máximo duas vezes pior que a resposta ótima. Tal algoritmo, obtido por Graham [5], é um simples algoritmo guloso e deu início à área de algoritmos de aproximação.

Capítulo 2

Ferramentas úteis

O objetivo principal desse capítulo é demonstrar o Teorema 2.8, uma desigualdade do tipo Chernoff originalmente obtida por Hoeffding [6]. O teorema afirma que se $X_1, \ldots, X_n \in [0, 1]$ são variáveis aleatórias independentes e $X = \sum_{i=1}^{n} X_i$, a probabilidade de X ser maior que a média decai exponencialmente.

A prova que faremos segue [12] e fornece uma decomposição que será útil para aplicar uma modificação do Lema Local de Lovász (Lema 3.5 no Capítulo 4). Uma ideia similar também servirá como base para a prova dos teoremas principais do Capítulo 5. A demonstração aqui dada também tem a vantagem de ser adaptável para variáveis aleatórias com dependência limitada.

2.1 Funções simétricas e convexidade

Definição 2.1. Seja $S_j \colon \mathbb{R}^n \to \mathbb{R}$ dado por

$$S_j(z_1, \ldots, z_n) = \sum_{i_1 < \cdots < i_j} z_{i_1} \cdots z_{i_j}$$

para $j > 0$ e $S_0(z_1, \ldots, z_n) \equiv 1$.

Definição 2.2. Para cada real x e um inteiro não-negativo r definimos

$$\binom{x}{r} = \frac{1}{r!} \prod_{i=0}^{r-1} (x - i).$$

Em particular, $\binom{x}{0} = 1$.

Definição 2.3. Seja x um número real e r um inteiro não-negativo. Definimos

$$[x]_r = x(x-1)(x-2) \cdots (x-r+1).$$

Em particular, $[x]_0 = 1$.

Lema 2.4. *Para cada $i = 1, \ldots, n$, seja z_i um número satisfazendo $0 \leq z_i \leq 1$. Se $a \in \mathbb{R}$ e $j \in \mathbb{Z}$ são tais que $a \geq j \geq 0$ e $\sum_{i=1}^{n} z_i \geq a$, então $S_j(z_1, \ldots, z_n) \geq \binom{a}{j}$.*

Demonstração. Seja $K = \{z \in [0,1]^n : \sum_{i=1}^n z_i \geq a\}$, e defina a função

$$\begin{aligned} S_j \colon K &\to \mathbb{R} \\ (z_1, \ldots, z_n) &\mapsto \sum_{1 \leq i_1 < \cdots < i_j \leq n} z_{i_1} \cdots z_{i_j}. \end{aligned}$$

Como K é compacto e S_j é uma função continua, ela admite um mínimo z^* em K. Além disso, como S_j é crescente em cada uma das variáveis, podemos diminuir alguns dos z_i^* (se necessário) para que valha $\sum_{i=1}^n z_i^* = a$. Iremos mostrar agora que podemos supor que z^* tem no máximo uma coordenada não-inteira. Para isso, seja $(z_1, \ldots, z_n)$ um vetor e p e q índices com $0 < z_p \leq z_q < 1$. Então

$$S_j(z_1, \ldots, z_n) = \sum_{|\Lambda| = j} z_\Lambda = A_2 z_p z_q + A_1(z_p + z_q) + A_0,$$

onde $A_\ell = \sum_S \prod_{s \in S} z_s$, com S percorrendo os subconjuntos de tamanho $j - \ell$ de $\{1, \ldots, n\} \setminus \{p, q\}$. Definindo $\varepsilon = \min\{z_p, 1 - z_q\} > 0$, podemos calcular

$$\begin{aligned} S_j(z_1, \ldots, z_p - \varepsilon, \ldots, z_q + \varepsilon, \ldots, z_n) &= A_2(z_p - \varepsilon)(z_q + \varepsilon) + A_1(z_p + z_q) + A_0, \\ &= S_j(z_1, \ldots, z_n) - \varepsilon A_2(z_q - z_p + \varepsilon) \\ &\leq S_j(z_1, \ldots, z_n). \end{aligned}$$

Assim, existe um mínimo da função S_j com no máximo uma coordenada inteira. Iremos supor que z^* tem no máximo uma coordenada não-inteira.

Se $a = \sum_{i=1}^n z_i^* \in \mathbb{Z}$, então a propriedade acima implica que $z_i^* \in \{0, 1\}$ para todo i. Para provar o resultado nesse caso, note que a é o número de coordenadas do vetor z que não são iguais a zero. Assim,

$$S_j(z^*) = \sum_{1 \leq i_1 < \cdots < i_j \leq n} z_{i_1}^* \cdots z_{i_j}^* = \binom{a}{j}.$$

Agora suponhamos que $a \notin \mathbb{Z}$. Defina $a_1 = \lfloor a \rfloor$, $a_2 = a - a_1 > 0$ e seja p o único índice tal que $z_p^* = a_2$. Assim,

$$\begin{aligned} S_j(z^*) = \sum_{1 \leq i_1 < \cdots < i_j \leq n} z_{i_1}^* \cdots z_{i_j}^* &= \sum_{\substack{1 \leq i_1 < \cdots < i_j \leq n \\ \neq p}} z_{i_1}^* \cdots z_{i_j}^* + a_2 \sum_{\substack{1 \leq i_1 < \cdots < i_{j-1} \leq n \\ \neq p}} z_{i_1}^* \cdots z_{i_j}^* \\ &= \binom{a_1}{j} + a_2 \binom{a_1}{j-1}. \end{aligned}$$

Vamos mostrar que

$$\binom{a_1}{j} + a_2 \binom{a_1}{j-1} \geq \binom{a_1 + a_2}{j} = \binom{a}{j}. \tag{2.1}$$

Primeiro vamos demostrar por indução em j que

$$[a_1]_j + a_2 j [a_1]_{j-1} \leq [a_1 + a_2]_j = [a]_j,$$

onde $[x]_r = x(x-1)\cdots(x-r+1)$ e $[x]_0 = 1$. Para $j = 1$,

$$[a_1]_1 + a_2[a_1]_0 = a_1 + a_2 \cdot 1 = \binom{a_1+a_2}{1} = \binom{a}{1}.$$

Suponhamos válido para $j-1$. Agora vamos mostrar que vale para j. Temos,

$$[a_1]_j + a_2 j[a_1]_{j-1} = a_1[a_1-1]_{j-1} + ja_2[a_1-1]_{j-2}a_1 - a_2a_1[a_1-1]_{j-2} + a_2a_1[a_1-1]_{j-2}$$

e tem-se que $a_2 < 1$, $a_2[a_1]_{j-1} \geq a_2[a_1+a_2-1]_{j-1}$. Logo,

$$\begin{aligned}[a_1]_j + a_2 j[a_1]_{j-1} &\geq a_1([a_1-1]_{j-1} + (j-1)a_2[a_1-1]_{j-2}) + a_2[a_1]_{j-1} \\ &\geq a_1([a_1-1]_{j-1} + (j-1)a_2[a_1-1]_{j-2}) + a_2[a_1+a_2-1]_{j-1}.\end{aligned}$$

Por hipótese indutiva,

$$\begin{aligned}[a_1]_j + a_2 j[a_1]_{j-1} &\geq a_1[a_1-1+a_2]_{j-1} + a_2[a_1+a_2-1]_{j-1} \\ &= [a_1+a_2-1]_{j-1}(a_1+a_2) \\ &= a[a-1]_{j-1} \\ &= [a]_j.\end{aligned}$$

Agora aplicando em (2.1) tem-se,

$$S_j(z_1^*, \ldots, z_n^*) = \binom{a_1}{j} + a_2\binom{a_1}{j-1} = \frac{1}{j!}([a_1]_j + a_2 j[a_1]_{j-1}) \geq \frac{1}{j!}[a]_j = \binom{a}{j}.$$

Como z^* era mínimo, provamos que

$$S_j(z_1, \ldots, z_n) \geq S_j(z_1^*, \ldots, z_n^*) \geq \binom{a}{j},$$

como queríamos. $\square$

Lema 2.5. *Sejam $z_1, \ldots, z_n$ números reais não-negativos. Se $a = \sum_{i=1}^n z_i$, então*

$$S_j(z_1, \ldots, z_n) \leq \binom{n}{j}\left(\frac{a}{n}\right)^j.$$

Demonstração. Primeiro vamos a demostrar que $S_j(z_1, \cdots, z_n)$ alcança sua valor máximo quando $z_1 = \cdots = z_n = \dfrac{a}{n}$. Suponha que para algum vetor $z \in \mathbb{R}_+^n$ com $\sum_{h=1}^n z_h = a$ existem $1 \leq p, q \leq n$ tais que $z_p < z_q$. Definimos z' de modo que $z'_p = z_p + \varepsilon$ e $z'_q = z_q - \varepsilon$ para qualquer $\varepsilon < z_q - z_p$, e $z'_j = z_j$ para todo $j \in [n] - \{p, q\}$. Assim,

$$S_j(z_1, \ldots, z_n) = \sum_{|\Lambda|=j} z_\Lambda = A_2 z_p z_q + A_1(z_p + z_q) + A_0,$$

onde $A_\ell = \sum_S \prod_{s \in S} z_s$, com S percorrendo os subconjuntos de tamanho $j - \ell$ de $\{1, \ldots, n\} \setminus$

$\{p,q\}$. Podemos calcular

$$\begin{aligned} S_j(z'_1,\dots,z'_p+\varepsilon,\dots,z'_q-\varepsilon,\dots,z'_n) &= A_2(z'_p+\varepsilon)(z'_q-\varepsilon)+A_1(z'_p+z'_q)+A_0,\\ &= S_j(z'_1,\dots,z'_n)+\varepsilon A_2(z'_q-z'_p-\varepsilon)\\ &\geq S_j(z_1,\dots,z_n). \end{aligned}$$

Com isso mostramos o que queremos. Agora vamos a mostrar a desigualdade do lema. Pelo que já mostramos, podemos supor que $z_{i_j}=\dfrac{a}{n}$ para todo $1\leq j\leq n$. Assim, $z_{i_1}\cdots z_{i_j}\leq\left(\dfrac{a}{n}\right)^j$ para todo $1\leq i_1<\cdots<i_j\leq n$, e portanto temos que

$$S_j\left(z_1,\cdots,z_n\right)=\sum_{i_1<\cdots<i_j} z_{i_1}\cdots z_{i_j}\leq\sum_{i_1<\cdots<i_j}\left(\frac{a}{n}\right)^j=\binom{n}{j}\left(\frac{a}{n}\right)^j.$$

Assim obtemos o desejado. $\square$

2.2 Desigualdades de concentração

Nessa seção, estudaremos cotas para probabilidades do tipo $\mathcal{P}\big(X\geq\mathbf{E}\left[X\right]+k\big)$, onde X é soma de variáveis aleatórias independentes definidas no intervalo $[0,1]$. Nosso objetivo será cotar tal probabilidade em termos da função G abaixo, e faremos isso no Teorema 2.8.

Definição 2.6. Para $\mu>0$ e $\delta\geq 0$, definimos a função G por

$$G(\mu,\delta)=\left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu.$$

Antes de enunciar e provar o resultado principal dessa seção, precisaremos de um lema técnico sobre números binomiais.

Lema 2.7. *Seja $k\in\mathbb{N}\cup\{0\}$ e $n,q\in\mathbb{R}$ com $n\geq q\geq k$. Então*

$$\binom{n}{k}\Big/\binom{q}{k}\leq\left(\frac{n}{q}\right)^k\left(\frac{n}{n-k}\right)^{n-k}\left(\frac{q-k}{q}\right)^{q-k}.$$

Demonstração. Observe que a afirmação a ser provada é equivalente a

$$\frac{\binom{n}{k}(n-k)^{n-k}}{\binom{q}{k}(q-k)^{q-k}}\leq\frac{n^n}{q^q}.$$

Faremos uma prova por indução sobre k. O caso $k=0$ é trivial, pois $\binom{x}{0}=1$ pela Definição 2.2. Para o passo indutivo, iremos mostrar que

$$\frac{\binom{n}{k+1}(n-k-1)^{n-k-1}}{\binom{q}{k+1}(q-k-1)^{q-k-1}}\leq\frac{\binom{n}{k}(n-k)^{n-k}}{\binom{q}{k}(q-k)^{q-k}}.$$

Usando que $\binom{x}{k+1} = \frac{x-k}{k+1}\binom{x}{k}$ e simplificando termos, isso é equivalente a mostrar que

$$\left(1 - \frac{1}{n-k}\right)^{n-k-1} \leq \left(1 - \frac{1}{q-k}\right)^{q-k-1},$$

o que é verdade se $n \geq q \geq k+1$ pois a função $\left(1 - \frac{1}{x}\right)^{x-1}$ é decrescente para $x \geq 1$. $\square$

Teorema 2.8. *Dadas variáveis aleatórias* $X_1, \ldots, X_n \in [0,1]$, *seja* $X = \sum_{i=1}^{n} X_i$ *e* $\mu = \mathbf{E}[X]$. *Então:*

(a) *Para qualquer* $q \geq 0$, *qualquer evento* Z *com* $\mathcal{P}(Z) > 0$ *e qualquer inteiro* $k \leq q$,

$$\mathcal{P}(X \geq q \mid Z) \leq \mathbf{E}\left[S_k(X_1, \ldots, X_n)/\binom{q}{k} \;\middle|\; Z\right].$$

(b) *Seja* k *inteiro positivo e* $\delta = \mu/k$. *Se os* X_i *são independentes, então*

$$\mathcal{P}(X \geq \mu + k) \leq \mathbf{E}\left[S_k(X_1, \ldots, X_n)\Big/\binom{\mu+k}{k}\right] \leq G(\mu, \delta),$$

onde G *é como na Definição 2.6.*

Demonstração. Pelo Lema 2.4 temos que

$$\left[\sum_{i=1}^{n} X_i \geq q\right] \subset \left[S_k(X_1, \ldots, X_n)\Big/\binom{q}{k} \geq 1\right].$$

Assim, pela desigualdade de Markov,

$$\mathcal{P}\left(\sum_{i=1}^{n} X_i \geq q \;\middle|\; Z\right) \leq \mathcal{P}\left(S_k(X_1, \ldots, X_n)\Big/\binom{q}{k} \geq 1 \;\middle|\; Z\right) \leq \mathbf{E}\left[S_k(X_1, \ldots, X_n)\Big/\binom{q}{k} \;\middle|\; Z\right],$$

o que prova o item (a). Para provar o item (b), note que pela independência de X_i, tem-se que

$$\mathbf{E}\left[\sum_{i_1 < \cdots < i_k} X_{i_1} \cdots X_{i_k}\right] = \mathbf{E}[X_{i_1}] \cdots \mathbf{E}[X_{i_k}].$$

Tome $q = (1+\delta)\mu = \mu + k$. Pela definição da função S_k e pelo Lema 2.5, obtemos que

$$\mathbf{E}\left[S_k(X_1, \ldots, X_n)\Big/\binom{q}{k}\right] = S_k(\mathbf{E}[X_1], \ldots, \mathbf{E}[X_n])\Big/\binom{q}{k} \leq \left(\frac{\mu}{n}\right)^k \binom{n}{k}\Big/\binom{q}{k},$$

Aplicando o Lema 2.7 e lembrando que $k = \delta\mu$, obtemos

$$\begin{aligned}\left(\frac{\mu}{n}\right)^k \binom{n}{k} \Big/ \binom{q}{k} &\leq \left(\frac{\mu}{q}\right)^k \left(\frac{n}{n-k}\right)^{n-k} \left(\frac{q-k}{q}\right)^{q-k} \\ &= \left(\frac{1}{1+\delta}\right)^{\delta\mu} \left(\frac{n}{n-\delta\mu}\right)^{n-\delta\mu} \left(\frac{1}{1+\delta}\right)^{\mu} \\ &= \left(1 + \frac{\delta\mu}{n-\delta\mu}\right)^{n-\delta\mu} \left(\frac{1}{1+\delta}\right)^{(1+\delta)\mu}.\end{aligned}$$

Usando que $1 + x \leq e^x$, temos que

$$\left(1 + \frac{\delta\mu}{n-\delta\mu}\right)^{n-\delta\mu} \left(\frac{1}{1+\delta}\right)^{(1+\delta)\mu} \leq e^{\delta\mu} \left(\frac{1}{1+\delta}\right)^{(1+\delta)\mu} = G(\mu, \delta).$$

Combinando os resultados obtidos, provamos a parte (b) e concluímos a prova do Teorema 2.8. $\square$

2.3 Estimativas para G

No Teorema 2.8, obtemos estimativas para o desvio de uma soma de variáveis aleatórias independentes em termos da função G da Definição 2.6. Para uso futuro, coletamos aqui várias desigualdades úteis envolvendo tal função.

Lema 2.9. *Seja G como na Definição 2.6. Então,*

$$e^{-\delta^2\mu} \leq G(\mu, \delta) \leq e^{-\delta^2\mu/3} \qquad \text{se } 0 \leq \delta \leq 1, \tag{2.2}$$

$$e^{-(1+\delta)\ln(1+\delta)\mu} \leq G(\mu, \delta) \leq e^{-\frac{\delta\ln(1+\delta)}{2}\mu} \leq e^{-\delta\mu/3} \qquad \text{se } \delta \geq 1. \tag{2.3}$$

Demonstração. Vamos provar primeiro as cotas inferiores. Para demonstrar que $e^{-\delta^2\mu} \leq G(\mu, \delta)$, usamos que $1 + x \leq e^x$ para calcular

$$\frac{e^\delta}{(1+\delta)^{(1+\delta)}} \geq e^{\delta - \delta(1+\delta)} = e^{-\delta^2},$$

o que mostra a cota inferior de (2.2). Similarmente, para mostrar que $e^{-(1+\delta)\ln(1+\delta)\mu} \leq G(\mu, \delta)$ para $\delta \geq 1$, observamos que

$$\frac{e^\delta}{(1+\delta)^{(1+\delta)}} \geq \frac{1}{(1+\delta)^{(1+\delta)}} = e^{-(1+\delta)\ln(1+\delta)},$$

o que implica a cota interior de 2.3. Passemos às cotas superiores. Vamos provar que $G(\mu, \delta) \leq e^{-\delta^2\mu/3}$. Para isso notemos que $1 + x \geq 2^x$ para $0 \leq x \leq 1$. Assim, nesse intervalo,

$$((1+x)\ln(1+x) - x)' = \ln(1+x) \geq x\ln 2 = \left(\frac{x^2 \ln 2}{2}\right)'.$$

Integrando de 0 a δ, obtemos que, para $0 \leq \delta \leq 1$,

$$(1+\delta)\ln(1+\delta) - \delta \geq \frac{\delta^2 \ln 2}{2} \geq \frac{\delta^2}{3}.$$

Com isso, podemos calcular

$$\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}} = e^{\delta-(1+\delta)\ln(1+\delta)} \leq e^{-\delta^2/3},$$

o que nos dá a cota inferior de (2.2). Para a cota superior de (2.3), observemos que, por convexidade, $x \leq (1+x)\log(1+x)$, o que implica $x/(1+x) \leq \log(1+x)$. Assim, vale que

$$((1+x)\ln(1+x) - x)' = \ln(1+x) \geq \frac{1}{2}\left(\frac{x}{1+x} + \ln(1+x)\right) = \left(\frac{x\ln(1+x)}{2}\right)'.$$

Integrando de 0 a δ, obtemos que

$$(1+\delta)\ln(1+\delta) - \delta \geq \frac{\delta\ln(1+\delta)}{2}.$$

Assim, obtemos

$$\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}} = e^{\delta-(1+\delta)\ln(1+\delta)} \leq e^{-\frac{\delta\ln(1+\delta)}{2}}.$$

o que mostra a primeira cota superior de (2.3). Como $\delta \geq 1$, vale que $\ln(1+\delta)/2 \geq 1/3$, o que implica a última cota superior de (2.3) e conclui a prova do lema. □

Os próximos dois resultados são propriedades simples da função G que serão úteis posteriormente.

Lema 2.10. *G é uma função decrescente na segunda variável.*

Demonstração. Aplicando a derivada na segunda variável, obtemos

$$\left(\frac{\partial G(x,y)}{\partial y}\right) = -x\left(\frac{e^{y}}{(1+y)^{(1+y)}}\right)^{x}\big(\log(1+y)\big) < 0$$

para todo $x, y > 0$. □

Proposição 2.11. *Se $0 < \mu_1 \leq \mu_2$, então para qualquer $k > 0$,*

$$G(\mu_1, k/\mu_1) \leq G(\mu_2, k/\mu_2).$$

Demonstração. Observemos que a função $x \mapsto x\ln x$ é convexa para $x > 0$, já que

$$(x\ln x)'' = (\ln x + 1)' = \frac{1}{x} > 0.$$

Transladando, a função $f(x) := (1+x)\ln(1+x)$ também é convexa para $x > -1$, e portanto $f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$ para $x, y \geq 0$ e $0 \leq t \leq 1$. Aplicando tal desigualdade para

$x = 0$, $y = k/\mu_1$ e $t = \mu_1/\mu_2$ e usando que $f(0) = 0$, obtemos

$$\left(1 + \frac{k}{\mu_2}\right) \ln\left(1 + \frac{k}{\mu_2}\right) = f(ty) \leq tf(y) = \frac{\mu_1}{\mu_2}\left(1 + \frac{k}{\mu_1}\right) \ln\left(1 + \frac{k}{\mu_1}\right).$$

Multiplicando a desigualdade acima por $-\mu_2$,

$$-\mu_1\left(1 + \frac{k}{\mu_1}\right) \ln\left(1 + \frac{k}{\mu_1}\right) \leq -\mu_2\left(1 + \frac{k}{\mu_2}\right) \ln\left(1 + \frac{k}{\mu_2}\right)$$

Como exponencial é uma função crescente, podemos aplicá-la acima para obter

$$\left(\frac{1}{(1 + k/\mu_1)^{1+k/\mu_1}}\right)^{\mu_1} \leq \left(\frac{1}{(1 + k/\mu_2)^{1+k/\mu_2}}\right)^{\mu_2}.$$

Multiplicando ambos os lados por e^k, obtemos a desigualdade desejada. □

2.4 A função H

Frequentemente, aplicamos as estimativas obtidas no Teorema 2.8 com o objetivo de provar que uma união de vários eventos ocorre com probabilidade no máximo p, para um valor apropriado de p. A função H definida abaixo e as estimativas que obteremos ajudam tal processo e serão usadas nos enunciados e provas dos Teoremas 4.6 e 4.9.

Lema 2.12. *Para $\mu > 0$ e $0 < p < 1$, a função definida por*

$$\delta = H(\mu, p) = \begin{cases} 8 \cdot \frac{\ln(p^{-1})}{\mu} \Big/ \ln\left(\frac{\ln(p^{-1})}{\mu}\right) & \textit{se } \mu \leq \ln(p^{-1})/e, \\ 8 & \textit{se } \mu > \ln(p^{-1})/e \textit{ e } \frac{6}{\mu}\ln(\mu + p^{-1}) > 1 \\ \sqrt{\frac{6}{\mu}\ln(\mu + p^{-1})} & \textit{se } \mu > \ln(p^{-1})/e \textit{ e } \frac{6}{\mu}\ln(\mu + p^{-1}) \leq 1, \end{cases} \tag{2.4}$$

satisfaz $\lceil \delta\mu \rceil \cdot G(\mu, \delta) \leq p$, onde $G(\mu, \delta)$ é a função da Definição 2.6.

Note que, sob as condições do segundo caso de (2.4),

$$1 \leq \sqrt{\frac{6}{\mu}\ln(\mu + p^{-1})} < \sqrt{\frac{6}{\mu}\ln(\mu + e^{e\mu})} \leq \sqrt{6(1 + e)},$$

pois a função $x \mapsto \frac{1}{x}\ln(x + e^{ex})$ é decrescente e tende a $1 + e$ quando x tende a 0. Assim, nos últimos dois casos de (2.4), $H(\mu, p) = O\left(\sqrt{(1/\mu)\ln(\mu + p^{-1})}\right)$.

Demonstração. O máximo da função $(\ln x)/x$ no domínio $x > 0$ é atingido quando $x = e$. Por isso, no caso $\mu \leq \ln(p^{-1})/e$, vale que $\delta \geq 8e \geq 1$. Além disso,

$$\ln\left(\frac{x}{\ln x}\right) = \ln(x) \cdot \left(1 - \frac{\ln\ln x}{\ln x}\right) \geq \ln(x) \cdot \left(1 - \frac{1}{e}\right) \geq \frac{\ln(x)}{2}.$$

Em particular, tomando $x = \ln(p^{-1})/\mu$,

$$\ln(\delta) \geq \ln 8 + \frac{1}{2}\ln\left(\frac{\ln(p^{-1})}{\mu}\right) \geq 2 + \frac{1}{2}\ln\left(\frac{\ln(p^{-1})}{\mu}\right).$$

Assim, pelo Lema 2.9,

$$G(\mu,\delta) \leq \exp\left(-\frac{\delta\ln(\delta)}{2}\mu\right) \leq e^{-\delta\mu}\cdot\exp\left(-2\ln(p^{-1})\right) = e^{-\delta\mu}\cdot p^2.$$

Usando que $\lceil x\rceil e^{-x} \leq (x+1)e^{-x} \leq 1$, temos que

$$\lceil\delta\mu\rceil\cdot G(\mu,\delta) \leq \lceil\delta\mu\rceil e^{-\delta\mu}\cdot p^2 \leq p^2 \leq p.$$

Suponha agora que $\mu > \ln(p^{-1})/2$. Para o segundo caso, onde $\delta = 8$, uma verificação numérica mostra que

$$G(\delta,\mu) = e^{\left(\delta-(1+\delta)\ln(1+\delta)\right)\mu} \leq e^{-(\delta+e)\mu}.$$

Assim, usando novamente que $\lceil x\rceil e^{-x} \leq (x+1)e^{-x} \leq 1$ (pois $1+x \leq e^x$), temos

$$\lceil\mu\delta\rceil G(\mu,\delta) \leq \lceil\mu\delta\rceil e^{-(\delta+e)\mu} \leq e^{-e\mu} \leq p,$$

como desejado. Finalmente, se $\mu > \ln(p^{-1})/2$ e $\frac{6}{\mu}\ln(\mu+p^{-1}) \leq 1$, temos por (2.2) que

$$G(\mu,\delta) \leq e^{-\delta^2\mu/3} = \left(\mu+p^{-1}\right)^{-2}.$$

Assim, como $\lceil\mu\rceil \leq \mu+1$ e $p^{-1} > 1$,

$$\lceil\mu\delta\rceil\, G(\mu,\delta) \leq \lceil\mu\rceil\left(\mu+p^{-1}\right)^{-2} \leq \frac{\mu+1}{\mu+p^{-1}}\cdot\frac{1}{\mu+p^{-1}} < \frac{\mu+1}{\mu+1}\cdot\frac{1}{p^{-1}} = p,$$

como queríamos. □

2.5 Eventos monótonos

No Capítulo 5, lidaremos com a dependência entre os eventos através de uma desigualdade de concentração conhecida como desigualdade FKG. Nessa seção, os espaços de probabilidade serão espaços-produto, isto é, espaços de probabilidade $(\Omega, \mathcal{F}, \mathbf{P})$ onde $\Omega = \{0,1\}^\ell$. Dados os vetores binários $\vec{a} = (a_1, a_2, \cdots, a_l)$ e $\vec{b} = (b_1, b_2, \cdots, b_\ell)$ de Ω, a relação $\vec{a} \preceq \vec{b}$ vale se $a_i \leq b_i$ para todo $1 \leq i \leq \ell$.

Definição 2.13. Um evento A é crescente se, para quaisquer $\vec{a} \in A$ e $\vec{b} \in \Omega$ tais que $\vec{a} \preceq \vec{b}$, vale que $\vec{b} \in A$.

Definição 2.14. Um evento A é decrescente se, para quaisquer $\vec{a} \in A$ e $\vec{b} \in \Omega$ tais que $\vec{b} \preceq \vec{a}$, vale que $\vec{b} \in A$.

Definição 2.15. Uma função $\mu\colon L \to \mathbb{R}^+$, onde L é um reticulado finito, é chamada log-supermodular se

$$\mu(x)\mu(y) \leq \mu(x \vee y)\mu(x \wedge y),$$

para todo $x, y \in L$

Teorema 2.16. *Seja L um reticulado finito, e seja $\mu\colon L \to \mathbb{R}^+$ uma função log-supermodular. Então, para qualquer par de funções crescentes $f, g\colon L \to \mathbb{R}^+$ temos que*

$$\left(\sum_{x\in L}\mu(x)f(x)\right)\left(\sum_{x\in L}\mu(x)g(x)\right) \leq \left(\sum_{x\in L}\mu(x)f(x)g(x)\right)\left(\sum_{x\in L}\mu(x)\right).$$

Demonstração. Uma prova para este teorema pode ser encontrada em [1, Teorema 6.2.1]. □

Observação 2.17. No caso que f é crescente e g é decrescente (ou vice-versa), vale a desigualdade oposta:

$$\left(\sum_{x\in L}\mu(x)f(x)\right)\left(\sum_{x\in L}\mu(x)g(x)\right) \geq \left(\sum_{x\in L}\mu(x)f(x)g(x)\right)\left(\sum_{x\in L}\mu(x)\right).$$

Observação 2.18. A interseção de eventos crescentes é crescente e a interseção de eventos decrescentes é decrescente.

Teorema 2.19. *(Desigualdade FKG) Seja $(\Omega, \mathcal{F}, \mathbf{P})$ um espaço-produto.*

(a) *Sejam A e B eventos ambos crescentes ou ambos decrescentes. Então $\mathcal{P}(A \mid B) \geq \mathcal{P}(A)$.*

(b) *Seja I um evento crescente e D um evento decrescente.*

Então $\mathcal{P}(I \mid D) \leq \mathcal{P}(I)$ e $\mathcal{P}(D \mid I) \leq \mathcal{P}(D)$.

Demonstração. Dada $f\colon \Omega \to \mathbb{R}_+$, definimos $\langle f\rangle$ por

$$\langle f\rangle = \frac{\sum_{x\in\mathcal{A}} f(x)\,\mu(x)}{\sum_{x\in\mathcal{A}} \mu(x)},$$
$$\langle g\rangle = \frac{\sum_{x\in\mathcal{A}} g(x)\,\mu(x)}{\sum_{x\in\mathcal{A}} \mu(x)},$$
$$\langle fg\rangle = \frac{\sum_{x\in\mathcal{A}} f(x)\,g(x)}{\sum_{x\in\mathcal{A}} \mu(x)}.$$

Se f, g são ambos não decrescente ou não crescente temos que, pelo Teorema 2.16,

$$\langle fg\rangle \geq \langle f\rangle\langle g\rangle,$$

e se f é não decrescente e g é não crescente ou vice versa pela Observação 2.17, então

$$\langle fg\rangle \leq \langle f\rangle\langle g\rangle.$$

Por hipótese temos que I_i é não decrescente para todo $i \in [t]$. Então se I_i é não decrescente para $\vec{a}$ e $\vec{a} \preceq \vec{b}$, por definição, I_i é não decrescente para $\vec{b}$. Logo $\bigwedge_{j\in S} I_j$ é não decrescente para $\vec{a}$. Se $\vec{a} \preceq \vec{b}$ então $\bigwedge_{j\in S} I_j$ é não decrescente para b.

Analogamente para D_j que é não crescente para todo $i \in [t]$, por definição, se D_i é não crescente para $\vec{a}$ e $\vec{b} \preceq \vec{a}$ então D_i é não crescente para $\vec{b}$. Logo $\bigwedge_{j\in S} D_j$ é não crescente para $\vec{a}$, e daí se $\vec{b} \preceq \vec{a}$ então $\bigwedge_{j\in S} D_j$ é não crescente para $\vec{b}$, ou seja, $\bigwedge_{j\in S} D_j$ é não crescente.

Agora seja $\mu \equiv 1$ e f uma função característica de I_i e g uma função característica de $\bigwedge_{j\in S} I_j$ então,

$$\begin{aligned}
\langle f\rangle = \mathcal{P}\left(I_i\right) &= \frac{\sum_{x\in I_i} f\left(x\right)\mu\left(x\right)}{\sum_{x\in I_i}\mu\left(x\right)}\\
\langle g\rangle = \mathcal{P}\left(\bigwedge_{j\in S} I_i\right) &= \frac{\sum_{x\in I_i} g\left(x\right)\mu\left(x\right)}{\sum_{x\in I_i}\mu\left(x\right)}\\
\langle fg\rangle = \mathcal{P}\left(I_i \wedge \bigwedge_{j\in S} I_j\right) &= \frac{\sum_{x\in I_i\wedge\bigwedge_{j\in S}} f\left(x\right) g\left(x\right)\mu\left(x\right)}{\sum_{x\in I_i\wedge\bigwedge_{j\in S}}\mu\left(x\right)}.
\end{aligned}$$

Agora de forma análoga, seja f_1 uma função característica de D_i e g_1 uma função característica de $\bigwedge_{j\in S} D_j$ então,

$$\begin{aligned}
\langle f_1\rangle = \mathcal{P}\left(D_i\right) &= \frac{\sum_{x\in D_i} f_1\left(x\right)\mu\left(x\right)}{\sum_{x\in D_i}\mu\left(x\right)},\\
\langle g_1\rangle = \mathcal{P}\left(\bigwedge_{j\in S} D_j\right) &= \frac{\sum_{x\in\bigwedge_{j\in S} D_j} g_1\left(x\right)\mu\left(x\right)}{\sum_{x\in\bigwedge_{j\in S}}\mu\left(x\right)},\\
\langle f_1 g_1\rangle = \mathcal{P}\left(D_i \wedge \bigwedge_{j\in S} D_j\right) &= \frac{\sum_{x\in D_i\wedge\bigwedge_{j\in S} D_j} g_1\left(x\right) f_1\left(x\right)\mu\left(x\right)}{\sum_{x\in D_i\wedge\bigwedge_{j\in S} D_j}\mu\left(x\right)}.
\end{aligned}$$

Além disso temos que

$$\begin{aligned}
\langle f_1 g\rangle = \mathcal{P}\left(D_i \wedge \bigwedge_{j\in S} I_j\right) &= \frac{\sum_{x\in D_i\wedge\bigwedge_{j\in S} I_j} f_1\left(x\right) g\left(x\right)\mu\left(x\right)}{\sum_{x\in D_i\wedge\bigwedge_{j\in S} I_j}\mu\left(x\right)}\\
\langle f g_1\rangle = \mathcal{P}\left(I_i \wedge \bigwedge_{j\in S} D_j\right) &= \frac{\sum_{x\in I_i\wedge\bigwedge_{j\in S} D_j} f\left(x\right) g_1\left(x\right)\mu\left(x\right)}{\sum_{x\in I_i\wedge\bigwedge_{j\in S} D_j}\mu\left(x\right)}.
\end{aligned}$$

1. Queremos demostrar: $\mathcal{P}\left(I_i \,\middle|\, \bigwedge_{j\in S} I_j\right) \geq \mathcal{P}\left(I_i\right)$

Temos

$$\mathcal{P}\left(I_i \wedge \bigwedge_{j\in S} I_j\right) = \langle fg\rangle \;\geq\; \langle f\rangle\langle g\rangle = \mathcal{P}\left(I_i\right)\mathcal{P}\left(\bigwedge_{j\in S} I_j\right),$$

então

$$\frac{\mathcal{P}\left(I_i \wedge \bigwedge_{j\in S} I_j\right)}{\mathcal{P}\left(\bigwedge_{j\in S} I_j\right)} \geq \mathcal{P}\left(I_i\right) \Rightarrow \mathcal{P}\left(I_i \,\middle|\, \bigwedge_{j\in S}\bigwedge_{j\in S} I_j\right) \geq \mathcal{P}\left(I_i\right).$$

Asim como queríamos.

2. Queremos demostrar: $\mathcal{P}\left(D_i \;\middle|\; \bigwedge_{j\in S} D_j\right) \geq \mathcal{P}\left(D_i\right)$.

$$\mathcal{P}\left(D_i \wedge \bigwedge_{j\in S} D_j\right) = \langle f_1 g_1\rangle \;\geq\; \langle f_1\rangle\langle g_1\rangle = \mathcal{P}\left(D_i\right)\mathcal{P}\left(\bigwedge_{j\in S} D_j\right),$$

então

$$\frac{\mathcal{P}\left(D_i \wedge \bigwedge_{j\in S} D_j\right)}{\mathcal{P}\left(\bigwedge_{j\in S} D_j\right)} \;\geq\; \mathcal{P}\left(D_i\right) \Rightarrow \mathcal{P}\left(D_i \;\middle|\; \bigwedge_{j\in S} D_j\right) \;\geq\; \mathcal{P}\left(D_i\right).$$

Obtendo o desejado, os casos 3 e 4 são de forma análoga. $\square$

Capítulo 3

Lema Local de Lovász e Extensões

3.1 Visão geral

O Lema Local de Lovász é um resultado para provar que, sob certas condições, interseções de eventos "pouco dependentes" entre si ocorrem com probabilidade positiva. Ele foi originalmente provado em 1973 por László Lovász e Paul Erdős em [4], e frequentemente dá as melhores cotas conhecidas para vários resultados de existência em Combinatória (por exemplo, para números de Ramsey e Van der Waerden).

Nesse capítulo, veremos o caso geral e o caso simétrico do lema, com provas baseando-se em [1, Capítulo 5]. Além disso, veremos uma variação do lema provada em [15] que enfraquece as hipóteses de dependência dos eventos.

3.2 Lema Local de Lovász

Definição 3.1 (Digrafo de dependência)**.** Um grafo direcionado $D = (V, E)$ no conjunto de vértices $V = \{1, 2, 3, \dots, n\}$ é chamado digrafo de dependência para os eventos $A_1, A_2, \dots, A_n$ se para cada i, $1 \leq i \leq n$, o evento A_i é mutuamente independente da coleção $\{A_j : j \neq i, (i, j) \notin E\}$.

Lema 3.2 (Lema Local de Lovász, caso geral)**.** *Suponha que $D = (V, E)$ é um digrafo de dependência para os eventos acima e suponha que existem números reais $x_1, \dots, x_n$ tal que $0 \leq x_i < 1$ e $\mathcal{P}(A_i) \leq x_i \prod_{(i,j)\in E}(1 - x_i)$ para todo $1 \leq i \leq n$. Então*

$$\mathcal{P}\left(\bigwedge_{i=1}^{n} \overline{A}_i\right) \geq \prod_{i=1}^{n}(1 - x_i)\,.$$

Em particular, com probabilidade positiva nenhum evento A_i ocorre.

Demonstração. Faremos a demonstração por passos. Note primeiro que

$$\mathcal{P}\left(\bigwedge_{i=1}^{n} \overline{A}_i\right) = \big(1 - \mathcal{P}(A_1)\big)\left(1 - \mathcal{P}\left(A_2 \,\middle|\, \overline{A}_1\right)\right)\cdots\left(1 - \mathcal{P}\left(A_n \,\middle|\, \bigwedge_{i=1}^{n-1} \overline{A}_i\right)\right) \tag{3.1}$$

Assim, para provar o lema, iremos mostrar que, para todo $S \subset \{1,\ldots,n\}$,

$$\mathcal{P}\left(A_i \,\middle|\, \bigwedge_{j\in S} \overline{A_j}\right) \le x_i. \tag{3.2}$$

A prova dessa afirmação será por indução em $s = |S|$. Para $s = 0$, a afirmação vale por vacuidade, pois não há $i \notin S$. Para o passo indutivo, suponha que a afirmação é válida para conjuntos de tamanho menor que s. Tome um S de tamanho s, fixe um $i \notin S$ e defina

$$S_1 = \{j \in S : (i,j) \in E\} \qquad \text{e} \qquad S_2 = S \setminus S_1.$$

Então, podemos calcular

$$\mathcal{P}\left(A_i \,\middle|\, \bigwedge_{j\in S} \overline{A_j}\right) = \frac{\mathcal{P}\left(A_i \wedge \bigwedge_{j\in S_1} \overline{A_j} \wedge \bigwedge_{\ell\in S_2} \overline{A_\ell}\right)}{\mathcal{P}\left(\bigwedge_{j\in S_1} \overline{A_j} \wedge \bigwedge_{\ell\in S_2} \overline{A_\ell}\right)} = \frac{\mathcal{P}\left(A_i \wedge \bigwedge_{j\in S_1} \overline{A_j} \,\middle|\, \bigwedge_{\ell\in S_2} \overline{A_\ell}\right)}{\mathcal{P}\left(\bigwedge_{j\in S_1} \overline{A_j} \,\middle|\, \bigwedge_{\ell\in S_2} \overline{A_\ell}\right)}. \tag{3.3}$$

Para cotar o numerador do lado direito, observemos que como A_i é mutuamente independente da coleção $\{A_\ell : \ell \in S_2\}$, então

$$\mathcal{P}\left(A_i \wedge \bigwedge_{j\in S_1} \overline{A_j} \,\middle|\, \bigwedge_{\ell\in S_2} \overline{A_\ell}\right) = \frac{\mathcal{P}\left(A_i \wedge \bigwedge_{j\in S_1} \overline{A_j} \wedge \bigwedge_{\ell\in S_2} \overline{A_\ell}\right)}{\mathcal{P}\left(\bigwedge_{\ell\in S_2} \overline{A_\ell}\right)} \le \frac{\mathcal{P}\left(A_i \wedge \bigwedge_{\ell\in S_2} \overline{A_\ell}\right)}{\mathcal{P}\left(\bigwedge_{\ell\in S_2} \overline{A_\ell}\right)} = \mathcal{P}\left(A_i\right).$$

Assim, pela hipótese do lema,

$$\mathcal{P}\left(A_i \wedge \bigwedge_{j\in S_1} \overline{A_j} \,\middle|\, \bigwedge_{\ell\in S_2} \overline{A_\ell}\right) \le x_i \prod_{(i,j)\in E} \left(1 - x_j\right). \tag{3.4}$$

O denominador de (3.3), por outro lado, pode ser limitado por hipótese indutiva. Se $|S_1| = 0$ então o denominador é 1 e segue-se o solicitado. Se $S_1 = \{j_1, \ldots, j_r\}$, então

$$\begin{aligned}\mathcal{P}\left(\bigwedge_{i=1}^{r} \overline{A_{j_i}} \,\middle|\, \bigwedge_{\ell\in S_2} \overline{A_\ell}\right) &= \prod_{i=1}^{r} \mathcal{P}\left(\overline{A_{j_i}} \,\middle|\, \bigwedge_{k=1}^{i-1} \overline{A_{j_k}} \wedge \bigwedge_{\ell\in S_2} \overline{A_\ell}\right) \\ &= \prod_{i=1}^{r} \left(1 - \mathcal{P}\left(A_{j_i} \,\middle|\, \bigwedge_{k=1}^{i-1} \overline{A_{j_k}} \wedge \bigwedge_{\ell\in S_2} \overline{A_\ell}\right)\right).\end{aligned} \tag{3.5}$$

Como todos os termos do produtório estão condicionados a menos de $|S|$ eventos, podemos aplicar a hipótese indutiva para concluir que

$$\mathcal{P}\left(\overline{A_{j_1}} \wedge \overline{A_{j_2}} \wedge \cdots \wedge \overline{A_{j_r}} \,\middle|\, \bigwedge_{\ell\in S_2} \overline{A_\ell}\right) \ge \left(1 - x_{j_1}\right) \cdots \left(1 - x_{j_r}\right) = \prod_{(i,j)\in E} \left(1 - x_j\right). \tag{3.6}$$

Substituindo (3.4) e (3.6) em (3.3) temos que

$$\mathcal{P}\left(A_i \,\middle|\, \bigwedge_{j\in S} \overline{A}_j\right) \leq \frac{x_i \prod (1-x_j)}{\prod (1-x_j)} = x_i.$$

Combinando as equações (3.1) e (3.2), obtemos

$$\mathcal{P}\left(\bigwedge_{i=1}^{n} \overline{A}_i\right) \geq \prod_{i=1}^{n}(1-x_i) > 0,$$

o que conclui a prova do lema. □

Corolário 3.3 (Lema local de Lovász, caso simétrico)**.** *Sejam $E_1, \ldots, E_n$ eventos em um espaço de probabilidade. Suponhamos que cada evento E_i é mutuamente independente a todos os E_j, exceto em d eventos, e que $\mathcal{P}(E_i) \leq p$ para todo $1 \leq i \leq n$. Se $ep(d+1) \leq 1$, então*

$$\mathcal{P}\left(\bigwedge_{i=1}^{n} \overline{E}_i\right) > 0.$$

Demonstração. Se $d = 0$ então os eventos são independentes por hipótese. Assim,

$$\mathcal{P}\left(\bigwedge_{i=1}^{n} \overline{E}_i\right) = \prod_{i=1}^{n} \mathcal{P}\left(\overline{E}_i\right) = \prod_{i=1}^{n}\left(1-\mathcal{P}(E_i)\right) \geq (1-p)^n \geq \left(1-\frac{1}{e}\right)^n > 0.$$

Se $d \geq 1$, então existe um digrafo de dependência $D = (V, E)$ para os eventos $E_1, E_2, \ldots, E_n$ tal que $|\{j : (i,j) \in E\}| \leq d$. Tome $x_i = 1/(d+1) < 1$ para todo i. Usando o fato que para qualquer $d \geq 1$ vale

$$\left(1-\frac{1}{d+1}\right)^d > \frac{1}{e},$$

temos que

$$x_i \prod_{i=1}^{d}(1-x_i) = \left(1-\frac{1}{d+1}\right)^d \frac{1}{d+1} \geq \frac{1}{e}\cdot\frac{1}{d+1} \geq p \geq \mathcal{P}(E_i).$$

Assim, podemos aplicar o Lema 3.2 então

$$\mathcal{P}\left(\bigwedge_{i=1}^{n} \overline{E}_i\right) \geq \prod_{i=1}^{n}(1-x_i) = \left(\frac{d}{d+1}\right)^n > 0,$$

como desejado. □

O exemplo a seguir foi extraído de [16] e é uma aplicação elementar do Lema Local de Lovász.

Exemplo 3.4. Suponha que $Q_1, \ldots, Q_{11n}$ sejam $11n$ pontos ao redor de um círculo coloridos com n cores diferentes, de forma que cada cor seja aplicada a exatamente 11 pontos. Então há como escolher um ponto de cada cor de modo que não selecionemos dois pontos adjacentes.

Para ver isso, escolha, para cada uma das n cores, um ponto dessa cor ao acaso (com probabilidade 1/11). Os $11n$ eventos diferentes E_i que queremos evitar correspondem aos $11n$ pares de pontos adjacentes (Q_i, Q_{i+1}) no círculo, onde Q_{11n} é adjacente a Q_1. Para cada evento,

$$\mathcal{P}\left(E_i\right) = \begin{cases} \frac{1}{121} & \text{se } Q_i \text{ e } Q_{i+1} \text{ têm cores diferentes,} \\ 0 & \text{caso contrário.} \end{cases}.$$

Assim, tomando $p_i = 1/121$ para $i = 1, \ldots, 11n$, temos que $\mathcal{P}\left(E_i\right) \leq p_i$.

Dado um par (Q_i, Q_{i+1}) de pontos adjacentes, existem no máximo outros 21 pares de pontos que usam a cor de Q_i e 21 pares que usam a cor de Q_{i+1}. Um evento E_i é mutuamente independente de uma família $\mathcal{E}$ se as cores correspondentes aos pares de pontos em $\mathcal{E}$ não intersectarem as cores dos pontos de E_i. Pelo raciocínio anterior, E_i é mutuamente independente de todos exceto no máximo 42 outros eventos.

Tomando $d = 42$, e usando que $ep(d+1) \approx 0,966 < 1$, há uma probabilidade positiva (pelo Corolário 3.3) de que nenhum dos eventos ruins ocorra, o que significa que, em particular, existem n pontos não-adjacentes de cores diferentes.

Um importante fortalecimento do Lema Local de Lovász baseado em "cluster expansion" foi desenvolvido na UFMG, como parte da tese de doutorado de Rodrigo Bissacot Proença [11]. Aplicada ao exemplo acima, tal versão melhorada do Lema Local de Lovász permite melhorar a constante 11 para 8 no exemplo acima: Em qualquer coloração de $8n$ pontos num círculo que use cada uma das n cores exatamente 8 vezes, é possível escolher um ponto de cada cor sem selecionar dois pontos adjacentes.

3.3 Modificação do Lema Local de Lovász

Para aplicarmos o caso simétrico do Lema Local de Lovasz, precisamos que cada evento seja mutuamente independente de todos exceto d outros. Assim, se houver muita dependência entre os eventos e a probabilidade de cada evento for razoavelmente grande, torna-se impossível usar o caso simétrico do Lema Local de Lovász. Apresentaremos uma solução parcial através de uma extensão do Lema Local de Lovász que permite diminuir o valor de d utilizado. Vamos decompor (ou codificar) cada evento em termo de um conjunto de variáveis aleatórias que toma valores reais não negativos de tal modo que cada variável associada a um evento dependa somente de poucos outros eventos. Majora-se a probabilidade de cada evento pela soma dos valores esperados das variáveis aleatórias associadas ao evento. Se cada $C_{i,j}$ depender de poucos E_k (os $C_{i,j}$s podem ser altamente dependentes entre si, bem como os E_is), veremos no Teorema 3.5 que ainda é possível obter a mesma conclusão do Lema Local de Lovász. A demonstração de tal versão do Lema Local de Lovász foi feita em [15].

Teorema 3.5. *Dados os eventos $E_1, \ldots, E_m$ e denote, para $I \subset [m]$, $Z(I) = \bigwedge_{i \in I} \overline{E}_i$. Suponhamos que para algum inteiro positivo d podemos definir para cada $i \in [m]$ um número finito de variáveis aleatórias $C_{i,1}, C_{i,2}, \ldots$ não-negativas satisfazendo:*

(i) Qualquer $C_{i,j}$ é mutuamente independente a todos exceto no máximo d eventos E_k, $k \neq i$,

(ii) Para todo $I \subset [m] \setminus \{i\}$, $\mathcal{P}\left(E_i \mid Z(I)\right) \leq \sum_j \mathbf{E}\left[C_{i,j} \mid Z(I)\right]$,

(iii) Para todo $i \in [m]$, $ep_i(d+1) \leq 1$, onde $p_i = \sum_j \mathbf{E}\left[C_{i,j}\right]$.

Então

$$\mathcal{P}\left(\bigwedge_i \overline{E}_i\right) \geq \left(\frac{d}{d+1}\right)^m > 0.$$

Demonstração. Primeiramente vamos demonstrar, por indução em I, que

$$\mathcal{P}\left(E_i \;\middle|\; Z(I)\right) \leq ep_i \text{ para todo } i \notin I. \tag{3.7}$$

Se $|I| = 0$ então se tem pela hipótese (ii) que $\mathcal{P}\left(E_i \;\middle|\; Z(I)\right) = \mathcal{P}\left(E_i\right) \leq p_i \leq ep_i$.

Para o passo indutivo, seja $S_{i,j,I} \subset I$ um conjunto, garantido pela hipótese (i), de tamanho no máximo d tal que $C_{i,j}$ é independente de $\{E_k : k \in I \setminus S_{i,j,I}\}$. Se $S_{i,j,I} = \emptyset$ então

$$\mathbf{E}\left[C_{i,j} \;\middle|\; Z(I)\right] = \mathbf{E}\left[C_{i,j}\right]$$

visto que $C_{i,j}$ são independentes dos E_k para $k \neq i$. Denotando $S'_{i,j,I} = I \setminus S_{i,j,I}$, temos pela definição do esperança (no caso discreto, por simplicidade) que

$$\begin{aligned}
\mathbf{E}\left[C_{i,j} \;\middle|\; Z(I)\right] &= \sum_x x\mathcal{P}\left(C_{i,j} = x \;\middle|\; Z(S_{i,j,I}) \wedge Z(S'_{i,j,I})\right) \\
&= \sum_x x \frac{\mathcal{P}\left(C_{i,j} = x \wedge Z(S_{i,j,I}) \;\middle|\; Z(S'_{i,j,I})\right)}{\mathcal{P}\left(Z(S_{i,j,I}) \;\middle|\; Z(S'_{i,j,I})\right)} \\
&\leq \sum_x x \frac{\mathcal{P}\left(C_{i,j} = x \;\middle|\; Z(S'_{i,j,I})\right)}{\mathcal{P}\left(Z(S_{i,j,I}) \;\middle|\; Z(S'_{i,j,I})\right)} \\
&= \frac{\mathbf{E}\left[C_{i,j} \;\middle|\; Z(S'_{i,j,I})\right]}{\mathcal{P}\left(Z(S_{i,j,I}) \;\middle|\; Z(S'_{i,j,I})\right)}.
\end{aligned}$$

Vamos limitar o denominador da última igualdade. Para isso, seja $S_{i,j,I} = \{\ell_1, \ldots, \ell_s\}$. Então, analogamente a (3.5), o denominador acima é

$$\begin{aligned}
\mathcal{P}\left(Z(S_{i,j,I}) \;\middle|\; Z(S'_{i,j,I})\right) &= \mathcal{P}\left(\bigwedge_{k=1}^{s} \overline{E}_{\ell_k} \;\middle|\; Z(S'_{i,j,I})\right) \\
&= \prod_{k=1}^{s} \mathcal{P}\left(\overline{E}_{\ell_k} \;\middle|\; Z\left(\{\ell_1, \ldots, \ell_{k-1}\} \cup S'_{i,j,I}\right)\right) \\
&= \prod_{k=1}^{s} \left(1 - \mathcal{P}\left(E_{\ell_k} \;\middle|\; Z\left(\{\ell_1, \ldots, \ell_{k-1}\} \cup S'_{i,j,I}\right)\right)\right).
\end{aligned}$$

Por hipótese indutiva,

$$\prod_{k=1}^{s} \left(1 - \mathcal{P}\left(E_{\ell_k} \;\middle|\; Z\left(\{\ell_1, \ldots, \ell_{k-1}\} \cup S'_{i,j,I}\right)\right)\right) \geq \prod_{k=1}^{s} \left(1 - ep_{\ell_k}\right).$$

Pela hipótese (iii), temos que

$$\prod_{k=1}^{s}(1-ep_{\ell_k}) \geq \left(1-\frac{1}{d+1}\right)^s \geq \left(1-\frac{1}{d+1}\right)^d > \frac{1}{e}.$$

onde a última igualdade ocorre pois $\left(1-\frac{1}{x+1}\right)^x$ decresce a $1/e$ quando $x \to \infty$. Assim, usando que $C_{i,j}$ é independente de $Z(S'_{i,j,I})$ por definição,

$$\frac{\mathbf{E}\left[C_{i,j} \,\middle|\, Z(S'_{i,j,I})\right]}{\mathcal{P}\left(Z(S_{i,j}) \,\middle|\, Z(S'_{i,j,I})\right)} \leq e\mathbf{E}\left[C_{i,j} \,\middle|\, Z(S'_{i,j,I})\right] = e\mathbf{E}\left[C_{i,j}\right].$$

Combinando os resultados e usando as hipóteses (ii) e (iii), obtemos que

$$\mathcal{P}\left(E_i \mid Z(I)\right) \leq \sum_j \mathbf{E}\left[C_{i,j} \mid Z(I)\right] \leq e\sum_j \mathbf{E}\left[C_{i,j}\right] = ep_i,$$

provando (3.7) por indução. Para concluir a prova do lema, note que

$$\begin{aligned}\mathcal{P}\left(\bigwedge_{i\in[m]} \overline{E}_i\right) &= \mathcal{P}\left(\overline{E}_1\right)\mathcal{P}\left(\overline{E}_2 \,\middle|\, \overline{E}_1\right)\mathcal{P}\left(\overline{E}_3 \,\middle|\, \overline{E}_1 \wedge \overline{E}_2\right)\cdots\mathcal{P}\left(\overline{E}_m \,\middle|\, \bigwedge_{i=1}^{m-1}\overline{E}_i\right)\\ &= \prod_{i\in[m]}\left(1-\mathcal{P}\left(E_i \,\middle|\, Z\left([i-1]\right)\right)\right) \geq (1-ep_i)^m \geq \left(1-\frac{1}{d+1}\right)^m > 0,\end{aligned}$$

onde a primeira desigualdade segue de (3.7). Isso conclui a prova do lema. □

Capítulo 4

Problemas Minimax

Neste capítulo vamos abordar a problemas inteiros minimax (MIP), dados na Definição 4.3. Para isso vamos garantir pelo Teorema 4.6 que dado um MIP, arredondar aleatoriamente uma solução da relaxação linear produz uma solução viável (sem aumentar muito o valor da função objetivo) com probabilidade positiva. Para fazer a demonstração, vamos precisar dos Teoremas 3.5 e 2.8.

A qualidade da aproximação no Teorema do Teorema 4.6 dependerá do máximo número de linhas não nulas num "bloco de colunas" da matriz A, denotado por t. No Teorema 4.9, substituiremos tal dependência por um outro parâmetro, o máximo número de linhas não nulas numa coluna da matriz A, fornecendo um fator de aproximação muito melhor.

Ambos os teoremas usam a função H definida em 2.12 para quantificar a qualidade da aproximação.

4.1 Definição de problemas minimax (MIP)

Comecemos com um exemplo de um problema de otimização.

Exemplo 4.1. Seja um grafo não direcionado $G = (V, E)$, dois conjuntos $V_1, V_2 \subset V$, uma função $\rho\colon V_1 \to V_2$ e, para cada $i \in V_1$, um conjunto P_i de caminhos em G, cada um ligando i a $\rho(i)$. Para todo i, queremos escolher um caminho $p_i \in P_i$ tal que o número máximo de vezes que qualquer aresta de G é usada seja minimizado.

Para modelar o problema acima, podemos criar variáveis $x_{i,j}$ indicadoras correspondentes a escolher o j-ésimo caminho de P_i. A definição abaixo fornece um jeito de se referir às colunas de uma matriz A de um jeito compatível com os dois índices dos $x_{i,j}$. Informalmente, as colunas da matriz serão divididas em n blocos, onde o primeiro bloco consiste das primeiras ℓ_1 colunas, o segundo bloco das próximas ℓ_2 colunas, e assim por diante.

Definição 4.2. Dados $n \in \mathbb{N}$ e inteiros positivos $\ell_1, \ldots, \ell_n$ inteiros positivos, denote $N = \sum_{i=1}^{n} \ell_i$. Para uma matriz $A \in \mathbb{R}^{m \times N}$, a notação $A_{i,(r,s)}$ ("linha i e s-ésima coluna do r-ésimo bloco") refere-se ao elemento na posição com coordenadas $\left(i, s + \sum_{r'<r} \ell_{r'}\right)$. Além disso, dados valores $(x_{i,j} : i \in [n], j \in [\ell_i])$, denotamos por x o vetor N-dimensional dado por

$$x = \left(x_{1,1}, \ldots, x_{1,\ell_1},\ x_{2,1}, \ldots, x_{2,\ell_2},\ \ldots,\ x_{n,1}, \ldots, x_{n,\ell_n}\right).$$

A notação Ax, portanto, representa um vetor-coluna com m entradas dado por

$$(Ax)_i = \sum_{j=1}^{N} A_{i,j} x_j = \sum_{r=1}^{n} \sum_{s=1}^{\ell_r} A_{i,(r,s)} x_{r,s}.$$

Com essa notação, podemos definir os problemas inteiros minimax (MIP).

Definição 4.3 (MIP)**.** Seja $(\ell_i)_{i=1}^n$ uma sequência de inteiros positivos e $N = \sum_{i\in[n]} \ell_i$. Dadas as variáveis $\{x_{i,j} : i \in [n], j \in [\ell_i]\}$, uma variável extra W e uma matriz $A \in [0,1]^{m\times N}$, um programa inteiro minimax (MIP) é da forma

$$\begin{array}{llll} \text{minimizar} & W & & \\ \text{sujeito a} & \sum_{j=1}^{\ell_i} x_{i,j} = 1, & 1 \le i \le n & \\ & Ax \le \vec{W} & & , \\ \text{no domínio} & x_{i,j} \in \{0,1\}, & 1 \le i \le n, & 1 \le j \le \ell_i \\ & W \in \mathbb{R}_{\ge 0} & & \end{array}$$

onde $\vec{W}$ denota o vetor m-dimensional $(W, \ldots, W)^\intercal$ e $Ax \le \vec{W}$ denota que cada componente do vetor Ax é menor ou igual a W.

Observe que, apesar do lado direito da desigualdade $Ax \le \vec{W}$ conter uma variável, esse ainda é um programa linear. Para convertê-lo para uma forma mais tradicional, basta colocar uma coluna a mais na matriz A (com entradas iguais a -1) correspondente à variável W.

Definição 4.4. Denotamos por $g = g(A)$ como a maior soma de uma coluna de A, e $a = a(A)$ como o número máximo de entradas não nulas de uma coluna de A. Além disso, definimos

$$t_r = \left| \left\{ 1 \le i \le n : A_{i,(r,s)} \ne 0 \text{ para algum } 1 \le s \le \ell_r \right\} \right|,$$

o número de linhas que possuem alguma entrada não-nula em alguma das colunas correspondentes ao r-ésimo bloco de variáveis, e

$$t = t(A) = \max_{1 \le r \le n} t_r. \tag{4.1}$$

Como as entradas de A estão entre 0 e 1, vale que

$$g \le a \le t \le \min\{m, a \cdot \max_{i\in[n]} \ell_i\}. \tag{4.2}$$

Observação 4.5. Podemos formular o problema do exemplo 4.1 como um MIP. Seja $m = |E|$, $n = |V_1|$, e $\ell_i = |P_i|$, para $1 \le i \le n$. A variável $x_{i,j}$ será a variável indicadora correspondente a escolher o j-ésimo caminho de P_i. Se $E = \{e_1, \ldots, e_k\}$, a matriz A será dada por $a_{k,(i,j)} = 1$ se o j-ésimo caminho de P_i contém a aresta e_k e $a_{k,(i,j)} = 0$ caso contrário. Assim, a condição $Ax \le \vec{W}$ é equivalente a dizer que nenhuma aresta do grafo é usada mais de W vezes.

4.2 Arredondamento aleatório para MIPs

O resultado abaixo fornece boas soluções inteiras se o parâmetro t do MIP satisfaz $t \ll m$. Além disso, ele será utilizado na prova do Teorema 4.9. Sua prova se baseia na decomposição fornecida pelo Lema 2.8 parte (b) e na variação do Lema Local de Lovász provada no Teorema 3.5.

A relaxação linear de um MIP é o programa linear obtido trocando-se as restrições $x_{i,j} \in \{0,1\}$ por restrições $x_{i,j} \geq 0$ (a restrição $x_{i,j} \leq 1$ não é necessária porque $\sum_j x_{i,j} = 1$ para todo i). Existem algoritmos eficientes para resolver programas lineares como esse (ver, por exemplo, [9]).

Teorema 4.6 (Teorema 4.2 de[15])**.** *Dado um MIP em conformidade com a Definição 4.3, seja y^* é o valor ótimo da função objetivo para a relaxação linear. Então existe solução inteira com valor no máximo*

$$y^* + \left\lceil \min\{y^*, m\} \cdot H\left(\min\{y^*, m\}, 1/(et)\right) \right\rceil,$$

onde H é como definida no Lema 2.12.

Demonstração. Seja x^* uma solução ótima da relaxação linear do MIP, e y^* o valor correspondente da função objetivo. Por definição de MIP, vale que

$$Ax^* \leq \vec{y^*}, \tag{4.3}$$

onde $\vec{y^*} = y^*(1, \ldots, 1)^\intercal$ é um vetor m-dimensional. No nosso caso, é possível provar que a relaxação linear

$$\begin{array}{ll} \text{minimizar} & W \\ \text{sujeito a} & \sum_{j=1}^{\ell_i} x_{i,j} = 1, \quad 1 \leq i \leq n \\ & Ax \leq \vec{W} \\ & x_{i,j} \geq 0, \quad 1 \leq i \leq n, \quad 1 \leq j \leq \ell_i \end{array}$$

admite uma ótimo *básico*, isto é, uma solução em que um conjunto de N restrições linearmente independentes (dentre todas as $N + m + n$ restrições) vale com igualdade (ver Definição 4.4.2 de [9]). Deduz-se que no máximo $m + n$ das coordenadas são não-nulas numa solução básica. Iremos supor que x^* tem essa propriedade. Defina

$$n_1 = \left| \left\{ 1 \leq i \leq n : 0 < x^*_{i,j} < 1 \text{ para algum } 1 \leq j \leq \ell_i \right\} \right|,$$

o número de "blocos" de coordenadas que possuem entradas fracionárias. Como as coordenadas de cada bloco somam a 1, cada bloco possui pelo menos uma coordenada não-nula, e os blocos contados por n_1 possuem pelo menos duas coordenadas não-nulas. Como no máximo $m + n$ coordenadas são não-nulas, $n_1 + n \leq m + n$, e portanto $n_1 \leq m$.

A seguir, iremos descrever um procedimento de arredondamento aleatório. Como os blocos de x^* sem entradas fracionárias não precisarão ser arredondados, daqui em diante iremos supor, sem perda de generalidade, que tais blocos não existem. Como as entradas de A estão entre 0 e 1 e restam $n_1 \leq m$ blocos, tal suposição implica em $Ax^* \leq m$, e portanto em $y^* \leq m$. Isso explica o termo $\min\{y^*, m\}$ na fórmula do enunciado.

Nosso objetivo é mostrar que podemos arredondar a solução x^* de modo a obter uma solução inteira com a propriedade enunciada. Para fazer o arredondamento, lembre-se que $\sum_j x^*_{i,j} = 1$; definimos variáveis aleatórias auxiliares independentes Y_i ($i = 1, \ldots, n$) por $\mathcal{P}(Y_i = j) = x^*_{i,j}$, para $1 \leq j \leq \ell_i$. Nossa solução arredondada z será então dada por

$$z_{i,j} = \chi(Y_i = j).$$

Em palavras, isso significa que, para cada i, arredondamos x^*_{i,Y_i} para 1 e as demais $x^*_{i,j}$ para 0, onde Y_i é uma coluna do i-ésimo grupo escolhida ao acaso. Com esse arredondamento, temos que

$$\mathbf{E}\left[z_{i,j}\right] = \mathcal{P}\left(Y_i = j\right) = x^*_{i,j}.$$

Seja y o valor da função objetivo para o arredondamento aleatório acima. Definindo

$$k = \left\lceil y^* \cdot H\left(y^*, 1/et\right)\right\rceil,$$

queremos mostrar que $\mathcal{P}\left(y \leq y^* + k\right) > 0$. Para isso, definindo para $i = 1, \ldots, m$,

$$E_i \equiv \left[(Az)_i > (Ax^*)_i + k\right],$$

temos que

$$\bigwedge_{i\in[m]} \overline{E}_i \equiv \left[y \leq y^* + k\right].$$

Nosso objetivo é aplicar o Teorema 3.5 para mostrar que $\mathcal{P}\left(\wedge_{i=1}^m \overline{E}_i\right) > 0$. Por (4.3), temos que

$$\mu_i := \mathbf{E}\left[(Az)_i\right] = \mathbf{E}\left[\sum_{j=1}^{n}\sum_{k=1}^{\ell_j} a_{i,(r,s)} z_{(r,s)}\right] = \sum_{j=1}^{N} a_{i,j}\mathbf{E}\left[z_j\right] = \sum_{j=1}^{N} a_{i,j}x^*_{i,j} = (Ax^*)_i \leq y^*. \quad (4.4)$$

Definimos as variáveis aleatórias

$$Z_{i,r} = \sum_{s\in[\ell_r]} A_{i,(r,s)} z_{r,s} = \sum_{s\in[\ell_r]} A_{i,(r,s)}\chi(Y_r = s),$$

e observemos que, para cada i, as variáveis aleatórias $\{Z_{i,r} : r \in [n]\}$ em $[0, 1]$ são independentes pois $Z_{i,r}$ depende apenas da variável aleatória Y_r. Assim temos que

$$E_i \equiv \left[\sum_{r\in[n]} Z_{i,r} \geq (Ax^*)_i + k\right].$$

Seja $u = \binom{n}{k}$ e $\mathcal{S} = \{X_1, \ldots, X_u\}$ a família de todos os subconjuntos de tamanho k de $[n]$. Motivados pela prova do Teorema 2.8, definimos as variáveis aleatórias $\{C_{i,j} : i \in [m], j \in [u]\}$ por

$$C_{i,j} = \frac{1}{\binom{\mu_i+k}{k}} \prod_{v\in X_j} Z_{i,v},$$

de modo que $C_{i,j} \in [0, 1]$ e

$$\sum_j C_{i,j} = \frac{1}{\binom{\mu_i+k}{k}} S_k(Z_{i,1}, \ldots, Z_{i,n}).$$

Assim, pelo Teorema 2.8 com $\delta_i = k/\mu_i$,

$$\mathcal{P}\left(E_i \mid Z\right) \leq \sum_{j=1}^{n} \mathbf{E}\left[C_{i,j} \mid Z\right] \qquad \text{e} \qquad \mathcal{P}\left(E_i\right) \leq \sum_{j=1}^{n} \mathbf{E}\left[C_{i,j}\right] \leq G(\mu_i, \delta_i).$$

Pelas definição de δ_i e usando que $\mu_i \leq y^*$ por (4.4),

$$G\left(\mu_i, \delta_i\right) = G\left(\mu_i, \frac{k}{\mu_i}\right) \leq G\left(y^*, \frac{k}{y^*}\right). \tag{4.5}$$

Como $G(x, y)$ é decrescente na segunda variável pelo Lema 2.10, temos que

$$G\left(y^*, \frac{k}{y^*}\right) \leq G\left(y^*, H\left(y^*, 1/et\right)\right). \tag{4.6}$$

Finalmente, usando a definição de H,

$$G\left(y^*, H\left(y^*, 1/et\right)\right) \leq \frac{1}{\left\lceil y^* H\left(y^*, 1/et\right)\right\rceil} \cdot \frac{1}{et} = \frac{1}{etk}. \tag{4.7}$$

Combinando (4.5), (4.6) e (4.7), temos que

$$\mathcal{P}\left(E_i\right) \leq \frac{1}{etk} =: p_i,$$

Um dado $Z_{i,r}$ pode depender apenas dos eventos E_j tais que $A_{j,(r,s)} \neq 0$ para algum $1 \leq s \leq \ell_r$, pois do contrário a variável Y_r será multiplicada por zero. Lembrando de (4.1) e excluindo o caso $i \neq j$, vemos que cada $Z_{i,r}$ é independente de todos exceto $t-1$ eventos E_j. Assim, cada $C_{i,j}$, sendo produto de k fatores da forma $Z_{i,r}$, é independente de todos exceto $k(t-1)$ eventos E_j. Podemos então tomar $d = k(t-1)+1$ e calcular

$$ep_i(d+1) = ep_i(k(t-1)+1) = \frac{(k(t-1)+1)}{kt} = 1 - \frac{1}{t} + \frac{1}{kt} \leq 1,$$

pois $k \geq 1$. Pelo Teorema 3.5 temos que

$$\mathcal{P}\left(\bigwedge_{i \in [m]} \overline{E}_i\right) > 0.$$

como desejado. $\square$

Nosso próximo objetivo é substituir o parâmetro t no teorema acima pelo parâmetro a da Definição 4.4. Para isso, vamos primeiro investigar no lema abaixo como a função H, definida no Lema 2.12, se comporta com relação a esses parâmetros.

Lema 4.7. *Para todo $K_0 \geq 2$ existe $C > 0$ com a seguinte propriedade. Sejam a, t, μ com $1 < a \leq t$. Se $t \leq \max\left\{\mu^7, K_0, a^4\right\}$, então*

$$H(\mu, 1/et) \leq C \cdot H(\mu, 1/a).$$

Demonstração. Lembre que a função $H(\mu, p)$ foi definida em (2.4) em casos. Ela satisfaz $H \geq 8e$ no primeiro caso, $H = 8$ no segundo e $H \leq 6$ no terceiro. Para μ fixo, passamos do primeiro para o segundo para o terceiro caso a medida que aumentamos p. Dividimos em três casos de acordo com a magnitude relativa dos parâmetros.

Caso 1: $e^{e\mu} \leq a \leq et$. Seja $\varphi(x) = 8x/\ln x$. Lembre que $\varphi(x) \geq 8e$ para todo $x \geq e$. Além disso, como $\ln x$ é crescente, vale que $\varphi(kx) \leq k\varphi(x)$ para todo $k \geq 1$. Queremos mostrar que

$$\varphi\left(\frac{\ln(et)}{\mu}\right) \leq C\varphi\left(\frac{\ln(a)}{\mu}\right).$$

Se $t \leq a^4$, então

$$\varphi\left(\frac{\ln(et)}{\mu}\right) \leq \varphi\left(\frac{\ln(a^6)}{\mu}\right) = \varphi\left(6 \cdot \frac{\ln(a)}{\mu}\right) \leq 6 \cdot \varphi\left(\frac{\ln(a)}{\mu}\right).$$

Se $t \leq \mu^7$, podemos usar que $\varphi(x)$ é função crescente para $x \geq e$, e que $\ln(e\mu^7)/\mu \leq 3$, para concluir que, se $C \geq 3/e$,

$$\varphi\left(\frac{\ln(et)}{\mu}\right) \leq \varphi(3) < 8 \cdot 3 \leq C \cdot (8e) \leq C\varphi\left(\frac{\ln(a)}{\mu}\right).$$

Se $t \leq K_0$, então, como φ é crescente e $a \geq 2$,

$$\varphi\left(\frac{\ln(et)}{\mu}\right) \leq \varphi\left(\frac{\ln(eK_0)}{\mu}\right) \leq \frac{\ln(eK_0)}{\ln 2} \cdot \varphi\left(\frac{\ln(a)}{\mu}\right).$$

Caso 2: $a \leq e^{e\mu} \leq et$. Nesse caso, observemos que $2 \leq a \leq e^{e\mu}$ implica que $\mu \geq \ln(2)/e \geq 1/4$. Queremos mostrar que

$$8 \cdot \frac{\ln(et)}{\mu} \Bigg/ \ln\left(\frac{\ln(et)}{\mu}\right) \leq C\sqrt{\frac{6}{\mu}\ln(\mu + a)}.$$

Se $t \leq a^4$, então das hipóteses do caso concluímos que $t \leq e^{4e\mu}$ e $a \geq e^{(e\mu-1)/4}$. Assim, é possível mostrar que o lado esquerdo é no máximo uma constante e o lado direito é no mínimo uma constante.

Se $t \leq \mu^7$, então o lado esquerdo é no máximo uma constante. Além disso, como $e^{e\mu} \leq e\mu^7$, o que só pode ocorrer se $\mu \leq 4$, o lado direito é no mínimo uma constante.

Se $t \leq K_0$, temos que

$$8 \cdot \frac{\ln(et)}{\mu} \Bigg/ \ln\left(\frac{\ln(et)}{\mu}\right) \leq \frac{8\ln(eK_0)}{\mu} \leq C\sqrt{\frac{1}{\mu}},$$

pois μ é limitado inferiormente.

Caso 3: $a \leq et \leq e^{e\mu}$. Para o terceiro caso de (2.4), queremos mostrar que

$$\sqrt{\frac{1}{\mu}\ln(\mu + et)} \leq C\sqrt{\frac{1}{\mu}\ln(\mu + a)}.$$

Se $t \le a^4$, então como $3a^4 \le a^6$, temos

$$\ln(\mu + et) \le \ln(\mu + 3a^4) \le \ln(\mu + a^6) \le \ln\left((\mu + a)^6\right) = 6\ln(\mu + a)$$

Se $t \le \mu^7$, então, como $C \ge \sqrt{8}$ e $a \ge 2$, o resultado desejado segue de

$$\ln(\mu + et) \le \ln(\mu + e\mu^7) \le \ln\left((\mu + 2)^8\right) = 8\ln(\mu + 2).$$

Se $t < K_0$, podemos cotar

$$\ln(\mu + eK_0) = \ln(eK_0/2) + \ln\left(\frac{\mu}{eK_0} + 2\right) \le \left(1 + \frac{\ln(eK_0/2)}{\ln 2}\right) \cdot \ln(\mu + 2) \le C\ln(\mu + a).$$

□

Antes de provarmos o resultado principal dessa seção usando o Teorema 4.6 e o Lema 4.7, façamos uma rápida observação.

Observação 4.8. Se $0 \le x \le 1$ e $0 \le y \le 1/2$, então

$$\frac{1+x}{1-y} = (1+x)\left(1 + \frac{y}{1-y}\right) \le (1+x)(1+2y) = 1 + x + 2y + x(2y) \le 1 + 4\max(x, y).$$

Teorema 4.9 (Teorema 2.5 de [15])**.** *Para qualquer MIP existe uma solução inteira de valor no máximo* $y^* + O(1) + O(\min\{y^*, m\} \cdot H(\min\{y^*, m\}, 1/a))$*, onde* m *e* a *estão na Definição 4.3 e* t *é definido em* (4.1)*.*

Demonstração. Seja K_0 uma constante suficientemente grande, e y^* o valor da solução ótima da relaxação linear. Podemos assumir, como no teorema anterior, que $y^* \le m$. Seja $K_0 > 4$ uma constante a ser escolhida posteriormente; o Teorema 4.6 garante uma solução inteira de valor no máximo

$$y^* + \lceil y^* \cdot H(y^*, 1/et)\rceil.$$

Podemos supor que

$$t > \max\left((y^*)^7, K_0, a^4\right), \tag{4.8}$$

pois do contrário o Lema 4.7 mostra que a conclusão do Teorema 4.6 é mais forte que o que procuramos. Além disso, note que se $y^* < t^{-1/7}$, então vale que $y^* \le t^{-1/7} \le \ln(et)/e$ para $t \ge K_0 > 4$. Assim, a solução dada pelo Teorema 4.6 satisfaz os requisitos, pois

$$y^* \cdot H(y^*, 1/et) \le 8\ln(et) \Big/ \ln\left(\frac{\ln(et)}{y^*}\right) \le 8 \cdot \frac{\ln(et)}{\ln(t^{1/7}\ln(et))} \le 56 \cdot \frac{\ln(t) + 1}{\ln t} = O(1).$$

Iremos dividir em dois casos, de acordo com a magnitude de y^*.

Caso 1: $t^{-1/7} < y^* < 1$**.** Na prova que iremos expor, vamos definir eventos ruins E_i e usar o Teorema 2.8 para associar a cada evento ruim um número finito de variáveis aleatórias $C_{i,j}$ não negativas que verifiquem as hipóteses do Teorema 3.5. Seja

$$\alpha = \frac{\log^5 t}{y^*}.$$

Definimos $x'_{i,j} = \alpha x^*_{i,j}$ e geramos o vetor z arredondando cada coordenada de x'. Formalmente, z é dado por $z_{i,j} = \lfloor x'_{i,j} \rfloor + X_{i,j}$, onde $X_{i,j} \in \{0,1\}$ e $\mathcal{P}\left(X_{i,j} = 1\right) = x'_{i,j} - \lfloor x'_{i,j} \rfloor$. O arredondamento aleatório satisfaz

$$\mu_i := \mathbf{E}\left[(Az)_i\right] = \sum_{j=1} a_{i,j} \mathbf{E}\left[z_{i,j}\right] = \alpha\left(Ax^*\right)_i \leq \alpha y^*. \tag{4.9}$$

Denote também

$$\mu'_i := \mu_i - (A\lfloor x' \rfloor)_i = \mathbf{E}\left[(AX)_i\right]. \tag{4.10}$$

Iremos agora mostrar que as variáveis aleatórias $(Az)_i$ e $\sum_i z_{i,j}$ são concentradas. Para tal, seja $K_1 > 0$ suficientemente grande, e defina

$$k = \left\lceil K_1 \log^3 t \right\rceil = \left\lceil \alpha y^* \cdot \frac{K_1}{\log^2 t} \right\rceil.$$

Os eventos ruins serão definidos por

$$E_i \equiv \left[(Az)_i > \mu_i + k\right] \qquad i = 1, \ldots, m, \tag{4.11}$$

$$E_{i+m} \equiv \left[\left|\sum_j z_{i,j} - \alpha\right| > K_1 \cdot \sqrt{\alpha \log t}\right] \qquad i = 1, \ldots, n. \tag{4.12}$$

Há forte dependência entre os eventos. Para lidar com tal situação, iremos decompor os E_i $(1 \leq i \leq m)$. Para isso, defina as variáveis aleatórias

$$Z_{i,(r,s)} = A_{i,(r,s)} X_{r,s}.$$

Para um i fixo, as variáveis acima definidas são independentes, pois estamos arredondando cada coordenada independentemente (ao contrário do Teorema 4.6). Relembrando (4.10), temos que

$$E_i \equiv \left[\sum_{(r,s)} Z_{i,(r,s)} > \mu'_i + k\right].$$

Seja $u = \binom{N}{k}$ e $\mathcal{S} = \{X_1, \ldots, X_u\}$ a família de todos os subconjuntos de tamanho k de $\{(r,s) : 1 \leq r \leq n, 1 \leq s \leq \ell_r\}$. Definimos as variáveis aleatórias $\{C_{i,j} : i \in [m], j \in [u]\}$ por

$$C_{i,j} = \frac{1}{\binom{\mu'_i + k}{k}} \prod_{v \in X_j} Z_{i,v}.$$

Como antes, cada $C_{i,j}$ é um produto de k parcelas do tipo $Z_{i,v}$. Temos que:

- Para $1 \leq j \leq m$, um dado $Z_{i,v}$ pode depender apenas dos eventos E_j tais que $A_{j,v} \neq 0$. Lembrando da Definição 4.4 e excluindo o caso $i = j$, temos $a - 1$ eventos nessa categoria.
- Para $1 \leq j \leq n$, um dado $Z_{i,v}$ pode depender apenas dos eventos E_{m+i} tais que $(i,a) \in X_v$ para algum a. Como $|X_v| = k$, temos k eventos nessa categoria.

Assim, $C_{i,j}$ depende de no máximo $k(a + k - 1)$ eventos ruins. Além disso, para $1 \leq i \leq m$, definimos $C_{m+i,1} = \chi(E_{m+i})$. Um raciocínio análogo mostra que cada $C_{m+i,1}$, depende de no máximo t dos E_is pois, para r fixo, as variáveis $z_{(r,s)}$ são independentes dos E_i correspondentes a linhas que só tem zeros no r-ésimo grupo.

Pelo Teorema 2.8 com $\delta_i = k/\mu_i$ e pela Proposição 2.11 (lembrando que $\mu_i' \leq \mu_i \leq \alpha y^*$ por (4.9) e (4.10)), temos que

$$\mathcal{P}(E_i) \leq G\left(\mu_i', \frac{k}{\mu_i'}\right) \leq G\left(\alpha y^*, \frac{k}{\alpha y^*}\right).$$

Pelo Lema 2.9 com $\mu = \alpha y^*$ e $\delta = k/\alpha y^*$,

$$G\left(\alpha y^*, \frac{k}{\alpha y^*}\right) \leq \exp\left(-\min\left\{\frac{\delta\mu}{3}, \frac{\delta^2\mu}{3}\right\}\right) \leq \exp\left(-\frac{1}{3}\min\left\{K_1 \log^3 t, (K_1)^2 \log t\right\}\right).$$

Assim, temos que $\mathcal{P}(E_i) < t^{-K_1/3}$. Segue-se que, se K_1 for maior que uma constante absoluta,

$$ep_i(d+1) = et^{-K_1/3}k(t+k-1) \leq 1.$$

Pelo Teorema 3.5 temos que $\mathcal{P}\left(\bigwedge_{i=1} \overline{E}_i\right) > 0$. Podemos então fixar uma realização de z tal que nenhum dos E_i ocorra. Como as desigualdades em (4.12) não valem e z tem coordenadas inteiras, para cada i existem no máximo $\alpha + K_1\sqrt{\alpha \log t}$ valores de j tal que $z_{i,j}$ é não-nulo. Além disso, se definirmos x'' por

$$x''_{i,j} = \frac{1}{\sum_k z_{i,k}} \cdot z_{i,j},$$

temos que, usando que (4.11) e (4.12) não valem,

$$(Ax'')_i = \frac{1}{\sum_k z_{i,k}}(Az)_i \leq \frac{\mu_i + k}{\alpha - K_1\sqrt{\alpha \log t}} \leq \frac{y^*\left(1 + \frac{K_1}{\log^2 t} + \frac{1}{\alpha y^*}\right)}{1 - K_1\sqrt{(\log t)/\alpha}}$$

Pela Observação 4.8 e usando que $1/(\log^2 t) \geq \sqrt{(\log t)/\alpha}$,

$$\frac{y^*\left(1 + \frac{K_1}{\log^2 t} + \frac{1}{\alpha y^*}\right)}{1 - K_1\sqrt{(\log t)/\alpha}} \leq y^*\left(1 + \frac{6K_1}{\log^2 t}\right).$$

Resumindo: Começando com um vetor x^* com valor objetivo y^*, obtemos um novo vetor x'' com valor objetivo $(y'')^*$ que satisfaz

$$(y'')^* \leq y^*\left(1 + \frac{6K_1}{\log^2 t}\right) \tag{4.13}$$

tal que no máximo

$$\alpha + K_1\sqrt{\alpha \log t} = O(\alpha)$$

coordenadas de x'' são não nulas. Portanto, se eliminarmos da matriz A as colunas correspondentes às entradas nulas de x'', teremos um novo MIP em que $\ell_i = O(\alpha)$ para todo i. Podemos tomar o vetor x'' sem as entradas não-nulas como uma solução da relaxação linear do novo problema, lembrando que o valor objetivo de tal solução satisfaz 4.13. Qualquer solução inteira do novo problema pode ser transformada numa solução inteira do MIP original colocando zeros nas posições apropriadas. Além disso, lembrando de (4.2), vemos que o novo MIP satisfaz $t = O(a\alpha)$. Assim, por (4.8), $t_{\text{novo}} = O(t^{1/4+1/7} \log^5 t)$, e portanto tal procedimento reduziu o valor de t.

Para resolver o novo MIP, iteraremos o procedimento até que (4.8) seja falsa (ou seja, até podermos aplicar o Teorema 4.6 para obter uma solução inteira diretamente). Para entender como as várias alteram o valor da função objetivo, seja $A^{(i)}$ a matriz da i-ésima iteração ($A^{(0)} = A$); definiremos a sequência

$$\begin{aligned} t_0 &= t, \\ t_{i+1} &= t_i^{1/4+1/7} \log^5 t_i. \end{aligned}$$

Pelas considerações anteriores, $t(A^{(i)}) \leq t_i$. Seja L o primeiro índice tal que $t_L \leq K_0$; por causa de (4.8), iremos fazer no máximo L iterações. Se K_0 é grande o suficiente, então para $i < L$,

$$\log t_{i+1} = \frac{11}{28} \log t_i + 5 \log(\log t_i)) \leq \frac{\log t_i}{2}.$$

Assim, $L \leq \log_2 \log t$. Podemos reescrever a equação acima e obter

$$\frac{1}{\log^2 t_i} \leq \frac{1}{4} \cdot \frac{1}{\log^2 t_{i+1}}.$$

Combinando tal decaimento com (4.13), o valor objetivo final é no máximo

$$y^* \prod_{i=0}^{L-1} \left(1 + \frac{6K_1}{\log^2 t_i}\right) \leq y^* \exp\left(\sum_{i=0}^{L-1} \frac{6K_1}{\log^2 t_i}\right) \leq y^* \exp\left(\frac{6K_1}{\log^2 t_L} \sum_{i=1}^{L} \frac{1}{4^i}\right) \leq y^* \exp\left(\frac{6K_1}{\log^2 K_0}\right). \tag{4.14}$$

Logo, o vetor obtido depois das iterações tem valor objetivo no máximo $y^* + O(1)$, e não satisfaz (4.8); assim, uma aplicação do Teorema 4.6 fornece uma solução de valor

$$y^* + O(1) + C \cdot H(y^* + O(1), 1/a),$$

pelo Lema 4.7. Para concluir a prova, observamos que H é decrescente na primeira variável.
Caso 2: $1 \leq y^* < t^{1/7}$**.** Tomamos

$$\alpha = \left(y^*\right)^2 (\log^5 t)$$

e para um K_1 suficientemente grande definimos

$$k = \left\lceil K_1 (y^*)^{3/2} \log^3 t \right\rceil = \left\lceil \alpha y^* \cdot \frac{K_1}{(y^*)^{3/2} \log^2 t} \right\rceil.$$

Definindo os eventos E_i como anteriormente, obtemos pelo mesmo argumento que

$$\mathcal{P}(E_i) \leq G\left(\alpha y^*, \frac{k}{\alpha y^*}\right).$$

Pelo Lema 2.9 com $\mu = \alpha y^*$ e $\delta = k/\alpha y^*$,

$$G\left(\alpha y^*, \frac{k}{\alpha y^*}\right) \leq \exp\left(-\min\left\{\frac{\delta\mu}{3}, \frac{\delta^2\mu}{3}\right\}\right) \leq \exp\left(-\frac{1}{3}\min\left\{K_1(y^*)^{3/2}\log^3 t, K_1^2 \log t\right\}\right).$$

Pelo Teorema 3.5 temos que $\mathcal{P}\left(\bigwedge_{i=1} \overline{E}_i\right) > 0$. Podemos então tomar um z para o qual nenhum dos eventos ruins vale e normalizá-lo, como antes, para obter um vetor x'' que satisfaz

$$(Ax'')_i \leq \frac{\mu_i + k}{\alpha - K_1\sqrt{\alpha \log t}} \leq \frac{y^*\left(1 + \frac{K_1}{(y^*)^{3/2}\log^2 t} + \frac{1}{\alpha y^*}\right)}{1 - K_1\sqrt{(\log t)/\alpha}}.$$

Pela Observação 4.8 e usando que $1/((y^*)^{3/2}\log^2 t) \leq \sqrt{(\log t)/\alpha} = 1/(y^* \log^2 t)$ (o contrário do caso anterior!), temos que

$$\frac{y^*\left(1 + \frac{K_1}{(y^*)^{3/2}\log^2 t} + \frac{1}{\alpha y^*}\right)}{1 - K_1\sqrt{(\log t)/\alpha}} \leq y^*\left(1 + \frac{6K_1}{y^* \log^2 t}\right).$$

Deletando colunas de A correspondentes a zeros de x'', como antes, obtemos um novo MIP que satisfaz $t_{\text{novo}} = O(a\alpha) = O(t^{1/4+2/7}\log^5 t)$. Como $1/4 + 2/7 < 1$, o mesmo decaimento exponencial de $\log t_i$ do caso anterior continua ocorrendo. Assim, depois de todas as iterações, o análogo de (4.14) é que o valor final da função objetivo é no máximo

$$y^* \prod_{i=0}^{L-1}\left(1 + \frac{6K_1}{y^* \log^2 t_i}\right) \leq y^* \exp\left(\frac{6K_1}{y^* \log^2 K_0}\right).$$

Como $e^x \leq 1 + 2x$ se $0 \leq x \leq 1$, vemos que

$$y^* \exp\left(\frac{6K_1}{y^* \log^2 K_0}\right) = y^* + O(1).$$

Assim, as iterações não mudam o valor da função objetivo por mais que uma constante. Uma aplicação do Teorema 4.6, como antes, completa a prova.

$\square$

Capítulo 5

Problemas de cobertura

Vamos fazer com essas ideias as definições multicritérios de forma geral CIP dada pela Definição 5.1. Os principais resultados são os Teoremas 5.16, sobre boas aproximações para CIPs e um teorema análogo (Teorema 5.19). Para isso, usaremos o Lema 5.8. Além disso, desenvolveremos uma versão algorítmica do resultado.

Definição 5.1 (Programa inteiro de cobertura (CIP))**.** Dado $A \in [0,1]^{m\times n}$, $b \in [1,\infty)^m$ e $c \in [0,1]^n$ com $\max_j c_j = 1$, um programa inteiro de cobertura (CIP) é um programa da forma

$$\begin{array}{ll} \text{minimizar} & c^\intercal \cdot x \\ \text{sujeito a} & Ax \geq b \\ & x_i \in \mathbb{Z}_{\geq 0},\ 1 \leq i \leq n. \end{array}$$

Se $A \in \{0,1\}^{m\times n}$, suporemos também que b é um vetor de inteiros assume valores inteiros. Definimos $B = \min_i b_i$ e a como o máximo número de entradas não-nulas em qualquer coluna de A. Um CIP é chamado sem peso se $c_j = 1$ para todo j, e ponderado caso contrário.

Exemplo 5.2 (Cobertura de um hipergrafo por arestas)**.** Seja $\mathcal{H} = \big(V, \mathcal{E} = \{e_1,\dots,e_m\}\big)$ um hipergrafo. O problema de cobertura por arestas consiste em encontrar o menor número de arestas cuja união é V. O CIP que o modela tem uma variável para cada aresta e uma inequação para cada vértice, e é dado por:

$$\begin{array}{ll} \text{minimizar} & \displaystyle\sum_{i=1}^{m} x_i \\ \text{sujeito a} & Ax \geq (1,\dots,1)^\intercal \\ & x_i \in \{0,1\}, \qquad 1 \leq i \leq m, \end{array}$$

onde

$$A = \big(a_{i,j}\big) = \begin{cases} 1 & \text{se } v_i \in e_j \\ 0 & \text{se } v_i \notin e_j \end{cases},$$

é a transposta da matriz de incidência de $\mathcal{H}$. Para esse problema, o parâmetro a é o maior número de vértices numa aresta de $\mathcal{H}$.

Definição 5.3 (Arredondamento aleatório generalizado)**.** Dado um vetor $p = (p_1,\dots,p_n) \in [0,1]^n$ e um parâmetro $\alpha > 1$, o arredondamento aleatório generalizado do vetor-solução x^* é o vetor z dado por $z_j = \lfloor \alpha x_j^* \rfloor + X_j$, onde $X_1, \ldots, X_n \in \{0,1\}$ são variáveis aleatórias independentes com $\mathcal{P}\left(X_j = 1\right) = p_j$.

Definição 5.4 (Arredondamento aleatório padrão)**.** O arredondamento aleatório padrão é definido tomando-se $p_j = \alpha x_j^* - \lfloor \alpha x_j^* \rfloor$ no arredondamento aleatório generalizado.

Para o arredondamento aleatório padrão, vale que

$$\mathbf{E}\left[z_j\right] = \lfloor \alpha x_j^* \rfloor + p_j = \alpha x_j^*. \tag{5.1}$$

Definição 5.5. A função g é dada por

$$g(B, \alpha) = \left(\alpha \cdot e^{-(\alpha-1)}\right)^B.$$

Frequentemente, a função g será aplicada com parâmetro B da Definição 5.1 e o α correspondente a um arredondamento aleatório (ver Definição 5.3).

Observação 5.6. Por convexidade da função $y \mapsto e^y$, temos que $1 + y \leq e^y$ e portanto $e^{-y} \leq 1/(1+y)$. Tomando $y = x/(1+x)$ e aplicando logaritmo, concluímos que

$$\ln\left(1 - \frac{x}{1+x}\right) < -\frac{x}{1+x}.$$

Assim,

$$-\ln(1+x) = \ln\left(1 - \frac{x}{1+x}\right) \leq -\frac{x}{1+x}.$$

Tal desigualdade será útil na demonstração do Lema 5.7.

Lema 5.7. *Sejam X_j ($1 \leq j \leq n$) variáveis aleatórias correspondentes ao arredondamento randomizado padrão com parâmetro α. Seja s o vetor dado por $s_j = \lfloor \alpha x_j^* \rfloor$ e $X = (X_1, \cdots, X_n)$. Para $1 \leq i \leq m$, defina*

$$\mu_i = \mathbf{E}\left[A_i \cdot X\right], \qquad \delta_i = 1 - \frac{b_i - A_i \cdot s}{\mu_i}, \qquad E_i \equiv [A_i \cdot X < (1 - \delta_i)\, \mu_i],$$

onde A_i denota a i-ésima linha da matriz A. Então

$$\mathcal{P}\left(E_i\right) \leq \frac{\mathbf{E}\left[(1-\delta_i)^{A_i \cdot X}\right]}{(1-\delta_i)^{(1-\delta_i)\mu_i}} \leq g\left(B, \alpha\right),$$

onde B é como na Definição 5.1.

Demonstração. Note primeiro que $\delta \geq 0$, pois

$$\mu_i + A_i \cdot s = \mathbf{E}\left[A_i \cdot X\right] + A_i \cdot s = \mathbf{E}\left[A_i(X+s)\right] = \alpha A_i \cdot x^* \geq \alpha b_i. \tag{5.2}$$

Além disso, podemos supor que $\delta_i < 1$, pois do contrário o vetor s já é solução inteira do CIP, o evento E_i não ocorre e a desigualdade é trivialmente satisfeita. Assim, iremos assumir que

$$b_i - A_i \cdot s > 0.$$

Pela desigualdade de Markov,

$$\mathcal{P}\left(E_i\right) = \mathcal{P}\left((1-\delta_i)^{A_i \cdot X} \geq (1-\delta_i)^{(1-\delta_i)\mu}\right) \leq \frac{\mathbf{E}\left[(1-\delta_i)^{A_i \cdot X}\right]}{(1-\delta_i)^{(1-\delta_i)\mu_i}}.$$

Denote $\mathcal{P}\left(X_j = 1\right)$ por p_j. Como $A_{ij}, \delta_i \in [0,1]$, temos que $(1-\delta_i)^{A_{ij}} \leq 1 - A_{ij}\delta_i$ pela desigualdade de Bernoulli (Lema 1.10). Assim,

$$(1-p_j) + p_j(1-\delta_i)^{A_{ij}} = 1 + p_j((1-\delta_i)^{A_{ij}} - 1) \leq 1 + p_j(-A_{ij}\delta_i) \leq e^{-p_j A_{ij}\delta_i}.$$

Podemos então calcular

$$\mathbf{E}\left[(1-\delta_i)^{A_i \cdot X}\right] = \prod_{j=1}^{N} \mathbf{E}\left[(1-\delta_i)^{A_{ij}X_j}\right] = \prod_{j=1}^{N}\left((1-p_j) + p_j(1-\delta_i)^{A_{ij}}\right) \leq e^{-\delta_i \mu_i}.$$

Para terminar a prova do teorema, portanto, basta mostrar que

$$\left(\frac{e^{-\delta_i}}{(1-\delta_i)^{(1-\delta_i)}}\right)^{\mu_i} \leq g(B,\alpha) = \left(\alpha \cdot e^{-(\alpha-1)}\right)^{B}. \tag{5.3}$$

Defina

$$\varphi(x) = \frac{x^{b_i - A_i \cdot s}}{(b_i - A_i \cdot s)^{b_i - A_i \cdot s}} e^{b_i - x - A_i \cdot s},$$

e observe que

$$\left(\frac{e^{-\delta_i}}{(1-\delta_i)^{1-\delta_i}}\right)^{\mu_i} = \frac{\mu_i^{b_i - A \cdot s}}{(b_i - A_i \cdot s)^{b_i - A_i \cdot s}} e^{b_i - \mu_i - A_i \cdot s} = \varphi(\mu_i). \tag{5.4}$$

Note que a função $\varphi(x)$ é decrescente para $x > b_i - A_i \cdot s$, pois

$$\varphi'(x) = \frac{x^{b_i - A_i \cdot s}}{(b_i - A_i \cdot s)^{b_i - A_i \cdot s}} e^{b_i - x - A_i \cdot s}\,(b_i - A_i \cdot s - x) < 0.$$

Assim, lembrando de (5.2), temos que

$$\varphi(\mu_i) \leq \varphi(\alpha b_i - A_i \cdot s). \tag{5.5}$$

Defina então a função

$$\psi(y) = \left((1+y)^{(\alpha-1)/y}\, e^{-(\alpha-1)}\right)^{b_i}$$

e observe que

$$\varphi(\alpha b_i - A_i \cdot s) = \left(\frac{\alpha b_i - A_i \cdot s}{b_i} - A_i \cdot s\right)^{b_i - A_i \cdot s} e^{-(\alpha-1)b_i} = \psi\left(\frac{(\alpha-1)b_i}{b_i - A_i \cdot s}\right).$$

O próximo passo será mostrar que ψ é decrescente para $y > 0$, o que implicará

$$\psi\left(\frac{(\alpha-1)b_i}{b_i - A_i \cdot s}\right) \leq \psi(\alpha - 1) = \left(\alpha e^{-(\alpha-1)}\right)^{b_i}. \tag{5.6}$$

Para isso, podemos calcular

$$\ln \psi(y) = (1-\alpha)\, b_i + (\alpha - 1)\, b_i \frac{\ln(1+y)}{y},$$

e portanto, usando a Observação 5.6,

$$\frac{d}{dy}\big(\ln \psi(y)\big) = (\alpha - 1)\, b_i \left(\frac{\frac{1}{1+y} y - \ln(1+y)}{y^2} \right) < 0,$$

provando (5.6). Finalmente, seja $z = \alpha - 1$. Usando que $1 + z \leq e^z$ e que $B = \min b_i$, obtemos

$$\left(\alpha e^{-(\alpha-1)}\right)^{b_i} \leq \left(\alpha \cdot e^{-(\alpha-1)}\right)^{B}. \tag{5.7}$$

Combinando as equações (5.4), (5.5), (5.6) e (5.7), demonstramos (5.3) e concluímos a prova. □

5.1 Aproximação não construtiva

Vamos trabalhar com as aproximações vinculadas a CIPs multicritério desenvolvidas no Teorema 5.19, onde vamos observar formas mais concretas no Corolário 5.21. No desenvolvimento do Teorema 5.19 encontramos argumentos dados pelo Teorema 5.16 nas quais as ideias são generalizadas no Teorema 5.19.

Voltamos com nossas variáveis aleatórias X_j e eventos E_i com a ajuda da desigualdade FKG cada vez que os eventos da forma "$\overline{E}_i$" ou $[X_j = 1]$ são não decrescentes como uma função do vetor $(X_1, \ldots, X_n)$.

Lema 5.8. *Sejam E_i eventos decrescentes para $i = 1, \ldots, m$. Para todo B_1, $B_2 \subset [m]$ tal que $B_1 \cap B_2 = \emptyset$ e para qualquer $B_3 \subset [n]$,*

$$\mathcal{P}\left(\bigwedge_{i \in B_1} \overline{E}_i \,\middle|\, \left(\bigwedge_{j \in B_2} \overline{E}_j\right) \wedge \left(\bigwedge_{k \in B_3} (X_k = 1)\right)\right) \geq \prod_{i \in B_1} \mathcal{P}\left(\overline{E}_i\right).$$

Demonstração. Como $\overline{E}_i$ e $[X_k = 1]$ são eventos crescentes, então ambos os eventos $\bigwedge_{i \in B_1} \overline{E}_j$, $(\bigwedge_{j \in B_2} \overline{E}_j) \wedge (\bigwedge_{k \in B_3} (X_k = 1))$ são eventos crescentes. Assim,

$$\mathcal{P}\left(\bigwedge_{i \in B_1} \overline{E}_i \,\middle|\, \left(\bigwedge_{j \in B_2} \overline{E}_j\right) \wedge \left(\bigwedge_{k \in B_3} (X_k = 1)\right)\right) \geq \mathcal{P}\left(\bigwedge_{i \in B_1} \overline{E}_i\right),$$

pelo Teorema 2.19. Seja $h \in B_1$ qualquer; então, também pelo Teorema 2.19,

$$\mathcal{P}\left(\bigwedge_{i \in B_1} \overline{E}_i\right) = \mathcal{P}\left(\bigwedge_{i \in B_1 - \{h\}} \overline{E}_i \,\middle|\, \overline{E}_h\right) \cdot \mathcal{P}\left(\overline{E}_h\right) \geq \mathcal{P}\left(\bigwedge_{i \in B_1 - \{h\}} \overline{E}_i\right) \mathcal{P}\left(\overline{E}_h\right). \tag{5.8}$$

Assim, procedendo indutivamente, obtemos

$$\mathcal{P}\left(\bigwedge_{i\in B_1}\overline{E}_i\right) \geq \prod_{i\in B_1}\mathcal{P}\left(\overline{E}_i\right). \tag{5.9}$$

Por (5.8) e (5.9) temos a desigualdade. □

Definição 5.9. Para qualquer s e quaisquer $j_1,\ldots,j_s$, o conjunto $\mathcal{R}\left(j_1,\ldots,j_s\right)$ é o conjunto de índices i tal que pelo menos um dos $a_{i,j_1},\ldots,a_{i,j_s}$ é não-nulo, isto é,

$$\mathcal{R}\left(j_1,\ldots,j_s\right) = \bigcup_{h=1}^{s}\left\{i : a_{i,j_h} \neq 0\right\}$$

Observação 5.10. Observe que

$$\left|\mathcal{R}\left(j_1,\ldots,j_s\right)\right| = \left|\bigcup_{h=1}^{s}\left\{i : a_{i,j_h} \neq 0\right\}\right| \leq a\cdot s,$$

já que a é o número máximo de entradas não nulos de qualquer coluna de A.

Exemplo 5.11. Seja A uma matriz 5×5 definida por

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Alguns exemplos para $s = 2$ são

$$\mathcal{R}(1,2) = \{1,2\}$$
$$\mathcal{R}(1,3) = \{1,3,4\}$$
$$\mathcal{R}(1,4) = \{1,2,3,4\}.$$

O lema abaixo fornece uma cota superior numa probabilidade relacionada ao CIP, e será usado para obter boas soluções inteiras para CIPs (Teorema 5.16) e CIPs multicritério (Teorema 5.19); os eventos em questão são ambos crescentes, mas a desigualdade tem sentido oposto à da desigualdade FKG. Sua prova é similar à do Lema Local de Lovász.

Lema 5.12. *Seja z um vetor obtido através de arredondamento aleatório generalizado com parâmetro p, e sejam $X_1,\ldots,X_n$ as variáveis subjacentes ao arredondamento. Suponha que $\mathcal{P}\left(\bigwedge_{i=1}^{n}\overline{E}_i\right) > 0$, onde*

$$E_i = \left[(Az)_i < b_i\right].$$

Para qualquer q e quaisquer $j_1 < \cdots < j_q$, vale que

$$\mathcal{P}\left(X_{j_1} = X_{j_2} = \cdots = X_{j_q} = 1 \;\middle|\; \bigwedge_{i=1}^{m} \overline{E}_i\right) \leq \left(\prod_{t=1}^{q} p_{j_t}\right) \Big/ \prod_{i \in \mathcal{R}\left(j_1,\ldots,j_q\right)} \left(1 - \mathcal{P}\left(E_i\right)\right).$$

Além disso, se p corresponde ao arredondamento aleatório padrão, temos que

$$\prod_{i \in \mathcal{R}\left(j_1,\ldots,j_q\right)} \left(1 - \mathcal{P}\left(E_i\right)\right) \geq \left(1 - g\left(B, \alpha\right)\right)^{aq},$$

onde g é a função da Definição 5.5.

Demonstração. Sejam $Q = \mathcal{R}\left(j_1, \ldots, j_q\right)$, $Q' = [m] \setminus Q$, $Z_1 \equiv \left(\bigwedge_{i \in Q} \overline{E}_i\right)$ e $Z_2 \equiv \left(\bigwedge_{i \in Q'} \overline{E}_i\right)$. Definamos

$$Y = \prod_{t=1}^{q} X_{j_t}.$$

Por definição de $\mathcal{R}$, $a_{i,j_t} = 0$ para todo $i \in Q'$ e todo $t \in \{1, \ldots, q\}$. Assim, os $\{Xj_t\}_{t=1}^{q}$ são independentes dos eventos E_i quando $i \in Q'$, e portanto $Y = \prod_{t=1}^{q} X_{j_t}$ e $Z_2 \equiv \bigwedge_{i \in Q'} \overline{E}_i$ são independentes. Com isso, podemos cotar

$$\mathcal{P}\left(Y = 1 \mid Z_1 \wedge Z_2\right) = \frac{\mathcal{P}\left(\left(Y = 1\right) \wedge \left(Z_1 \wedge Z_2\right)\right)}{\mathcal{P}\left(Z_1 \wedge Z_2\right)} = \frac{\mathcal{P}\left(\left(Y = 1\right) \wedge Z_1 \;\middle|\; Z_2\right)}{\mathcal{P}\left(Z_1 \mid Z_2\right)} \leq \frac{\mathcal{P}\left(Y = 1 \mid Z_2\right)}{\mathcal{P}\left(Z_1 \mid Z_2\right)}.$$

Como Y e Z_2 são independentes, e Y é produto de variáveis independentes,

$$\frac{\mathcal{P}\left(Y = 1 \mid Z_2\right)}{\mathcal{P}\left(Z_1 \mid Z_2\right)} = \frac{\mathcal{P}\left(Y = 1\right)}{\mathcal{P}\left(Z_1 \mid Z_2\right)} = \frac{\prod_{t=1}^{q} \mathcal{P}\left(X_{j_t} = 1\right)}{\mathcal{P}\left(Z_1 \mid Z_2\right)}.$$

Além disso, pela desigualdade de correlação do Lema 5.8, temos que

$$\frac{\prod_{t=1}^{q} \mathcal{P}\left(X_{j_t} = 1\right)}{\mathcal{P}\left(Z_1 \mid Z_2\right)} \leq \frac{\prod_{t=1}^{q} \mathcal{P}\left(X_{j_t} = 1\right)}{\prod_{i \in Q}\left(1 - \mathcal{P}\left(E_i\right)\right)}.$$

Combinando as desigualdades obtidas, provamos a primeira parte. Para a segunda parte, temos pelo Lema 5.7 que, para todo $i \in Q$,

$$\mathcal{P}\left(E_i\right) \leq g\left(B, \alpha\right).$$

Então

$$\prod_{i \in \mathcal{R}\left(j_1,\ldots,j_q\right)} \left(1 - \mathcal{P}\left(E_i\right)\right) \geq \left(1 - g\left(B, \alpha\right)\right)^{|Q|} \geq \left(1 - g\left(B, \alpha\right)\right)^{aq},$$

usando a Observação 5.10 e o fato de que $0 < g\left(B, \alpha\right) < 1$. □

Lema 5.13. *A função g da Definição 5.5 satisfaz*

$$g(B, \alpha) \leq e^{-B(\alpha-1)^2/(2\alpha)}.$$

Demonstração. Para provar o lema, precisamos estimar $\alpha e^{-(\alpha-1)}$. Para fazê-lo, note que

$$(x \ln x)' = 1 + \ln x \leq x = \left(\frac{x^2-1}{2}\right)',$$

onde a desigualdade decorre da concavidade da função logaritmo, pois $x-1$ é a reta tangente a $\ln(x)$ em $x=1$. Podemos então integrar no intervalo $[1,\alpha]$ para obter

$$\alpha \ln \alpha \leq \frac{\alpha^2-1}{2}.$$

Assim, podemos calcular

$$\alpha e^{-(\alpha-1)} = \exp\big((\ln \alpha) - (\alpha-1)\big) \leq \exp\left(\frac{\alpha^2-1}{2\alpha} - (\alpha-1)\right) = \exp\left(-\frac{(\alpha-1)^2}{2\alpha}\right),$$

e portanto

$$g(B,\alpha) = (\alpha e^{-(\alpha-1)})^B \leq e^{-B(\alpha-1)^2/(2\alpha)} \text{ para } B \geq 0,$$

como queríamos. □

Proposição 5.14. *Sejam α e β definidos por*

$$\begin{cases} \alpha = \frac{6\ln(a+1)}{B}, \beta = 2 & \textit{se } \ln(a+1) \geq B, \\ \alpha = \beta = 1 + 6\sqrt{\frac{\ln(a+B)}{B}} & \textit{se } \ln(a+1) < B. \end{cases} \tag{5.10}$$

Então vale que $\beta\left(1 - g(B,\alpha)\right)^\alpha > 1$.

Demonstração. Se $\ln(a+1) \geq B$, se tem por hipótese que $\alpha = \frac{6\ln(a+1)}{B} \geq 6$. Iremos verificar primeiro que

$$\alpha \leq e^{\alpha/2-1}. \tag{5.11}$$

Para isso, seja f a função dada por $f(x) = x - e^{x/2-1}$. Se $x \geq 4$, então f é decrescente, pois

$$f'(x) = 1 - \frac{1}{2}e^{x/2-1} < 1 - \frac{e}{2} < 0.$$

Assim, se $x \geq 6$, $f(x) < f(6) = 6 - e^2 < 0$, de onde concluímos (5.11). Assim, relembrando a definição de g,

$$g(B,\alpha) = \left(\alpha e^{-(\alpha-1)}\right)^B \leq \left(e^{\alpha/2-1-(\alpha-1)}\right)^B = e^{-B\alpha/2} = e^{-3\ln(a+1)} = (a+1)^{-3}.$$

Consequentemente, usando a desigualdade de Bernoulli (Proposição 1.10),

$$\left(1 - g(B,\alpha)\right)^a \geq \left(1 - (a+1)^{-3}\right)^a \geq 1 - a(a+1)^{-3}. \tag{5.12}$$

Vemos então que

$$a(a+1)^{-3} \leq \frac{a}{(a+1)^2} = \frac{1}{a+1}\left(1 - \frac{1}{a+1}\right) \leq \frac{1}{4}.$$

Substituindo em (5.12), temos que

$$\beta\left(1-g\left(B,\alpha\right)\right)^{a}>\frac{3}{4}\beta>1,$$

o que prova que a escolha de α e β no primeiro caso de (5.10) satisfaz a conclusão desejada. Passemos ao caso em que $\ln\left(a+1\right)<B$; pelo Lema 5.13 e usando que $\alpha<7$ nesse caso,

$$g\left(B,\alpha\right)\leq\exp\left(\frac{-B\left(\alpha-1\right)^{2}}{2\alpha}\right)<\exp\left(-\frac{B\left(\alpha-1\right)^{2}}{14}\right).$$

Substituindo o valor de α de (5.10), obtemos que

$$\exp\left(-\frac{B\left(\alpha-1\right)^{2}}{14}\right)=\exp\left(-\frac{B}{14}\cdot\frac{36\ln\left(a+B\right)}{B}\right)=\left(e^{\ln(a+B)}\right)^{-18/7}=\left(a+B\right)^{-18/7}.$$

Assim, multiplicando por β, temos que

$$\begin{aligned}\beta\left(1-g\left(B,\alpha\right)\right)^{a}&=\left(1+6\sqrt{\ln\left(a+B\right)/B}\right)\left(1-g\left(B,\alpha\right)\right)^{a}\\&>\left(1+6\sqrt{\ln\left(a+B\right)/B}\right)\left(1-\left(a+B\right)^{-18/7}\right)^{a}\\&\geq\left(1+6\sqrt{\ln\left(a+B\right)/B}\right)\left(1-a\left(a+B\right)^{-18/7}\right).\end{aligned}\tag{5.13}$$

Queremos mostrar que o lado direito é maior que 1. Expandindo, isso é equivalente a mostrar que

$$6\sqrt{\ln\left(a+B\right)/B}\geq a\left(a+B\right)^{-18/7}+6\sqrt{\ln\left(a+B\right)/B}\cdot a\left(a+B\right)^{-18/7}.\tag{5.14}$$

Para simplificar as contas, note que, como $\ln(a+B)<1$, temos que

$$12a(a+B)^{-18/7}\geq a(a+B)^{-18/7}+6\sqrt{\ln(a+B)/B}\cdot a(a+B)^{-18/7}.$$

Assim, para mostrar (5.14), é suficiente verificar que

$$6\sqrt{\frac{\ln(a+B)}{B}}>12a(a+B)^{-18/7}.\tag{5.15}$$

Para isso, seja $x=a+B\geq 2$. Assim,

$$x^{36/7}\ln x\geq x^{4}\ln 2\geq x^{3}\cdot 2\ln 2\geq 3a^{2}B\cdot 2\ln 2\geq 4a^{2}B,$$

que é equivalente a (5.15) reordenando termos e tirando raiz quadrada. Assim, provamos (5.14) e portanto o lado direito de (5.13) é maior que 1. Em outras palavras,

$$\beta\left(1-g\left(B,\alpha\right)\right)^{a}>1,$$

como queríamos. □

Lema 5.15. *Seja x^* uma solução da relaxação linear de um CIP, e z seu arredondamento aleatório padrão com parâmetro α. Então, com probabilidade positiva $z_j>x_j^*$ ocorre para todo j.*

Demonstração. Suponha primeiro que αx_j^* não é inteiro. Então pela Definição 5.4 o arredondamento aleatório satisfaz

$$\mathcal{P}\left(z_j = \lceil \alpha x_j^* \rceil\right) = 1 - (\alpha x_j^* - \lfloor \alpha x_j^* \rfloor) > 0.$$

Se o evento acima ocorre, então $z_j \geq \alpha x_j^* > x_j^*$, como queríamos. Por outro lado, se αx_j^* é inteiro, vale deterministicamente que $z_j = \alpha x_j^* > x_j^*$, como afirmado. □

Teorema 5.16 (Teorema 2.6 de [15])**.** *Para cada CIP, existe uma solução viável de valor no máximo*

$$y^* \left(1 + O\left(\max\left(\ln{(a+1)}/B, \sqrt{\ln{(a+B)}/B}\right)\right)\right),$$

onde y^* é o valor ótimo da função objetivo para a relaxação linear.

Demonstração. Sejam α e β dadas pela Proposição 5.14. Seja x^* uma solução ótima da relaxação linear, e z obtido de x^* por arredondamento aleatório padrão com parâmetro α. Vamos demostrar que, com probabilidade positiva,

$$z^* \leq y^* \left(1 + O\left(\max\left\{\frac{\ln{(a+1)}}{B}, \sqrt{\frac{\ln{(a+B)}}{B}}\right\}\right)\right),$$

onde $z^* = c^\intercal z$ é o valor da função objetivo para z. Para tal arredondamento aleatório, sejam E_i e μ_i como no Lema 5.7. Por (5.1), temos $\mathbf{E}\left[z_v\right] = \alpha y_v^*$ para $v = 1, \ldots, n$. Definamos

$$Z \equiv \bigwedge_{i=1}^{m} \overline{E_i} \text{ e}$$

$$\mu := \mathbf{E}\left[c^\intercal z\right] = \mathbf{E}\left[\sum_v c_v z_v\right] = \sum_v c_v \mathbf{E}\left[z_v\right] = \sum_v c_v \alpha y_v^* = \alpha y^*. \tag{5.16}$$

Vamos verificar que $\mathcal{P}\left(Z\right) > 0$ quando $\alpha > 1$. Pelo Lema 5.15 temos que, com probabilidade positiva vale que $A_i z > A_i x^* > b_i$ para todo i, e portanto $\mathcal{P}\left(Z\right) > 0$ como queríamos. Podemos então calcular, pela desigualdade de Markov,

$$\mathcal{P}\left(c^\intercal z > y^* \alpha\beta \mid Z\right) \leq \frac{\mathbf{E}\left[c^\intercal z \mid Z\right]}{y^* \alpha\beta} = \frac{1}{\mu\beta} \cdot \sum_{j=1}^{n} c_j \mathbf{E}\left[z_j \mid Z\right] = \frac{1}{\mu\beta} \cdot \sum_{j=1}^{n} c_j \mathcal{P}\left(X_j = 1 \mid Z\right).$$

Para estimar o lado direito, usaremos o Lema 5.12 com $q = 1$ e $j_1 = j$. Obtemos, usando (5.16), que

$$\frac{\sum_{j=1}^{n} c_j \mathcal{P}\left(X_j = 1 \mid Z\right)}{\mu\beta} \leq \frac{\sum_{j=1}^{n} c_j \mathcal{P}\left(X_j = 1\right)}{\mu\beta\left(1 - g\left(B, \alpha\right)\right)^a} \leq \frac{1}{\beta\left(1 - g\left(B, \alpha\right)\right)^a} < 1,$$

onde a última desigualdade segue da Proposição 5.14. Assim,

$$\mathcal{P}\left(c^\intercal z \leq y^* \alpha\beta \mid Z\right) > 0,$$

como queríamos. □

5.2 CIPs multicritério

Definição 5.17 (CIP multicritério)**.** Informalmente, um CIP multicritério é uma variação do CIP em que, ao invés de buscarmos minimizar um único valor $c^\intercal z$, temos vários vetores dados não negativos $c_1, \ldots, c_\ell$ e queremos minimizar alguma função dos $c_i^\intercal \cdot x$; por exemplo, podemos ter como objetivo minimizar $\max_i c_i^\intercal z$.

Como na Definição 5.1, um CIP multicritério é definido por parâmetros $A \in [0,1]^{m \times n}$ e $b \in [1, \infty)^m$. Além disso, para todo $1 \le i \le \ell$, seja $c_i \in [0,1]^n$ um vetor com $\max_j c_{i,j} = 1$, onde $c_{i,j}$ denota a j-ésima coordenada do vetor c_i.

Definição 5.18 (As funções ch e ch′)**.** Considere o arrendondamento aleatório generalizado com algum parâmetro p, e $X_1, \ldots, X_n$ variáveis aleatórias subjacentes ao arredondamento. Para cada $i \in [m]$, sejam s e δ_i definidos como no Lema 5.7 (em particular δ_i é definido em termos de μ_i, que depende das variáveis do arredondamento aleatório *padrão*). Definimos

$$\mathrm{ch}_i\left(p\right) = \frac{\mathbf{E}\left[\left(1-\delta_i\right)^{A_i \cdot X}\right]}{\left(1-\delta_i\right)^{b_i - A_i \cdot s}} = \frac{\prod_{j \in [n]} \mathbf{E}\left[\left(1-\delta_i\right)^{A_{i,j} X_j}\right]}{\left(1-\delta_i\right)^{b_i - A_i \cdot s}}$$

e

$$\mathrm{ch}_i'\left(p\right) = \min\{\mathrm{ch}_i\left(p\right), 1\}.$$

Teorema 5.19. *Suponhamos um CIP multicritério, bem como parâmetros $\lambda_1, \ldots, \lambda_\ell > 0$. Seja z o arredondamento aleatório generalizado com parâmetro $p = (p_1, \ldots, p_n)$, e $\mathcal{A}$ dado por*

$$\mathcal{A} \equiv \left[[Az \ge b] \wedge \left(\bigwedge_{i=1}^{\ell} [c_i^\intercal z \le \lambda_i] \right) \right].$$

Então para qualquer sequência de inteiros positivos $(k_1, \ldots, k_\ell)$ tais que $k_i \le \lambda_i$ vale que

$$\mathcal{P}\left(\mathcal{A}\right) \ge \Phi(p),$$

onde Φ é dado por

$$\Phi\left(p\right) = \left(\prod_{r \in [m]} \left(1 - \mathrm{ch}_r'\left(p\right)\right) \right) - \sum_{i=1}^{\ell} \frac{1}{\binom{\lambda_i}{k_i}} \sum_{j_1 < \cdots < j_{k_i}} \left(\prod_{t=1}^{k_i} c_{i,j_t} \cdot p_{j_t} \right) \prod_{r \notin \mathcal{R}\left(j_1, \ldots, j_{k_i}\right)} \left(1 - \mathrm{ch}_r'\left(p\right)\right). \tag{5.17}$$

Demonstração. Note primeiro que

$$\Phi(p) = \left(\prod_{r \in [m]} \left(1 - \mathrm{ch}_r'\left(p\right)\right) \right) \left(1 - \sum_{i=1}^{\ell} \frac{1}{\binom{\lambda_i}{k_i}} \sum_{j_1 < \cdots < j_{k_i}} \left(\prod_{t=1}^{k_i} c_{i,j_t} p_{i,j_t} \right) \prod_{r \in \mathcal{R}\left(j_1, \ldots, j_{k_i}\right)} \left(1 - \mathrm{ch}_r'\left(p\right)\right)^{-1} \right).$$

Defina, para $1 \le i \le m$ e para $1 \le j \le \ell$,

$$E_i \equiv \left[\left(Az\right)_i < b_i\right] \qquad \text{e} \qquad \mathcal{E}_j \equiv \left[c_j^\intercal z > \lambda_j\right].$$

Pelo Lema 5.7 e pela definição de ch$'$ (Definição 5.18) temos que

$$\mathcal{P}\left(E_i\right) \leq \mathrm{ch}'_i\left(p\right).$$

Observemos que se $\mathrm{ch}'_r\left(p\right) = 1$ para algum r então $\Phi(p) = 0$, e portanto o resultado desejado é imediato. Assim, podemos supor que $\mathrm{ch}'_r\left(p\right) < 1$ para todo r.

Seja $Z \equiv \bigwedge_{r\in[m]} \overline{E}_r$. Note que

$$\mathcal{P}\left(\mathcal{A}\right) = \mathcal{P}\left(Z\right)\mathcal{P}\left(\bigwedge_i \overline{\mathcal{E}}_i \,\middle|\, Z\right) = \mathcal{P}\left(Z\right)\left(1 - \mathcal{P}\left(\bigvee_i \mathcal{E}_i \,\middle|\, Z\right)\right) \geq \mathcal{P}\left(Z\right)\left(1 - \sum_i \mathcal{P}\left(\mathcal{E}_i \mid Z\right)\right). \tag{5.18}$$

Iremos cotar $P(Z)$ e $\mathcal{P}\left(\mathcal{E}_i \mid Z\right)$. Pela desigualdade FKG temos

$$\mathcal{P}\left(Z\right) = \mathcal{P}\left(\bigwedge_{r\in[m]} \overline{E}_r\right) \leq \prod_{r=1}^{m} \mathcal{P}\left(\overline{E}_r\right) = \prod_{r\in[m]} \left(1 - \mathcal{P}\left(E_r\right)\right).$$

Além disso, pelo Teorema 2.8 parte(a),

$$\begin{aligned}
\mathcal{P}\left(\mathcal{E}_i \mid Z\right) &\leq \mathbf{E}\left[S_{k_i}\left(c_{i,1}X_1, \ldots, c_{i,n}X_n\right) \Big/ \binom{\lambda_i}{k_i} \,\middle|\, Z\right] \\
&= \sum_{j_1<\cdots<j_{k_i}} \left[c_{i,j_1}\cdots c_{i,j_{k_i}}\mathcal{P}\left(X_{j_1} = X_{j_2} = \cdots X_{j_{k_i}} = 1 \,\middle|\, \bigwedge_{i=1}^{m} \overline{E}_i\right) \Big/ \binom{\lambda_i}{k_i}\right] \\
&\leq \frac{\sum_{j_1<\cdots<j_{k_i}} \left[\left(c_{i,j_1}, \cdots c_{i,j_{k_i}} \prod_{t=1}^{k_i} p_{j_t}\right) \Big/ \binom{\lambda_i}{k_i}\right]}{\prod_{i\in\mathcal{R}\left(j_1,\ldots,j_{k_i}\right)} \left(1 - \mathcal{P}\left(E_i\right)\right)} \text{(pelo Lema 5.12 parte 1)} \\
&= \frac{1}{\binom{\lambda_i}{k_i}} \sum_{j_1<\cdots<j_{k_i}} \left(\prod_{t=1}^{k_i} c_{i,j_t}p_{i,j_t}\right)\left(\prod_{i\in\mathcal{R}\left(j_1,\ldots,j_{k_i}\right)} \left(1 - \mathcal{P}\left(E_i\right)\right)^{-1}\right).
\end{aligned}$$

Substituindo as duas desigualdades em em (5.18),

$$
\begin{aligned}
\mathcal{P}(\mathcal{A}) &\geq \left(\prod_{r\in[m]}\left(1-\mathcal{P}(E_r)\right)\right)\left(1-\sum_i \mathcal{P}(\mathcal{E}_i \mid Z)\right)\\
&\geq \left(\prod_{r\in[m]}\left(1-\mathcal{P}(E_r)\right)\right)\left(1-\sum_{i=1}^{l}\frac{1}{\binom{\lambda_i}{k_i}}\sum_{j_1<\cdots<j_{k_i}}\left(\prod_{t=1}^{k_i} c_{i,j_t}p_{i,j_t}\right)\left(\prod_{i\in\mathcal{R}\left(j_1,\ldots,j_{k_i}\right)}\left(1-\mathcal{P}(E_i)\right)^{-1}\right)\right)\\
&\geq \left(\prod_{r\in[m]}\left(1-\mathrm{ch}_r'(p)\right)\right)\left(1-\sum_{i=1}^{l}\frac{1}{\binom{\lambda_i}{k_i}}\sum_{j_1<\cdots<j_{k_i}}\left(\prod_{t=1}^{k_i} c_{i,j_t}p_{i,j_t}\right)\left(\prod_{i\in\mathcal{R}\left(j_1,\ldots,j_{k_i}\right)}\left(1-\mathrm{ch}_r'(p)\right)^{-1}\right)\right)\\
&= \Phi(p),
\end{aligned}
$$

onde a última desigualdade vale por (5.18). □

Corolário 5.20. *Suponha que valem as hipóteses do Teorema 5.19 e que p corresponde ao arredondamento aleatório padrão. Seja* $\lambda_i = \nu_i(1+\gamma_i)$ *para cada* $i \in [\ell]$ *onde* $\nu_i = \mathbf{E}\left[c_i^\intercal \cdot z\right] = c_i^\intercal \cdot \alpha x^*$ *e* $\gamma_i > 0$ *é algum parâmetro. Então*

$$
\Phi(p) \geq \left(1-g(B,\alpha)\right)^m \cdot \left(1-\sum_{i=1}^{\ell}\frac{\binom{n}{k_i}\cdot\left(\frac{\nu_i}{n}\right)^{k_i}}{\binom{\nu_i(1+\gamma_i)}{k_i}}\cdot\left(1-g(B,\alpha)\right)^{-ak_i}\right). \tag{5.19}
$$

Em particular, se o lado direito de (5.19) *é positivo, então* $\mathcal{P}(\mathcal{A}) > 0$ *para o arrendondamento aleatório padrão.*

Demonstração. Temos, pela definição de ch (Definição 5.18) e pelo Lema 5.7

$$
\mathrm{ch}_r'(p) \leq \mathrm{ch}_r(p) = \frac{\mathbf{E}\left[(1-\delta_r)^{A_r\cdot X}\right]}{(1-\delta_r)^{b_r - A_r\cdot s}} \leq g(B,\alpha).
$$

Assim $1-\mathrm{ch}_r'(p) \geq 1-g(B,\alpha)$, e portanto

$$
\begin{aligned}
\Phi(p) &= \left(\prod_{r\in[m]}\left(1-\mathrm{ch}_r'(p)\right)\right)\left(1-\sum_{i=1}^{\ell}\frac{1}{\binom{\lambda_i}{k_i}}\sum_{j_1<\cdots<j_{k_i}}\left(\prod_{t=1}^{k_i} c_{i,j_t}p_{i,j_t}\right)\left(\prod_{i\in\mathcal{R}\left(j_1,\ldots,j_{k_i}\right)}\left(1-\mathrm{ch}_r'(p)\right)^{-1}\right)\right)\\
&\geq \left(1-g(B,\alpha)\right)^m\left(1-\sum_{i=1}^{\ell}\frac{1}{\binom{\lambda_i}{k_i}}\sum_{j_1<\cdots<j_{k_i}}\left(\prod_{t=1}^{k_i} c_{i,j_t}p_{i,j_t}\right)\left(1-g(B,\alpha)\right)^{-ak_i}\right),
\end{aligned} \tag{5.20}
$$

já que $|\mathcal{R}(j_1,\ldots,j_{k_i})| \leq ak_i$ pela Observação 5.10. Como $\sum_j c_{i,j}p_{i,j} = \nu_i$, temos pelo Teorema 2.5

que

$$\sum_{j_1<\cdots<j_{k_i}} \left(\prod_{t=1}^{k_i} c_{i,j_t} p_{i,j_t}\right) \leq \binom{n}{k_i} \left(\frac{\nu_i}{n}\right)^{k_i}.$$

Logo, de (5.20) se tem que

$$\Phi(p) \geq \big(1 - g(B,\alpha)\big)^m \left(1 - \sum_{i=1}^{\ell} \frac{1}{\binom{\lambda_i}{k_i}} \binom{n}{k_i} \left(\frac{\nu_i}{n}\right)^{k_i} \big(1 - g(B,\alpha)\big)^{-ak_i}\right),$$

como afirmado. □

Corolário 5.21. *Existe uma constante $K' > 0$ tal que o seguinte é válido. Suponha que dado um CIP multicritério com a notação como na parte 2 do Teorema 5.19. Defina*

$$\alpha = K' \max\left\{\frac{\ln(\alpha) + \ln\ln(2\ell)}{B}, 1\right\}.$$

Se $\nu_i \geq \log^2(2\ell)$ para $i = 1, \ldots, \ell$, então o arrendondamento aleatório padrão produz uma solução viável z tal que $c_i^{\intercal} \cdot z \leq 3\nu_i$, $\forall\ i$, com probabilidade positiva.

Demonstração. Para aplicar o Teorema 5.19, definimos $\gamma_i = 2$ e $k_i = \lceil \ln 2\ell \rceil$. Obtemos então que

$$\Phi(p) \geq \big(1 - g(B,\alpha)\big)^m \left(1 - \sum_{i=1}^{\ell} \frac{1}{\binom{3\nu_i}{k_i}} \binom{n}{k_i} \left(\frac{\nu_i}{n}\right)^{k_i} \big(1 - g(B,\alpha)\big)^{-ak_i}\right).$$

Basta então mostrar que o lado direito é positivo, isto é, que

$$\sum_{i=1}^{\ell} \frac{1}{\binom{3\nu_i}{k_i}} \binom{n}{k_i} \left(\frac{\nu_i}{n}\right)^{k_i} \big(1 - g(B,\alpha)\big)^{-ak_i} < 1.$$

Iremos mostrar que cada parcela do somatório é menor que $1/\ell$. Usando que $\binom{n}{k_i} \leq n^{k_i}/k_i!$, basta mostrarmos que

$$\frac{\nu_i^{k_i}/k_i!}{\binom{3\nu_i}{k_i}} \cdot \big(1 - g(B,\alpha)\big)^{-a\cdot k_i} < \frac{1}{\ell}, \tag{5.21}$$

onde $k_i = \lceil \ln 2\ell \rceil$. Assim, vamos verificar duas observações para aproveitar a demostração da equação (5.21).

a. Como $k_i \sim \ln(\ell)$ e $\nu_i \geq \log^2(2\ell)$ então

$$\binom{3\nu_i}{k_i} = \frac{1}{k_i!} \prod_{j=0}^{k_i-1} (3\nu_i - j) = \frac{1}{k_i!} (3\nu_i)^{k_i} \cdot e^{\Theta\left(\sum_{j=0}^{k_i-1} j/3\nu_i\right)} = \Theta\left(\big(1/k_i!\big)\,(3\nu_i)^{k_i}\right)$$

b. Vamos demostrar que $\big(1 - g(B,\alpha)\big)^{-a\cdot k_i} \sim 1$.

$$\big(1 - g(B,\alpha)\big)^{-a\cdot k_i} = \left(1 - \left(\alpha e^{-(\alpha-1)}\right)^B\right)^{-a\cdot k_i} = \left(1 - \alpha^B e^{-(\alpha-1)B}\right)^{-a\cdot k_i}.$$

Para um K' suficientemente grande temos o procurado.

Então com as observações a e b temos que para cada i

$$\frac{\nu_i/k_i!}{\binom{3\nu_i}{k_i}} \cdot \big(1 - g(B,\alpha)\big)^{-a\cdot k_i} = \frac{\nu_i/k_i!}{\Theta\Big(\big(1/k_i!\big)\,(3\nu_i)^{k_i}\Big)} \cdot 1 < \frac{1}{\ell},$$

assim temos a equação (5.21). Agora temos que

$$\sum_{i=1}^{\ell} \frac{\binom{n}{k_i} \cdot \big(\nu_i/n\big)^{k_i}}{\binom{3\nu_i}{k_i}} \cdot \big(1 - g\,(B,\alpha)\big)^{-a\cdot k_i} < \sum_{i=1}^{\ell} \frac{1}{\ell} = 1,$$

logo $\mathcal{P}\,(\mathcal{A}) > 0$ como queríamos. □

5.3 Versão construtiva

Podemos mostrar, para muitos problemas, que o arredondamento aleatório produz as soluções garantidas pelos Teoremas 5.16 e 5.19 com probabilidade exponencialmente baixa. Portanto, é interessante considerar o problema de obter algoritmos randomizados que rodem em tempo polinomial (ou próximo disso) e que encontrem as soluções dadas pelos teoremas da seção anterior.

Nossa estratégia será a seguinte: Dado um vetor de probabilidades $p = (p_1, p_2, \ldots, p_n)$ tal que $\Phi(p) > 0$, iremos considerar os vetores $p' = (p_1, p_2, \ldots, p_{n-1}, 0)$ e $p'' = (p_1, p_2, \ldots, p_{n-1}, 1)$. Se conseguirmos determinar que $\Phi(p') > 0$ ou $\Phi(p'') > 0$, então podemos alterar o valor de p_n e repetir o processo com as demais variáveis. Quando todos os p_is estiverem em $\{0, 1\}$, temos uma solução para o problema dada por $x_i = p_i$.

Definição 5.22. Dado um vetor $p \in [0,1]^n$, denotaremos por p' e p'' os vetores

$$p' = (p_1, p_2, \ldots, p_{n-1}, 0) \text{ e } p'' = (p_1, p_2, \ldots, p_{n-1}, 1).$$

Além disso, definimos vetores q, q' e q'' de modo que, para cada $1 \le i \le m$,

$$q_i = \mathrm{ch}'(p), \qquad q_i' = \mathrm{ch}_i'(p'), \qquad q_i'' = \mathrm{ch}_i'(p'').$$

Apresentamos agora um lema útil sobre esses vetores.

Lema 5.23. *Para todo $p \in [0,1]^n$, os vetores q, q' e q'' da Definição 5.22 satisfazem as seguintes propriedades para todo $i \in [m]$:*

(a) $0 \le q_i'' \le q_i' \le 1$.

(b) $q_i \ge p_n q_i'' + (1 - p_n)\, q_i'$,

(c) $q_i' = q_i'' = q_i$ *se* $i \notin \mathcal{R}(n)$.

Demonstração. Sejam $X_1, \ldots, X_n$ as variáveis aleatórias subjacentes ao arredondamento aleatório generalizado com parâmetro p (isto é, $\mathcal{P}(X_i = 1) = p_i$). Temos pela Definição 5.18 que

$$\begin{aligned}
\mathrm{ch}_i(p'') &= \frac{(1-\delta_i)^{A_{i,n}\cdot 1}\prod_{j\in[n-1]}\mathbf{E}\left[(1-\delta_i)^{A_{i,j}X_j}\right]}{(1-\delta_i)^{b_i - A_i\cdot s}} \\
&\leq \frac{(1-\delta_i)^{A_{i,n}\cdot 0}\prod_{j\in[n-1]}\mathbf{E}\left[(1-\delta_i)^{A_{i,j}X_j}\right]}{(1-\delta_i)^{b_i - A_i\cdot s}} = \mathrm{ch}_i(p'),
\end{aligned}$$

assim $\mathrm{ch}_i(p'') \leq \mathrm{ch}_i(p')$. Pela definição de q' e q'', tomando mínimos temos que $0 \leq q_i'' \leq q_i' \leq 1$, provando (a). Para demonstrar (b), dividiremos em casos. Se $q_i = 1$, então

$$q_i = 1 = p_n + (1-p_n) \geq p_n q_i'' + (1-p_n)\, q_i'.$$

Por outro lado, se $q_i < 1$, então

$$\begin{aligned}
q_i = \mathrm{ch}_i(p) &= \frac{\prod_{j\in[n]}\mathbf{E}\left[(1-\delta_i)^{A_{i,j}X_j}\right]}{(1-\delta_i)^{b_i - A_i\cdot s}} \\
&= \left(p_n(1-\delta_i)^{A_{i,n}\cdot 1} + (1-p_n)(1-\delta_i)^{A_{i,n}\cdot 0}\right)\cdot \frac{\prod_{j\in[n-1]}\mathbf{E}\left[(1-\delta_i)^{A_{i,j}X_j}\right]}{(1-\delta_i)^{b_i - A_i\cdot s}} \\
&= p_n\,\mathrm{ch}_i(p'') + (1-p_n)\,\mathrm{ch}_i(p') \\
&\geq p_n\,\mathrm{ch}_i'(p'') + (1-p_n)\,\mathrm{ch}_i'(p') \\
&= p_n q_i'' + (1-p_n) q_i',
\end{aligned}$$

provando (b). Para provar (c), note que, como p, p' e p'' são todos iguais nas primeiras $n-1$ coordenadas, então apenas o fator $j = n$ do produtório muda quando trocamos p por p' ou p'' em

$$ch_i(p) = \frac{\prod_{j\in[n]}\mathbf{E}\left[(1-\delta_i)^{A_{i,j}\cdot X_j}\right]}{(1-\delta_i)^{b_i - A_i\cdot s}}.$$

Se $i \notin \mathcal{R}(n)$, então $A_{i,n} = 0$ e portanto os expoentes de tal fator em $\mathrm{ch}_i(p)$, $\mathrm{ch}_i(p')$ e $\mathrm{ch}_i(p'')$ são todos iguais a zero, provando o item (c). $\square$

Uma manipulação algébrica permite reescrever a condição (b) do Lema 5.23 dos seguintes dois modos equivalentes:

$$q_i' - q_i \leq p_n(q_i' - q_i'') \tag{5.22}$$

$$q_i - q_i'' \geq (1-p_n)(q_i' - q_i''). \tag{5.23}$$

Definição 5.24. Dados um vetor $r = (r_1, r_2, \ldots, r_m)$ e um conjunto $U \subset [m]$, definimos a função

$$f(U, r) = \prod_{i\in U}(1 - r_i).$$

Em particular, $f(U, r) = 1$ se $U = \emptyset$.

Lema 5.25. *Dado $p \in [0,1]^n$, sejam q, q' e q'' como na Definição 5.22. Defina*

$$\Delta(U) = (1 - p_n) \cdot f(U, q') + p_n \cdot f(U, q'') - f(U, q) \tag{5.24}$$

Se $U \subset V \subset [m]$, então

(a) $\Delta(U) \geq 0$.

(b) $\Delta(U)/f(U,q) \leq \Delta(V)/f(V,q)$.

Para melhor motivar a definição e o lema acima, iremos primeiro usar o Lema 5.25 para provar o teorema principal dessa seção, enunciado abaixo. Ele implica que se $\Phi(p) > 0$, então ou $\Phi(p') > 0$ ou $\Phi(p'') > 0$, como queríamos.

Teorema 5.26. *Seja $p \in [0,1]^n$ um vetor de probabilidades, Φ dado por* (5.17) *e p', p'' dados por* (5.22)*. Se $\Phi(p) > 0$, então*

$$\Phi(p) \leq (1 - p_n)\, \Phi\left(p'\right) + p_n \Phi\left(p''\right). \tag{5.25}$$

Demonstração. Considerando que $p_1, \ldots, p_{n-1}$ são constantes, podemos agrupar termos em (5.17) para afirmar que existem constantes não-negativas $u_1, \ldots, u_t$ e $v_1, \ldots, v_{t'}$ e subconjuntos $U_1, \ldots, U_t$ e $V_1, \ldots, V_{t'}$ de $[m]$ tais que

$$\Phi\,(p) = f\left([m], q\right) - \left(\sum_i u_i \cdot f(U_i, q)\right) - \left(p_n \cdot \sum_j v_j \cdot f(V_j, q)\right), \tag{5.26}$$

$$\Phi\left(p'\right) = f\left([m], q'\right) - \sum_i u_i \cdot f\left(U_i, q'\right), \tag{5.27}$$

$$\Phi\left(p''\right) = f\left([m], q''\right) - \left(\sum_i u_i \cdot f\left(U_i, q''\right)\right) - \left(\sum_j v_j \cdot f\left(V_j, q''\right)\right). \tag{5.28}$$

com a propriedade de que

$$V_j \cap \mathcal{R}(n) = \emptyset. \tag{5.29}$$

Queremos provar que

$$(1 - p_n)\, \Phi\left(p'\right) + p_n \Phi\left(p''\right) - \Phi(p) \geq 0. \tag{5.30}$$

Substituindo (5.26), (5.27) e (5.28) em (5.30), os termos multiplicados por v_j se cancelam porque $q_i = q''_i$ se $i \in V_j$ (por (5.29) e item (c) do Lema 5.23). Assim, lembrando da definição de Δ no Lema 5.25, concluímos que podemos reescrever (5.30) como

$$\Delta\left([m]\right) - \sum_i u_i \cdot \Delta\left(U_i\right) \geq 0. \tag{5.31}$$

Pelo Lema 5.25, temos que, para todo U_i, $\Delta(U_i)/f(U_i,q) \leq \Delta([m])/f([m],q)$. Assim,

$$\begin{aligned}\Delta\big([m]\big) - \sum_i u_i \cdot \Delta(U_i) &= \frac{\Delta\big([m]\big)}{f([m],q)} \cdot f\big([m],q\big) - \sum_i \left(u_i \cdot \frac{\Delta(U_i)}{f(U_i,q)} \cdot f(U_i,q)\right) \\ &\geq \frac{\Delta\big([m]\big)}{f\big([m],q\big)} \cdot \left[f\big([m],q\big) - \sum_i u_i \cdot f(U_i,q)\right].\end{aligned}$$

O lado direito é produto de dois fatores. Para estimar o primeiro fator, tomamos $U = \emptyset$ e $V = [m]$ no Lema 5.25 para obter

$$\frac{\Delta([m])}{f([m],q)} \geq \frac{\Delta(\emptyset)}{f(\emptyset,q)} = 0.$$

Para o segundo fator, notamos que por (5.26) e por hipótese, temos que

$$f\big([m],q\big) - \sum_i u_i \cdot f(U_i,q) \geq \Phi(p) > 0.$$

Assim, mostramos (5.31), o que conclui a prova do teorema. □

Resta então provar o Lema 5.25.

Prova do Lema 5.25. Faremos uma prova de (a) por indução na cardinalidade de U. Se $|U| = 0$ temos

$$\Delta(U) = (1-p_n) + p_n - 1 = 0.$$

Suponhamos que, para algum $s < m$, o item (a) valha para todos os conjuntos de cardinalidade s. Seja $w \notin U$. Iremos provar que

$$\Delta\big(U \cup \{w\}\big) \geq (1-q_w)\Delta(U). \tag{5.32}$$

Para isso, usamos as desigualdades (5.22) e (5.23) para calcular

$$\begin{aligned}\Delta\big(U \cup \{w\}\big) - (1-q_w) \cdot \Delta(U) &= (1-p_n)(q_w - q'_w)f(U,q') + p_n(q_w - q''_w)f(U,q'') \\ &\geq -(1-p_n)p_n(q'_w - q''_w)f(U,q') + p_n(1-p_n)(q'_w - q''_w)f(U,q'') \\ &= p_n(1-p_n)(q'_w - q''_w)(f(U,q'') - f(U,q')) \geq 0,\end{aligned}$$

onde a última desigualdade vem do Lema 5.23(a) e da definição de f. Como $\Delta(U) \geq 0$ por hipótese de indução, então $\Delta(U \cup \{w\}) \geq 0$, provando que (a) vale para todo conjunto de tamanho $s+1$. Para provar (b), note que se $w \notin U$, então

$$\frac{f(U \cup \{w\}, q)}{f(U,q)} = 1 - q_w.$$

Substituindo em (5.32), concluímos que para todo U e todo $w \notin U$,

$$\frac{\Delta(U,q)}{f(U,q)} \leq \frac{\Delta(U \cup \{w\})}{f(U \cup \{w\}, q)}.$$

Se $V \setminus U = \{w_1, \ldots, w_k\}$, podemos iterar o procedimento e obter que

$$\frac{\Delta(U,q)}{f(U,q)} \leq \frac{\Delta(U \cup \{w_1\})}{f(U \cup \{w_1\}, q)} \leq \cdots \leq \frac{\Delta(U \cup \{w_1, \ldots, w_k\})}{f(U \cup \{w_1, \ldots, w_k\}, q)} = \frac{\Delta(V)}{f(V,q)},$$

provando o item (b). $\square$

Conclusão

Estimamos a aplicação do Lema Local de Lovász e suas extensões em problemas de Programação Inteira, desde a teoria até a implementação algorítmica. Resultados significativos relacionados aos Problemas Minimax e de Cobertura são destacados, ilustrando a utilidade de abordagens não construtivas e multifatoriais no campo da otimização combinatória. A análise detalhada de casos específicos, como a cobertura de um hipergrafo por arestas, sublinha a versatilidade e aplicabilidade dos métodos estudados. Além disso, fica evidente a importância dos conceitos de eventos monotônicos e estimativas para uma compreensão profunda das técnicas discutidas nos capítulos, consolidando assim a relevância das estratégias probabilísticas na resolução eficiente de problemas complexos de Programação Inteira.

Referências Bibliográficas

[1] N. Alon and J. H. Spencer. *The Probabilistic Method.* Wiley Publishing, 4th edition, 2016.

[2] M. H. d. Carvalho, M. R. Cerioli, R. Dahab, P. Feofiloff, C. G. Fernandes, C. E. Ferreira, K. S. Guimarães, F. K. Miyazawa, J. C. d. Pina Jr, J. A. R. Soares, and Y. Wakabayashi. *Uma introdução sucinta a algoritmos de aproximação*. IMPA, 2001.

[3] S. A. Cook. The complexity of theorem-proving procedures. In *Proceedings of the third annual ACM symposium on Theory of computing*, pages 151–158. ACM, 1971.

[4] P. Erdős and L. Lovász. Problems and results on 3-chromatic hypergraphs and some related questions. In *Colloquia Mathematica Societatis János Bolyai 10. Infinite and Finite Sets, Keszthely (Hungria)*, 1973.

[5] R. L. Graham. Bounds for certain multiprocessing anomalies. *Bell System Technical Journal*, 45(9):1563–1581, 1966.

[6] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[7] R. M. Karp. Reducibility among combinatorial problems. In *Complexity of computer computations*, pages 85–103. Springer, 1972.

[8] L. A. Levin. Universal sequential search problems. *Problemy peredachi informatsii*, 9(3):115–116, 1973.

[9] J. Matousek and B. Gärtner. *Understanding and using linear programming.* Springer Science & Business Media, 2007.

[10] P. Mukhopadhyay. *An introduction to the theory of probability.* World Scientific, 2012.

[11] R. B. Proença. *Técnicas para convergência da Expansão do Gás de Polímeros e uma aplicação ao Método Probabilístico.* PhD thesis, Universidade Federal de Minas Gerais, 2009.

[12] J. P. Schmidt, A. Siegel, and A. Srinivasan. Chernoff–hoeffding bounds for applications with limited independence. *SIAM Journal on Discrete Mathematics*, 8(2):223–250, 1995.

[13] A. Schrijver. *Combinatorial optimization: polyhedra and efficiency*, volume 24. Springer Science & Business Media, 2003.

[14] M. Sipser. *Introduction to the Theory of Computation.* Course Technology, third edition, 2013.

[15] A. Srinivasan. An extension of the Lovász local lemma, and its applications to integer programming. *SIAM Journal on Computing*, 36(3):609–634, 2006.

[16] Wikipedia contributors. Lovász local lemma — Wikipedia, the free encyclopedia, 2019. [Online; accessed 26-November-2019].