



UNIVERSIDADE FEDERAL DO RECÔNCAVO DA BAHIA
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
MESTRADO PROFISSIONAL EM MATEMÁTICA EM
REDE NACIONAL



JOILSON DE LIMA SILVA

O Problema de Transporte sob as Perspectivas da Programação Linear e da Teoria do Transporte Ótimo

CRUZ DAS ALMAS - BAHIA

Fevereiro de 2022

JOILSON DE LIMA SILVA

O Problema de Transporte sob as Perspectivas da Programação Linear e da Teoria do Transporte Ótimo

Dissertação apresentada ao Mestrado Profissional em Matemática em Rede Nacional do Centro de Ciências Exatas e Tecnológicas da Universidade Federal do Recôncavo da Bahia como parte dos requisitos para a obtenção do título de Mestre em Matemática.

Orientador: Prof. Dr. João Tiago Assunção Gomes

CRUZ DAS ALMAS - BAHIA

Fevereiro de 2022

FICHA CATALOGRÁFICA

S586p	<p>Silva, Joilson de Lima. O problema de transporte sob as perspectivas da programação linear e da teoria do transporte ótimo / Joilson de Lima Silva. _ Cruz das Almas, Bahia, 2022. 100f.; il.</p> <p>Dissertação (Mestrado) – Universidade Federal do Recôncavo da Bahia, Centro de Ciências Exatas e Tecnológicas, Mestrado Profissional em Matemática – PROFMAT.</p> <p>Orientador: Prof. Dr. João Tiago Assunção Gomes.</p> <p>1. Matemática – Estudo e ensino. 2. Modelagem matemática – Programação linear. 3. Transportes – Análise. I. Universidade Federal do Recôncavo da Bahia, Centro de Ciências Exatas e Tecnológicas. II. Título.</p> <p>CDD: 510.7</p>
-------	--

Ficha elaborada pela Biblioteca Central de Cruz das Almas - UFRB.
Responsável pela Elaboração - Antonio Marcos Sarmento das Chagas (Bibliotecário - CRB5 / 1615).
(os dados para catalogação foram enviados pelo usuário via formulário eletrônico).

JOILSON DE LIMA SILVA

O Problema de Transporte sob as Perspectivas da Programação Linear e da Teoria do Transporte Ótimo

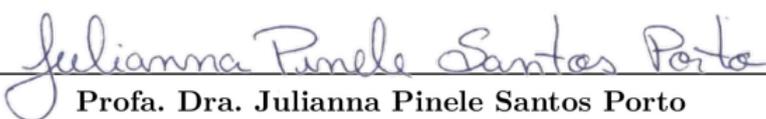
Dissertação apresentada ao Mestrado Profissional em Matemática em Rede Nacional do Centro de Ciências Exatas e Tecnológicas da Universidade Federal do Recôncavo da Bahia como parte dos requisitos para a obtenção do título de Mestre em Matemática.

Aprovado em: 04 de Fevereiro de 2022

BANCA EXAMINADORA:



Prof. Dr. João Tiago Assunção Gomes
Orientador



Profa. Dra. Julianna Pinele Santos Porto
Examinadora - UFRB



Prof. Dr. Cleber Fernando Colle
Examinador - UERJ

CRUZ DAS ALMAS - BAHIA

Fevereiro de 2022

Dedico esse trabalho, em especial, à memória de nosso saudoso amigo e meu parceiro de jornada, Luís Mário, que ao transcender desta vida carregou consigo a grandiosidade de um cidadão do bem; dedico também, a todos que me acompanharam nessa trajetória.

Agradecimentos

Agradeço primeiramente a Deus, por ter me concebido a dádiva da vida e saúde para trilhar esse caminho em busca dos meus objetivos.

À minha esposa Daiane Machado, aos meus filhos Jônathas Kelvin, Arthur Henrique e Deivid Willian, aos meus pais Lelinha e José, aos meus irmãos Jean, Jocileide e Jamilson; e a todos os outros familiares e amigos que sempre me apoiaram e me deram forças para seguir em frente nessa jornada.

Aos meus grandiosos professores por transmitirem seus conhecimentos e lições de vida, em especial ao meu orientador, professor Dr. João Tiago Assunção Gomes, por ter aceitado a missão de me guiar nesse momento final de minha jornada e ter conduzido esta função com muito amor e dedicação de sua grande maestria.

Aos Professores Dra. Julianna Pinele Santos Porto e Dr. Cleber Fernando Colle, membros da Banca Examinadora, por aceitarem o convite para avaliar e contribuírem com este trabalho.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela grandiosíssima ajuda financeira que tanto contribuiu em diversas etapas ao longo do curso.

Aos meus colegas de turma, especialmente os que se tornaram grandes amigos: Renata Sande e meu companheiro de trabalho Franquilande Aragão, vocês certamente deixaram essa longa caminhada mais suave, significativa e divertida.

“A Matemática é o alfabeto com o qual Deus escreveu o Universo.”

(Galileu Galilei)

Resumo

O objetivo deste trabalho é analisar os modelos matemáticos para o problema de transporte, a partir de duas teorias distintas, como metodologia capaz de trazer situações reais para o fazer matemática na sala de aula. A Programação Linear será o primeiro ponto de vista para analisar certos problemas de transporte associados à distribuição de produtos. Outra modelagem para tal problema, envolvendo probabilidades, será estudada pela Teoria de Transporte Ótimo, a qual permitirá caracterizar noções geométricas sobre tais elementos. Por fim, destacamos algumas aplicações do problema de transporte à comparação de histogramas associados à propagação da COVID-19 no Brasil com o intuito de reconhecer o fenômeno de “onda” epidemiológica.

Palavras-chave: COVID-19. Distância. Método *Simplex*. Modelagem Matemática. Pesquisa Operacional. Probabilidade. Problema de Transporte. Programação Linear. Teoria do Transporte Ótimo.

Abstract

The objective of this work is to analyze the mathematical models for the transport problem, from two different theories, as a methodology capable of bringing real situations to do mathematics in the classroom. Linear Programming will be the first point of view in order to analyze certain transport problems associated with product distribution. Another model for this problem, involving probabilities, will be studied by the Optimal Transport Theory, which will allow us to characterize geometrical notions about such elements. Finally, we highlight some applications of the transport problem to the comparison of histograms associated with the spread of the COVID-19 in Brazil in order to recognize the epidemiological “wave” phenomenon.

Keywords: COVID-19. Distance. Simplex Method. Mathematical Modeling. Operational Research. Probability. Transport Problem. Linear Programming. Optimal Transport Theory.

Lista de ilustrações

Figura 1 – Incluindo Dados no <i>Excel</i>	28
Figura 2 – Alimentando Informações no <i>Solver</i> I	28
Figura 3 – Alimentando Informações no <i>Solver</i> II	29
Figura 4 – Resolvendo com o <i>Solver</i>	30
Figura 5 – Solução do <i>Solver</i>	31
Figura 6 – Histograma do Lançamento de um Dado Honesto p	44
Figura 7 – Histograma do Lançamento de um Dado Viciado q	45
Figura 8 – Histograma do Lançamento de um Dado Viciado r	45
Figura 9 – Histograma Associado ao Vetor a	62
Figura 10 – Histograma Associado ao Vetor b	62
Figura 11 – Histograma Associado ao Vetor c	63
Figura 12 – Histograma Associado ao Vetor a	65
Figura 13 – Histograma Associado ao Vetor b	65
Figura 14 – Histograma dos Números de Casos da COVID-19	69
Figura 15 – Histograma dos Números de Óbitos da COVID-19	70
Figura 16 – Gráfico da Função Cosseno	71
Figura 17 – Histograma da Função Cosseno	72
Figura 18 – Gráfico da Função Cúspide Cúbica	73
Figura 19 – Histograma da Função Cúspide Cúbica	74
Figura 20 – Recorte da Planilha de Preparação para o <i>Solver</i>	77

Lista de quadros

Quadro 1 – Modelo Matemático do Exemplo Motivador	27
Quadro 2 – Formulação Teórica	32
Quadro 3 – Modelo Algébrico	33
Quadro 4 – Problema de Monge	41
Quadro 5 – Problema de Kantorovich	50
Quadro 6 – Problema de Kantorovich em Termos da Programação Linear	54
Quadro 7 – Problema Primal	94
Quadro 8 – Problema Dual	94

Lista de tabelas

Tabela 1 – Custos de Transporte	25
Tabela 2 – Números de Casos da COVID-19	69
Tabela 3 – Números de Óbitos da COVID-19	70
Tabela 4 – Valores para Função Cosseno	72
Tabela 5 – Valores para Função Cúspide Cúbica	73
Tabela 6 – Divisão Intervalos dos Histogramas para Cálculo de Distâncias	75
Tabela 7 – Distâncias de Wasserstein para o Histogramas de Casos da COVID-19	79
Tabela 8 – Distâncias Euclidianas para o Histogramas de Casos da COVID-19	79
Tabela 9 – Distâncias Wasserstein para o Histogramas de Óbitos da COVID-19	82
Tabela 10 – Distâncias Euclidianas para o Histogramas de Óbitos da COVID-19	82

Sumário

	INTRODUÇÃO	13
1	MODELAGEM MATEMÁTICA	17
1.1	Breve Histórico	18
2	PROGRAMAÇÃO LINEAR	21
2.1	Modelagem de uma Situação-Problema	21
2.2	Exemplo Motivador	24
2.3	Formulação Teórica	31
2.4	Método <i>Simplex</i>	33
3	TEORIA DO TRANSPORTE ÓTIMO	40
3.1	O Problema de Monge	40
3.2	O Problema de Kantorovich	42
3.3	Elo entre a Teoria de Transporte Ótimo e a Programação Linear	51
3.4	Distância de Wasserstein	54
4	APLICAÇÕES	67
4.1	Comparação de Histogramas da COVID-19	67
4.2	Outras Aplicações	83
5	CONCLUSÃO	86
	REFERÊNCIAS	88
	APÊNDICES	91
	APÊNDICE A – DUALIDADE	92
	APÊNDICE B – VETORES DE PROBABILIDADE DA SEÇÃO 4.1	95

Introdução

A modelagem matemática tem se firmado no contexto da educação matemática como sendo uma metodologia capaz de trazer situações reais para o fazer matemática na sala de aula. Este fato tem transformado esta ciência em um instrumento com o qual cada aluno se torne capaz de interpretar matematicamente a realidade que lhes rodeiam, permitindo desenvolver habilidades que lhes possibilitem melhores intervenções no mundo em que vivem.

Através da modelagem somos capazes de melhorar a motivação dos alunos, de forma que eles possam desenvolver atitudes críticas perante a sociedade, despertem o senso criativo e sejam impulsionados para utilizar novas estratégias, mesmo que estas sejam informais. Além disso, quando o ensino-aprendizagem ocorre vinculado a um contexto próximo da realidade do aluno, a disciplina estudada ganha uma nova conotação como matéria de aprendizagem.

A resolução de problemas via modelos matemáticos se apresenta como uma metodologia de ensino de Matemática que pode ser é muito eficaz, tendo em vista que exige a mobilização de diversos saberes até que se construa uma solução satisfatória.

Um professor de matemática tem, assim, uma grande oportunidade. Se ele preenche o tempo que lhe é concedido a exercitar seus alunos em operações rotineiras, aniquila o interesse e tolhe o desenvolvimento intelectual dos estudantes, desperdiçando, dessa maneira, a sua oportunidade. Mas se ele desafia a curiosidade dos alunos, apresentando-lhes problemas compatíveis com os conhecimentos destes e auxiliando-os por meio de indagações estimulantes, poderá incutir-lhes o gosto pelo raciocínio independente e proporcionar-lhes certos meios para alcançar este objetivo (POLYA, 1978).

Durante a resolução de problemas, é muito importante valorizar o processo escolhido pelo aluno para tentar chegar a uma resposta e, dessa forma, podemos abrir espaço para o surgimento de diferentes resoluções. Além disso, deve-se estabelecer a comparação entre tais soluções e oportunizar aos estudantes o compartilhamento dos caminhos seguidos até à solução final do problema.

É fundamental que a resolução de problemas seja baseada na apresentação de situações abertas e sugestivas, que sejam capazes de extrair dos alunos participação que desencadeia uma atitude ativa, combinada com um esforço efetivo na busca de suas próprias respostas para culminar na produção de seu próprio conhecimento.

Problema de Transporte

Bons exemplos usados no ensino de matemática do Ensino Médio e do Ensino Técnico, para auxiliar no desenvolvimento cognitivo, visando melhorar o processo de tomada de decisão dentro de uma perspectiva mais próxima da realidade de cada aluno, são encontrados nas áreas de logística e de transporte.

[...]Transportes. Para a maioria das firmas, o transporte é a atividade logística mais importante simplesmente porque ela absorve, em média, de um a dois terços dos custos logísticos. É essencial, pois nenhuma firma moderna pode operar sem providenciar a movimentação de suas matérias-primas ou de seus produtos acabados de alguma forma. Sua importância é sempre sublinhada pelos problemas financeiros colocados para muitas empresas quando há uma greve ferroviária nacional ou quando carreteiros autônomos paralisam suas atividades devido a aumentos de combustíveis. Não é incomum denominar tais eventos de desastres nacionais. Os mercados não podem ser atendidos e produtos permanecem no canal de distribuição para deteriorarem-se ou tornarem-se obsoletos. “Transporte” refere-se aos vários métodos para se movimentar produtos. Algumas das alternativas populares são os modos rodoviário, ferroviário e aeroviário. A administração da atividade de transporte geralmente envolve decidir-se quanto ao método de transporte, aos roteiros e à utilização da capacidade dos veículos. (BALLOU, 1993, p. 24)

O setor de transporte tem uma grande importância dentro dos diversos empreendimentos que necessitam fazer o deslocamento de seus produtos, uma vez que os custos do trinômio oferta-demanda-transporte incidem significativamente sobre o valor final dos produtos. Trata-se de um estudo direcionado à alocação de recursos, que tem por objetivo principal encontrar os melhores valores para uma função, que representa os menores custos envolvidos nas possibilidades de transporte desses recursos. Por isso, são importantes o uso de técnicas na otimização do transporte para possibilitar maior eficiência na tomada de decisão, durante a elaboração do plano de escoamento de tais produtos.

Neste trabalho, interpretaremos o problema de transporte sob o ponto de vista algébrico e algorítmico da Programação Linear e sob uma perspectiva probabilística e geométrica da Teoria do Transporte Ótimo.

Programação Linear

A Programação Linear é um ramo da Pesquisa Operacional que se dedica na investigação de modelos matemáticos para representar diversos problemas do cotidiano. De tal forma que a palavra “Programação” não se refere à programação de computadores, e sim, a um planejamento que deve ser estruturado em torno de um problema, enquanto que o adjetivo “Linear” significa que todos os modelos encontrados serão representados por funções lineares.

Segundo [Paiva \(2008\)](#), Programação Linear é uma técnica da matemática aplicada que constitui um dos ramos da investigação operacional. Envolve a pesquisa, o estudo e desenvolvimento de modelos de otimização sobre operações, sendo aplicada a problemas que envolvem a condução e coordenação de operações. Dentro do universo dos modelos de otimização, existem os modelos de programação matemática que permitem determinar em que condição é possível, ou não, otimizar um dado objetivo face a um conjunto de limitações.

De acordo com as ideias de [Passos \(2008\)](#), Programação Linear é uma técnica de otimização aplicada a sistemas de equações (ou inequações) lineares representativos de modelos previamente elaborados.

Teoria do Transporte Ótimo

Segundo ([PEYRÉ; CUTURI et al., 2019](#)), a Teoria do Transporte Ótimo está interessada no problema de como comparar duas distribuições de probabilidades - duas pilhas diferentes de areia de mesmo volume. São consideradas todas as maneiras possíveis de transformar, transportar ou remodelar a primeira pilha na segunda, usando as informações locais de quanto custa mover um grão de areia de um lugar para outro.

A teoria do transporte ótimo pode ser descrita informalmente usando as palavras do Matemático francês Gaspard Monge (1746-1818): Um trabalhador com uma pá na mão tem que mover um grande monte de areia deitado em um canteiro de obras. O objetivo do trabalhador é erguer com toda a areia uma nova pilha com uma forma prescrita (por exemplo, a de um castelo de areia gigante). Naturalmente, o trabalhador deseja minimizar seu esforço total, quantificado por exemplo, como a distância total ou o tempo gasto carregando pás cheias de areia. ([PEYRÉ; CUTURI et al., 2019](#), p. 1)

Para [Thorpe \(2018\)](#), existem duas maneiras de formular o problema de transporte na Teoria do Transporte Ótimo: O Problema de Monge e o Problema de Kantorovich. A formulação de Kantorovich pode ser vista como uma generalização do problema de Monge, cuja principal vantagem é teórica. Por exemplo, a existência de solução para o problema de Kantorovich é sempre garantida, enquanto que resultados de existência de solução para Monge são consideravelmente mais difíceis.

Além disso, existem diversas aplicações da Teoria do Transporte Ótimo na atualidade. Por exemplo, [Thorpe \(2018\)](#) pontua: recuperação de imagem, representação de sinal e imagem, problemas inversos, detecção de câncer, textura e modelagem de cores, registro de formas e imagens e aprendizado de máquina.

Estrutura do Trabalho

Iniciaremos esse trabalho fazendo uma abordagem sobre Modelagem Matemática, em seguida, realizamos um breve passeio pela evolução histórica da Pesquisa Operacional, onde mostramos os principais marcos do desenvolvimento desta teoria, da Programação Linear e da Teoria de Transporte Ótimo.

No Capítulo 2, detalharemos uma sequência de passos que seguiremos para nos levar a alcançar o objetivo definido pela Programação Linear. Apresentaremos um exemplo motivador, o qual resolveremos baseado nos passos citados anteriormente e que servirá de base para os outros problemas e modelos investigados neste texto. Ainda nesse Capítulo, estudaremos de forma concisa a formulação teórica de um problema de Programação Linear e seu principal meio de resolução, o Método *Simplex*.

Discutiremos no Capítulo 3 um pouco sobre a Teoria do Transporte Ótimo e, para tanto, faremos um relato histórico descrevendo o Problema de Monge e, em seguida, o Problema de Kantorovich. Veremos uma explanação acerca de algumas propriedades métricas do transporte ótimo, que serão úteis para introduzir (e demonstrar) a Métrica de Wasserstein no final do capítulo.

Por fim, no capítulo 4, apresentamos algumas aplicações do problema de transporte diretamente associado à comparação de histogramas. Para estas comparações, utilizamos a Métrica de Wasserstein com o intuito de reconhecer o fenômeno de “onda” epidemiológica nos histogramas associados à propagação da pandemia de COVID-19 no Brasil.

1 Modelagem Matemática

Ao passo que resolvemos problemas diretamente relacionados com nossa realidade através de um modelo matemático, estamos criando saber matemático-significativo. Muitas vezes dizemos, dentro deste contexto, que estamos produzindo conhecimento aplicado. Os termos “aplicações” e “modelagem” são frequentemente utilizados para apresentar várias formas de conectar a Matemática à realidade. Por outro lado, vale ressaltar que existe uma diferença entre estes dois termos: a modelagem matemática é entendida como o processo de criação de um padrão ou fórmula matemática, para explicação ou compreensão de um fenômeno natural que pode ser de qualquer área do conhecimento; enquanto que fazer uma aplicação é utilizar a Matemática para resolver problemas do mundo real, visando valorizar, principalmente, situações que são acessíveis a um determinado modelo matemático preexistente e muitas vezes já explorado.

Em função de tal diferença, alguns autores apresentam divergências de concepções em torno do uso da Modelagem Matemática em sala de aula. Tanto que, [Borba, Meneghetti e Hermini \(1999, p. 76\)](#) consideram que a modelagem pode ser vista como um esforço de descrever matematicamente um fenômeno que é escolhido pelos alunos com o auxílio do professor. Já para [Bassanezi \(2002, p. 38\)](#) um fenômeno a ser modelado deve servir de “força motivadora” para o aprendizado dos conteúdos matemáticos da mesma forma que, segundo ele, as discussões sobre o tema a ser trabalhado em aula, favorecem a formação do indivíduo como elemento ativo no seu contexto social. Com isso, entende-se que esta é uma forma de agregar significado matemático ao conteúdo trabalhado.

Dentre as principais áreas de conhecimento que respaldam a Modelagem Matemática, destacamos a Pesquisa Operacional que se vale de métodos quantitativos para auxiliar na busca da melhor maneira de se tomar uma decisão dentro de uma organização, levando em conta as condições impostas por sua realidade.

Podemos usar a Pesquisa Operacional, por exemplo, quando se pretende expressar as melhores decisões, como o custo mínimo ou lucro máximo de produção, considerando o fato de possuir recursos limitados, ou então, apresentar as melhores rotas em um problema de transporte de cargas, tendo que obedecer à limitação de oferta das origens e também sempre atender no mínimo a demanda de mercadoria dos destinos.

É bem verdade que existem diversos métodos e técnicas para se resolver problemas de Pesquisa Operacional como, por exemplo, Teoria dos Grafos, Teoria de Filas, Programação Linear, Programação Inteira, Programação Inteira Mista, Programação Não-Linear, dentre outras. Este trabalho dará ênfase a Programação Linear como principal técnica empregada para resolver problemas de Pesquisa Operacional, por isso, iremos focar nossas análises na

resolução de problemas cujo modelo matemático seja representado por expressões lineares.

1.1 Breve Histórico

O processo de desenvolvimento histórico da Modelagem Matemática é marcado por uma larga colaboração entre cientistas e militares que sempre buscaram alcançar as melhores estratégias durante as batalhas para que assim obtivessem a vitória. Nesta seção, listamos de forma concisa importantes fatos históricos associados ao conteúdo deste trabalho (consulte ([GRANJA; RUIZ, 2006](#)) e ([VILLANI, 2009](#), Páginas 29 até 37)).

Muitos estudiosos consideram o início da Pesquisa Operacional no século III A.C, durante a Segunda Guerra Púnica, tendo como base fundamentadora as propostas de soluções de Arquimedes para defender a cidade de Siracusa, sitiada pelos romanos. Entre suas invenções se encontravam as catapultas de guerra com as quais podiam lançar pedras de um quarto de tonelada a um quilômetro de distância e um sistema de espelhos para incendiar as navegações inimigas concentrando os raios do sol.

Nos séculos XVII e XVIII, Isaac Newton, Gottfried Leibniz, Joseph-Louis Lagrange e vários membros da família Bernoulli, obtiveram, à luz de seus trabalhos, resultados sobre máximos e mínimos de funções. Enquanto que o matemático francês Jean-Baptiste Fourier apresentou algoritmo para resolver sistemas de inequações, uma ideia preliminar do que hoje conhecemos como Programação Linear. Já no final do século XVIII, Gaspard Monge conseguiu desenvolver os primeiros passos do método gráfico através do avanço da Geometria Descritiva.

Frederick Taylor, durante o século XIX, conseguiu mostrar através de seus estudos que durante o trabalho de escavação das minas a única variável realmente significativa era o peso combinado da pá e sua carga, sendo assim, conseguiu maximizar o desempenho dos mineiros. A partir disso, foi possível desenhar pás de acordo com os diferentes tipos de materiais que seriam usados.

Dentre os especialistas, há alguns que consideram Charles Babbage como o pai da Pesquisa Operacional, atribuição conferida por conta de seus estudos investigativos acerca dos custos de transporte e triagens do correio feitas para o Uniform Penny Post da Inglaterra em 1840.

Durante a Primeira Guerra Mundial, o engenheiro inglês Frederick Lanchester estudou sobre o poder balístico das forças opositoras e conseguiu desenvolver, a partir de um sistema de equações diferenciais, a Lei Quadrada de Lanchester. Através desta lei era possível determinar o resultado de uma batalha militar em função da força numérica relativa e a capacidade de fogo dos combatentes.

A primeira abordagem feita para o Problema de Transporte foi em 1781, por Gaspard

Monge, motivado por um problema de engenharia. Em 1930, o russo Lev Nikolaevitch Tolstói publicou o artigo “Métodos de Encontrar a Quilometragem Mínima no Transporte de Carga no Espaço”, que foi publicado na coleção Volume de Planejamento de Transporte I para o Comissariado Nacional de Transporte da União Soviética.

Em 1939, o matemático russo Leonid Kantorovich e o economista holandês Tjalling Koopmans deram importantes contribuições para a logística e a economia ao sugerir uma nova abordagem ao problema proposto por Monge. Com isso, estes tipos de problemas passaram a serem conhecidos como “Problemas de Transporte de Kantorovich-Koopmans”.

No transcorrer da década de 1940, Kantorovich e Koopmans desenvolveram, cada um com sua própria visão independente, o problema do transporte, iniciando áreas de pesquisa que mais tarde viriam a ser chamadas de “Teoria de Transporte Ótimo” e de “Programação Linear”.

Contudo, os grandes estudiosos consideram que o maior marco do desenvolvimento da Pesquisa Operacional acontece durante a II Guerra Mundial, na batalha da Grã-Bretanha, quando o governo britânico buscou alguma forma para melhorar a insuficiência da capacidade aérea de suas forças armadas. Mesmo contando com uma grande experiência em combate, resolveu convocar cientistas de diversas disciplinas para solucionar o problema e extrair o máximo de benefício dos radares recém-inventados que dispunham. Com essa força tarefa, conseguiram determinar a localização ideal das antenas para melhor distribuição de sinal e, com isso, foi possível dobrar a eficácia do sistema de defesa aérea evitando a derrota.

No ano de 1942, a frota de submarinos alemães começou um bloqueio à Grã-Bretanha, por meio de ataques aos comboios de navios que transportavam suprimentos dos Estados Unidos, impedindo-os que chegassem ao seu destino. O *Anti-Submarine Warfare Operations Research Group (ASWORG)*¹ conseguiu apresentar modelos matemáticos desses navios que consideravam as diversas restrições impostas pela realidade, como a velocidade máxima atingida pelos navios, a capacidade de suprimentos a ser transportado e o combustível necessário para chegar ao seu destino. Foram capazes também de realizar essas representações matemáticas para frota de submarinos alemã e, com isso, conseguiram modelar a guerra naval, chegando a desenvolver estratégias que reduziram o número de navios americanos afundados e ainda aumentaram a destruição de submarinos alemães.

Ao passo que a Inglaterra conseguiu avaliar a interferência dessas novas estratégias, prontamente se criaram grupos de “Pesquisa de Operações” tentando obter melhores resultados na batalha. Em 1947, os Estados Unidos formou um grupo de trabalho dedicado a melhorar o processo de planejamento em larga escala o projeto: *Scientific Computation Of Optimum Programs (SCOOP)*².

¹ Tradução livre do autor: Grupo de Pesquisa de Operações de Guerra Anti-Submarino

² Tradução livre do autor: Computação Científica de Programação Otimizada

Em 1949, Dantzig apresentou um método numérico de resolução do problema de transporte, dando os primeiros passos para o desenvolvimento da Programação Linear. Curiosamente, algumas técnicas associadas a este método já haviam sido trabalhadas por Kantorovich. Tanto que até hoje, ainda existem autores que divergem sobre qual destes deve ser considerado o verdadeiro fundador da Programação Linear.

No Brasil, a Pesquisa Operacional apareceu em meados do século XX quando, no final da década de 50, foi introduzida em alguns trabalhos de profissionais empresariais. Apesar disso, teve seu principal berço por meio do desenvolvimento acadêmico, segundo Lóss (1981). Em 1959, o Instituto Tecnológico de Aeronáutica (ITA) deu início ao ensino das primeiras disciplinas de Pesquisa Operacional (tais como teoria de jogos, simulação, filas e estatística) ministradas para o primeiro curso de Engenharia de Produção do Brasil, que teve início em 1957. Enquanto isso a Universidade de São Paulo (USP) implantou a disciplina de Programação Linear. Concomitantemente, a Pontifícia Universidade Católica do Rio de Janeiro (PUC-RJ), adquiria seu primeiro computador e um grupo de pesquisadores começaram a trabalhar com modelos e aplicações de técnicas de Pesquisa Operacional. A partir disso o Instituto de Matemática Pura e Aplicada (IMPA) e a Escola Nacional de Ciências Estatísticas (ENCE) mostraram maior interesse em desenvolver pesquisas na área de Programação Linear. Na década de 60, a empresa Petróleo Brasileiro S.A. (PETROBRAS) fundou a Sociedade Brasileira de Pesquisa Operacional (SOBRAPO), a qual se configurou no primeiro grupo brasileiro de estudos em Pesquisa Operacional. Segundo esta sociedade, a Pesquisa Operacional é uma ciência aplicada voltada para a resolução de problemas reais, tendo como principal foco avaliar linhas de ação alternativas para encontrar as soluções que melhor servem aos objetivos dos indivíduos ou organizações.

As contribuições do problema de transporte à economia foi o principal tópico do *Prêmio de Ciências Econômicas em Memória de Alfred Nobel* de 1975 concedido à Leonid Kantorovich e Tjalling Koopmans.

A partir da década de 1990, a Teoria de Transporte Ótimo foi revisitada tanto por matemáticos teóricos, pelas aplicações ao tópico de Equações Diferenciais Parciais, como por matemáticos computacionais, devido as aplicações a área de *Machine Learning*³. Vale ressaltar que pesquisadores deste tópico, como Cédric Villani (em 2010) e Alessio Figalli (em 2018), foram agraciados com o *International Medal for Outstanding Discoveries in Mathematics*⁴, também conhecida como Medalha *Fields*, que é a mais importante honraria da Matemática da atualidade.

³ Tradução livre do autor: Aprendizagem de Máquina

⁴ Tradução livre do autor: Medalha Internacional de Descobrimientos Proeminentes em Matemática

2 Programação Linear

Neste capítulo, estudaremos o problema de transporte sob o ponto de vista da Programação Linear. De acordo com [Marins \(2011\)](#) a Programação Linear tem como objetivo obter a melhor solução para problemas cujo modelo possuam expressões lineares. Desta forma, a otimização ocorrerá por meio do estudo de uma função linear, chamada de função objetivo, que deverá atender à um sistema de igualdades e/ou desigualdades lineares, às quais denominamos de restrições.

Inicialmente, discutiremos o processo de modelagem e as etapas de resolução de um problema da Programação Linear através de um exemplo de transporte associado a distribuição de produtos. Nas seções [2.3](#) e [2.4](#), serão apresentados os alicerces teóricos da Programação Linear que viabilizam a otimização dos problemas propostos neste texto. Listamos as seguintes referências ([MARINS, 2011](#); [GOLDBARG; LUNA, 2005](#); [BAZARAA; JARVIS; SHERALI, 2008](#); [DANTZIG; THAPA, 2006](#)) onde podem ser encontrado todo o conteúdo desenvolvido neste capítulo.

2.1 Modelagem de uma Situação-Problema

[Marins \(2011\)](#) defende que para se resolver um problema de Programação Linear são necessárias as etapas de observação, de formulação do problema, de construção do modelo científico e de resolução deste modelo por meio de método adequado, o qual pode ser determinístico ou probabilístico por se tratar de uma abordagem científica.

Desta forma, para resolvermos um problema de Programação Linear, primeiramente precisamos realizar uma abordagem para encontrarmos uma modelagem matemática para esse problema, para que tenhamos uma melhor compreensão da situação-problema e sejamos capazes de criar modelos que representem bem tais situações, e a partir de então, usarmos um método de resolução que nos leve à melhor solução do modelo.

Independentemente do método a ser aplicado para a análise e resolução de um Problema de Programação Linear (PPL), será necessário seguir algumas etapas básicas que, em linhas gerais, são apresentadas e detalhadas nesta seção.

Etapa 1. Compreensão da Situação-Problema

A definição e compreensão do problema a ser estudado trará uma visão exata dos objetivos a serem alcançados, uma vez que a qualidade e o nível de confiança do resultado final são diretamente proporcionais à precisão da análise inicial. Vale a pena salientar

que ao longo desse processo de entendimento inicial da situação-problema, pode haver suposições em que as informações não são bem estruturadas, e por isso, é importante saber que quanto mais objetivas forem as informações, melhor será o resultado final. Além disso, esta etapa é a principal base para definição das limitações impostas pela natureza do problema que irá configurar as chamadas restrições que devem serem atendidas.

Etapa 2. Construção de um Modelo Matemático

Uma vez compreendido a situação-problema, devemos seguir para a construção de um modelo, que deve ser feita através da representação da situação-problema e de suas restrições através de modelos matemáticos para nortear a busca por uma solução.

Tal modelo deve ser composto por um conjunto de variáveis envolvidas em equações e inequações que descrevem o problema e suas restrições. Vale a pena lembrar, que nem sempre um modelo matemático descreve a situação real em sua totalidade, pois pode haver variáveis que não serão incluídas por existir um número muito grande delas, o que dificulta a modelagem e uma possível resolução. A quantidade de variáveis a serem representadas matematicamente pode variar de acordo com as necessidades impostas pela situação-problema e estas são chamadas de variáveis de decisão.

De forma resumida, apresentamos a seguir uma sequência de passos que auxiliam na construção de um modelo matemático de Programação Linear que melhor represente a situação-problema de forma coerente, clara e concisa:

- 2A. Identificar as variáveis de decisão e representá-las em simbologia algébrica;
- 2B. Identificar as restrições do problema e expressá-las como equações ou inequações lineares em termos das variáveis de decisão;
- 2C. Identificar o objetivo de interesse no problema e representá-lo como função linear em termos das variáveis de decisão. O objetivo deve compreender a otimização desta função, isto é, maximizar ou minimizar tal função definida.

Em consonância com a sequência de passos apontados acima, é muito importante que nessa fase sejam feitas e respondidas muitas perguntas, tais como: Quais são as hipóteses que devem ser sugeridas? Quais decisões devem ser tomadas? Quais recursos afetam tais decisões?

Etapa 3. Resolução do Modelo

Nessa etapa, buscamos uma solução do modelo de Programação Linear que foi proposto. Esta solução é chamada de solução ótima e pode ser obtida através de técnicas

ou ferramentas de solução numérica e computacionais, que geralmente são baseadas em um algoritmo.

Um algoritmo bastante utilizado na Programação Linear é o Método *Simplex*, com uma boa contextualização em [Dantzig \(1963\)](#). Outro algoritmo muito poderoso foi proposto por [Karmarkar \(1984\)](#) e consiste no Método dos Pontos Interiores. Detalhes sobre ambos podem ser obtidos em [Hillier e Lieberman \(2013\)](#).

Uma poderosa ferramenta computacional muito utilizada na resolução de problemas de Programação Linear é o complemento *Solver*.

[Gomes Júnior e Souza \(2004\)](#) afirmam que:

O *Solver* faz parte de um conjunto de programas algumas vezes chamado de ferramentas de análise hipotética. Com o *Solver* você pode localizar um valor ideal para uma fórmula em uma célula, chamada de célula de destino, em uma planilha. O *Solver* trabalha com um grupo de células relacionadas direta ou indiretamente com a fórmula na célula de destino. O *Solver* ajusta os valores nas células variáveis que você especificar, chamadas de células ajustáveis, para produzir o resultado especificado por você na fórmula da célula de destino. Você pode aplicar restrições para restringir os valores que o *Solver* poderá usar no modelo e as restrições podem se referir a outras células que afetem a fórmula da célula de destino.

Etapa 4. Análise da Solução

A avaliação da solução ótima tem por objetivo principal determinar se as conclusões encontradas para o modelo proposto escolhido para representar a situação-problema, são adequadas ou não para o problema real. Caso a solução seja julgada inadequada, ou seja, apresente um resultado discrepante, será preciso redefinir o problema ou redefinir o modelo. Caso a solução seja apropriada, teremos uma solução adequada para o problema e, então, deve-se fazer a aplicação real da solução ótima encontrada.

Com base nesses passos, espera-se que tenhamos uma solução que seja condizente com o ambiente real, ou no mínimo, que a solução ótima seja exequível para apreciação dentro da realidade.

É importante frisarmos que a sequência de passos listados anteriormente é flexível, ou seja, não é necessário seguir todo o ciclo para que possamos concluir se a solução encontrada é boa ou ruim para representar o problema real. Assim, pode ser necessário fazer uma modificação no modelo, para que possamos encontrar uma adaptação melhor do modelo à realidade, ou mesmo, objetivando encontrarmos uma solução mais precisa para o problema.

2.2 Exemplo Motivador

Iremos abordar nessa seção, um problema de transporte que tem resolução capaz de ser facilmente entendida e que nos possibilite dialogar sobre suas restrições, o que denota uma abordagem com entendimento claro e possível de ser aplicado no Ensino Médio. O problema abaixo foi inspirado em (CAIXETA FILHO, 2001).

Situação-Problema. Suponha que uma empresa é responsável pela distribuição da produção agrícola de um produtor de arroz paulista. Para isso, deve-se levar em conta todo o custo do transporte da produção. Sabe-se que o produtor possui três locais de produção:

Fazenda Mariana, Fazenda Braga e Fazenda Falcão.

Considera-se ainda que cada local de produção possui uma oferta mensal de no máximo

100, 150 e 170 toneladas de arroz,

respectivamente. Esta produção deve ser transportada até dois diferentes centros de distribuição (CD), que se localizam em

Campinas e Sorocaba,

e que requerem, respectivamente, uma demanda mensal de

250 e 140 toneladas de arroz.

Como realizar esse transporte com o menor custos e atendendo às especificações do problema?

Etapa 1. Compreensão da situação-problema

Para se obter uma melhor compreensão, iremos enumerar tanto as origens quanto os destinos, com números naturais em ordem crescente. Assim, a Fazenda Mariana é 1, a Fazenda Braga é 2, a Fazenda Falcão recebe 3, o centro de distribuição de Campinas será 4 e, por fim, o centro de distribuição de Sorocaba é 5. Esta enumeração visa facilitar a identificação das variáveis e a associação com as trajetórias.

Etapa 2. Construção de um Modelo Matemático

Veja que para modelar este problema deve-se considerar cada variável de decisão como sendo a quantidade que deve ser transportada da origem i , para o destino j , representada por x_{ij} .

Como exemplo, seja a quantidade de arroz que sairá da Fazenda Braga (2) e abastecerá o centro de distribuição de Sorocaba (5). Na notação proposta, será a quantidade a ser transportada de 2 para 5. Logo, a origem será 2 e o destino será 5, obtendo a variável definida por x_{25} . Da mesma maneira, a quantidade de arroz que sairá da Fazenda Braga (1) para o centro de Campinas (4), terá origem 1 e destino 4, obtendo a variável x_{14} .

Observe que na modelagem desse problema, cada uma das três fazendas tem a opção de levar seus produtos a dois diferentes locais, o que gera seis rotas diretas diferentes, isto significa que estamos diante de um problema de seis variáveis.

Criamos a Tabela 1, para podermos realizar a análise dos custos de transporte por tonelada de arroz em todas as trajetórias possíveis do problema. Essa tabela é uma idealização baseada apenas em critérios não oficiais para representar o que chamamos de matriz dos custos de transporte.

Tabela 1 – Custos de Transporte

R\$ / Toneladas	CD Campinas (4)	CD Sorocaba (5)
Fazenda Mariana (1)	12,50	6,80
Fazenda Braga (2)	13,00	9,00
Fazenda Falcão (3)	10,00	8,90

Fonte: Elaborada pelo autor.

Assim, por exemplo, a quantidade a ser transportada de 1 para 5, cuja variável de decisão é x_{15} , terá um custo de $6,8x_{15}$, pois x_{15} representa a quantidade, em toneladas, de arroz a ser transportada e 6,8 representa segundo a matriz dos custos, o valor, em reais (R\$) do transporte de uma tonelada de arroz nesse possível caminho. Seguindo esse raciocínio teremos a função que representa o custo total do transporte (Z). Vale salientar, que se alguma rota acima não possuir transporte o valor de sua variável de decisão será zero; com isso, teremos a seguinte função objetivo,

$$Z = 12,5x_{14} + 6,8x_{15} + 13x_{24} + 9x_{25} + 10x_{34} + 8,9x_{35}.$$

Elaborada a função custo, iremos analisar as restrições do problema. A produção máxima da Fazenda Mariana (1) é de 100 toneladas de arroz, logo a soma das variáveis que começam com (1) é menor do que ou igual a 100, assim,

$$x_{14} + x_{15} \leq 100.$$

Analogamente, a Fazenda Braga (2) tem produção máxima de 150 toneladas e a Fazenda Falcão (3) possui produção máxima de 170 toneladas, com isso,

$$x_{24} + x_{25} \leq 150 \quad \text{e} \quad x_{34} + x_{35} \leq 170.$$

Atribuída as restrições de produção, analisaremos as restrições de demanda, onde iremos analisar a capacidade total de cada centro de distribuição.

O CD Campinas (4) tem capacidade máxima para 250 toneladas de arroz, logo, o somatório das variáveis que terminam com (4) terá que ser igual a 250, ou seja,

$$x_{14} + x_{24} + x_{34} = 250.$$

Da mesma forma, o CD Sorocaba (5) tem demanda máxima de 140 toneladas de arroz, o que gera a seguinte restrição,

$$x_{15} + x_{25} + x_{35} = 140.$$

Finalmente, acrescentam-se as condições de não negatividade, as quais determinam que todas as variáveis de decisão do problema não podem assumir valores negativos, vale lembrar que essas condições podem ser implementadas a partir de um link, durante a resolução do problema por meio do *Solver* que é um complemento do *Excel*.

Feita a análise do problema, apresentaremos abaixo o Quadro 1 com o modelo de Programação Linear para o mesmo, contendo a função objetivo, as restrições de produção, as restrições de demanda e as restrições de não negatividade.

Quadro 1 – Modelo Matemático do Exemplo Motivador

<p>Função Objetivo</p> <p>Minimizar $Z = 12,5x_{14} + 6,8x_{15} + 13x_{24} + 9x_{25} + 10x_{34} + 8,9x_{35}$.</p> <p>Sujeito às Restrições de Produção</p> $x_{14} + x_{15} \leq 100$ $x_{24} + x_{25} \leq 150$ $x_{34} + x_{35} \leq 170$ <p>Sujeito às Restrições de Demanda</p> $x_{14} + x_{24} + x_{34} = 250$ $x_{15} + x_{25} + x_{35} = 140$ <p>Sujeito às Restrições de Não Negatividade</p> $x_{14} \geq 0 \quad x_{15} \geq 0$ $x_{24} \geq 0 \quad x_{25} \geq 0$ $x_{34} \geq 0 \quad x_{35} \geq 0$ <p>Onde</p> <p>$\{x_{14}, x_{15}, x_{24}, x_{25}, x_{34}, x_{35}\}$ são as variáveis de decisão.</p>
--

Fonte: Elaborado pelo autor.

Etapa 3. Resolução do Modelo (via *Solver*)

Feito a modelagem do problema lançamos mão do *Excel* para adicionar os dados do problema. Em nosso exemplo, precisamos calcular os totais que irão atender às restrições do problema, esses totais estão nas células de *I6* (coluna I e linha 6) a *I12* (coluna I e linha 12), para tanto, usaremos a função do *Excel* conhecida como [*SOMARPRODUTO*], digitamos essa fórmula na célula *I6*, em seguida, escolhemos as células de *C6* a *H6*, que são onde tem os coeficientes das variáveis, depois digitamos o ponto e vírgula (;) e selecionamos às células que contém as variáveis que são de *C14* a *H14*. No final, apertamos a tecla *F4* para travarmos estas células e assim poderemos transferir esta fórmula corretamente para as demais células da coluna de Totais, agora basta teclarmos *ENTER*.

Para transferir a fórmula para as demais colunas basta selecionar a *I6*, clicar sobre o quadradinho em destaque no canto inferior direito dessa célula e arrastar o cursor sobre as demais células segurando o clique inicial, como mostra a Figura 1.

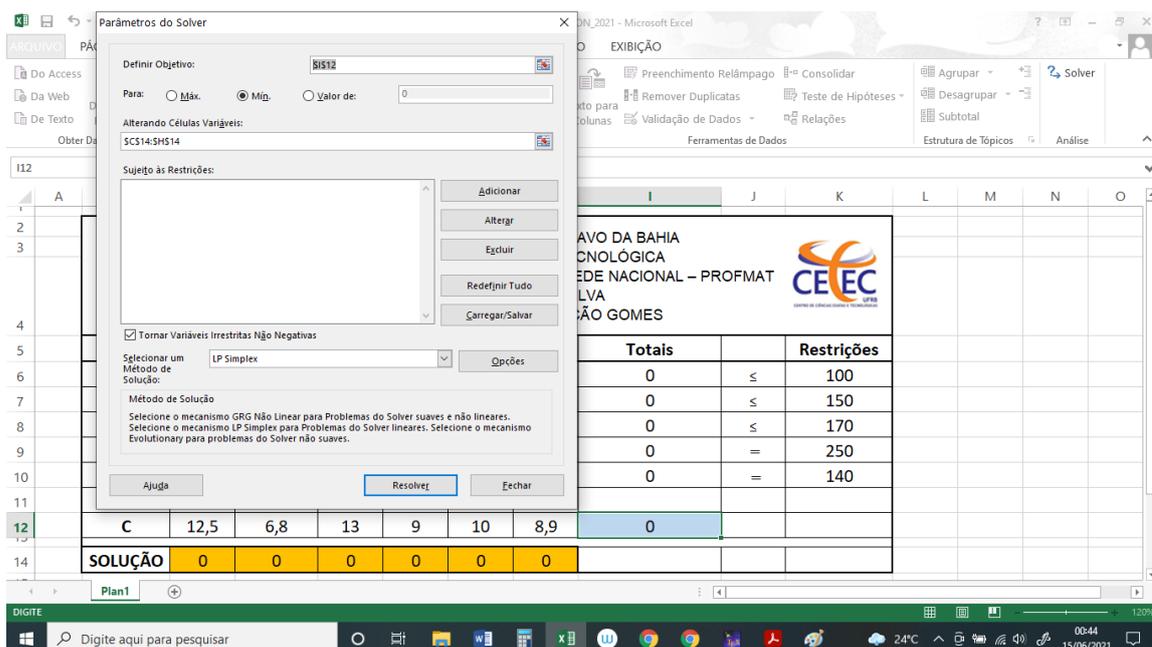
Figura 1 – Incluindo Dados no Excel



Fonte: Elaborada pelo autor.

Nesse momento, iremos usar o suplemento do Excel conhecido como Solver. Com esta ferramenta devidamente habilitada, iremos acessar o ícone DADOS na janela do Excel e no canto superior direito irá aparecer a janela com o símbolo do Solver sobre o qual iremos clicar para abrir a janela de alimentação de dados desta ferramenta, conforme Figura 2.

Figura 2 – Alimentando Informações no Solver I



Fonte: Elaborada pelo autor.

Com a janela do *Solver* aberta vamos informar os dados do problema. Aparecerá o campo “*Definir Objetivo:*”. Selecionamos este campo da janela e clicamos sobre a célula *I12*, que representa o total de nossa função custo. Em seguida, marcamos a forma de otimização que queremos que o programa realize, isto é, “*Min*” para minimizar o problema. Também devemos preencher a linha “*Alterando Células Variáveis:*”. Para isto selecionamos esta linha da janela, clicamos e arrastamos sobre as células que representam as variáveis, as quais são *C14* até *H14*.

Clicamos em adicionar e irá abrir uma nova janela na qual devemos informar as restrições do problema, ver Figura 3.

Figura 3 – Alimentando Informações no *Solver* II

The screenshot shows an Excel spreadsheet with the following data table:

	X14	X15	X24	X25	X34	X35	Totais	Restrições
R1	1	1	0	0	0	0	0	≤ 100
R2	0	0	1	1	0	0	0	≤ 150
R3	0	0	0	0	1	1	0	≤ 170
R4	1	0	1	0	1	0	0	= 250
R5	0	1	0	1	0	1	0	= 140
C	12,5	6,8	13	9	10	8,9	0	
SOLUÇÃO	0	0	0	0	0	0		

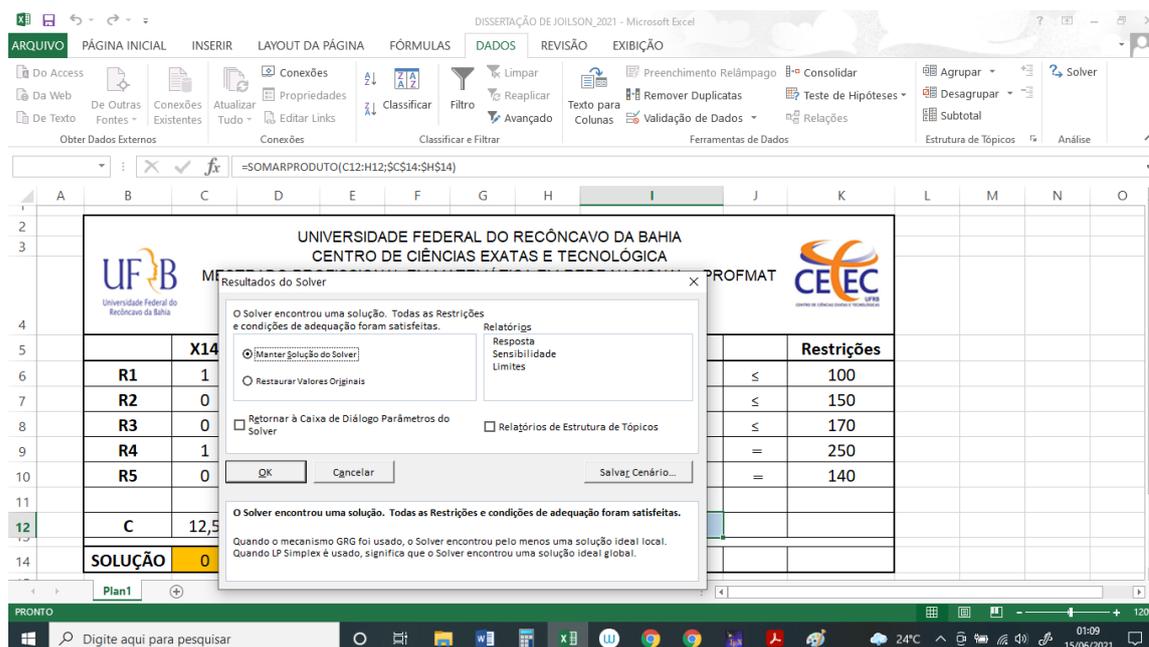
The dialog box 'Adicionar Restrição' is open, showing the reference cell '\$I\$6' and the constraint '= <= \$K\$6'. The spreadsheet also includes a header for 'UNIVERSIDADE FEDERAL DO RECÔNCAVO DA BAHIA' and a title 'MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL – PROFMAT'.

Fonte: Elaborada pelo autor.

Colocando o cursor sobre a linha “*Referência de Célula*” e clicamos sobre a célula *I6* da coluna (**TOTAIS**). Também escolhemos o símbolo referente a essa restrição clicando na seta da coluna central da janela “*Adicionar Restrições*”. Em seguida, colocamos o cursor sobre a coluna restrição da janela e depois clicamos sobre a célula da primeira restrição *J6*.

Como nosso problema tem outras restrições, clicamos em adicionar e repetimos o procedimento anterior para as demais restrições (veja Figura 3 para facilitar a compreensão) e, ao final da última restrição, clicamos em *OK*.

Ao clicarmos em *OK*, seremos encaminhados para a primeira janela do *Solver*. Agora, devemos clicar no botão *Resolver* e então abrirá uma ultima janela.

Figura 4 – Resolvendo com o *Solver*

Fonte: Elaborada pelo autor.

Nesta última etapa, o *Solver* nos fornece algumas opções, conforme mostra a Figura 4, temos a opção de: Manter a solução por ele encontrada; exibir alguns relatórios; ou restaurar as opções para valores iniciais. Iremos manter a solução e clicar *OK*. Logo em seguida, nos será exibida a solução ótima encontrada pelo *Solver*.

Etapa 4. Análise da Solução

Com a solução, consegue-se analisar exatamente como se proceder o transporte dos centros de distribuições até as respectivas fazendas, a fim de que o custo seja mínimo, ou seja, somos capazes de identificar a melhor quantidade a ser transportada de uma das fazendas para um determinado centro de distribuição; ou até mesmo, definir se vale a pena não fazer o transporte em uma determinada rota.

Ao observarmos os resultados (veja Figura 5) notamos que no campo correspondente à variável de decisão x_{15} , dado pela célula $D14$, aparece o valor 100. Isto significa que a informação da quantidade ótima de arroz a ser transportada de 1 para 5, ou mais precisamente, da fazenda Mariana para o centro de distribuição de Sorocaba, é de 100 toneladas. Da mesma forma, verificamos que o mais econômico é que não haja transporte da fazenda Mariana para o centro de distribuição de Campinas, pois a variável de decisão x_{14} , na célula $C14$, é igual a zero.

Figura 5 – Solução do Solver

The screenshot shows an Excel spreadsheet with the following data:

	X14	X15	X24	X25	X34	X35	Totais	Restrições
R1	1	1	0	0	0	0	100	≤ 100
R2	0	0	1	1	0	0	120	≤ 150
R3	0	0	0	0	1	1	170	≤ 170
R4	1	0	1	0	1	0	250	= 250
R5	0	1	0	1	0	1	140	= 140
C	12,5	6,8	13	9	10	8,9	3780	
SOLUÇÃO	0	100	80	40	170	0		

Fonte: Elaborada pelo autor.

Outra informação importante, é o fato de que para obter o menor custo ótimo de transporte, apenas a fazenda Braga utilizou 120 toneladas de arroz das 150 existentes, enquanto as demais fazendas transportaram o total de arroz produzido para suprir a capacidade total de $390 = (250+140)$ toneladas de arroz para os dois centros de distribuição. Por fim, obtemos a informação de que o menor custo total do transporte para atender às condições do problema será de

$$Z = R\$ 3.780,00.$$

2.3 Formulação Teórica

Estabelecidos os procedimentos práticos para resolução de um problema de transporte segundo à Programação Linear, a partir desta seção, iremos estudar as técnicas e resultados matemáticos que embasam tal teoria.

Em linhas gerais, Programação Linear pode ser entendida como uma técnica que pressupõe a relação linear entre as características do problema, buscando a solução de problemas que objetivam a otimização de um sistema de estudo. Essas características do problema são representadas e relacionadas por meio de uma série de equações e/ou inequações lineares.

Goldberg e Luna (2005), defendem que os modelos de Programação Linear constituem um tipo especial de modelos de otimização. Para que um determinado sistema

possa ser representado por meio de um modelo de Programação Linear, ele deve possuir as seguintes características:

- *Proporcionalidade*: a quantidade de recurso consumido por uma dada atividade deve ser proporcional ao nível dessa atividade na solução final do problema. Além disso, o custo de cada atividade é proporcional ao nível de operação da atividade.
- *Não Negatividade*: deve ser sempre possível desenvolver dada atividade em qualquer nível não negativo, e qualquer proporção de um dado recurso deve sempre poder ser utilizado.
- *Aditividade*: o custo total é a soma das parcelas associadas a cada atividade.
- *Separabilidade*: pode-se identificar de forma separada o custo (ou consumo de recursos) específico das operações de cada atividade.

Com isso, a formulação geral de um problema de Programação Linear, é apresentada no Quadro 2.

Quadro 2 – Formulação Teórica

<p>Função Objetivo</p> <p>Maximizar (ou Minimizar) $Z = c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + \dots + c_nx_n$</p> <p>Sujeita às Restrições</p> $a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n \leq b_1 \quad (\text{ou } \geq \text{ ou } =)$ $a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n \leq b_2 \quad (\text{ou } \geq \text{ ou } =)$ $a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \dots + a_{3n}x_n \leq b_3 \quad (\text{ou } \geq \text{ ou } =)$ \vdots $a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \dots + a_{mn}x_n \leq b_m \quad (\text{ou } \geq \text{ ou } =)$ <p>Onde</p> <p>$\{c_1, c_2, c_3, \dots, c_n\}$ são os coeficientes das variáveis na função objetivo Z; $\{x_1, x_2, x_3, \dots, x_n\}$ são as variáveis de decisão; $\{a_{11}, a_{12}, a_{13}, \dots, a_{mn}\}$ são os coeficientes das variáveis nas restrições; $\{b_1, b_2, b_3, \dots, b_m\}$ são os termos independentes das restrições.</p>

Fonte: Elaborado pelo autor.

Independente do método escolhido para resolução, um problema de Programação Linear é representado por um modelo simbólico estruturado com equações e/ou inequações lineares, donde também podemos propor uma reformulação matricial para representá-lo, ver Quadro 3.

Quadro 3 – Modelo Algébrico

<p>Função Objetivo</p> <p>Maximizar (ou Minimizar) $Z = CX$</p> <p>Sujeita às Restrições</p> $AX \leq B, \quad X \geq 0 \quad \text{e} \quad B \geq 0 \quad (\text{ou } \geq \text{ ou } =)$ <p>Onde</p> $C = (c_1, \dots, c_m), \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$
--

Fonte: Elaborado pelo autor.

Observação. Outras formas equivalentes de enunciar um problema de Programação Linear serão exploradas no Apêndice A por meio da noção de dualidade.

2.4 Método *Simplex*

De posse do modelo matemático que compreende um problema de Programação Linear, destacando a função objetivo e as restrições, precisamos utilizar um método de Programação Linear para encontrarmos a solução ótima. O processo que utilizamos para tal tarefa recebe o nome de Método *Simplex*.

O Método *Simplex* foi desenvolvido por George Dantzig e Tjalling Koopmans em 1946, quando trabalhavam para *United States Air Force (USAF)*¹. Este é considerado por muitos como um dos principais algoritmos inventados no século XX e trata-se de uma estratégia de manipulação interativa, utilizada para se determinar numericamente a solução ótima de um problema de Programação Linear.

O método consiste em procurar uma solução inicial e, em seguida, buscar outra solução que seja melhor que a anterior. Este processo se repete até se encontrar uma solução ótima. (GOLDBARG; LUNA, 2005) pontua que,

O *Simplex* é um algoritmo que se utiliza de um ferramental baseado na álgebra linear para determinar, por um método iterativo, a solução ótima de um Problema de Programação Linear. Sua concepção básica é simples e, por isso mesmo, eficiente. Em linhas bastante gerais, o algoritmo parte de uma solução viável do sistema de equações que constituem as restrições do Problema de Programação Linear, solução essa normalmente extrema (vértice). A partir dessa solução inicial vai identificando novas soluções viáveis de valor igual ou melhor que a corrente. O algoritmo, portanto, possui um critério de escolha que permite encontrar sempre novos e

¹ Tradução livre do autor: Força Aérea dos Estados Unidos

melhores vértices da envoltória convexa do problema, e um outro critério que consegue determinar se o vértice escolhido é ou não um vértice ótimo.

Para [Ploskas e Samaras \(2014\)](#) a escolha do elemento de articulação em cada iteração é um dos passos mais críticos para o algoritmo *Simplex*. A flexibilidade da seleção de variáveis que entram e saem permite desenvolver várias regras de articulações. Por este motivo, o Método *Simplex* é utilizado na otimização matemática sequencial para avaliar a melhor solução possível de um problema complexo, dadas determinadas condições operacionais e a quantidade de recursos.

Para resolver um modelo de Programação Linear usando o Método *Simplex*, as seguintes etapas são necessárias:

- Formulário padrão: uma representação matemática do objetivo do problema;
- Apresentar variáveis de folga: transforma as inequações em equações;
- Criar o quadro: uma definição das variáveis envolvidas no problema;
- Variáveis dinâmicas: identificação da variável de decisão;
- Criar um novo quadro: definição dos critérios de parada;
- Verificar a otimização: realização de iterações para avaliar o comportamento das variáveis e chegar na solução desejada;
- Identificar os valores ideais: apresentação da solução ótima.

Já vimos que a formulação padrão de um problema de Programação Linear se dá em torno de um sistema de equações e inequações, com isso, surgem os primeiros desafios do Método *Simplex*. A dificuldade desse tipo de problema pode ser, segundo [Goldbarg e Luna \(2005\)](#), dividida em dois blocos:

- 1) Como obter soluções factíveis básicas do sistema de equações;
- 2) Como evitar o teste de todas as soluções factíveis básicas possíveis para garantir a otimização do sistema.

A partir do modelo algébrico do problema de Programação Linear, apresentado no Quadro 3, iremos renomear alguns elementos para trabalharmos com o Método *Simplex*:

- $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ é a matriz dos coeficientes (ou matriz tecnológica);

- $C = (c_1, c_2, c_3, \dots, c_n)$ é o vetor de custos;
- $X = (x_1, x_2, x_3, \dots, x_n)^T$ é o vetor das variáveis ou incógnitas;
- $B = (b_1, b_2, b_3, \dots, b_m)^T$ é o vetor dos termos independentes.

Além disto, faz-se necessário apresentarmos algumas definições para o melhor entendimento do processo realizado pelo Método *Simplex*.

Definição 2.1 (Partição Básica de A). Toda matriz dos coeficientes A pode ser reorganizada como

$$A = [\bar{B} \mid N]$$

de modo que:

- $\bar{B}_{m \times m}$ representa a matriz básica que é formada por m elementos da matriz dos coeficientes $A_{m \times n}$. Esta matriz básica é invertível, pois possui posto m . Ela é representada como $\bar{B} = [a_{\bar{B}_1}, a_{\bar{B}_2}, \dots, a_{\bar{B}_m}]$, onde cada elemento $a_{\bar{B}_i}$ representa uma coluna da matriz A e os índices $\bar{B}_1, \bar{B}_2, \dots, \bar{B}_m$ são chamados de índices básicos e indicam qual é a coluna da matriz A que faz parte de \bar{B} ;
- $N_{m \times (n-m)}$ representa a matriz não-básica que é formada pelas $n - m$ colunas restantes da matriz A . É representada por $N = [a_{N_1}, a_{N_2}, \dots, a_{N_{n-m}}]$, onde cada elemento a_{N_i} representa uma coluna da matriz A e os índices N_1, N_2, \dots, N_{n-m} são chamados de índices não-básicos e indicam qual é a coluna da matriz A que este elemento pertencia.

Esta partição nas colunas da matriz A , por sua vez, estabelece uma partição no vetor X , dividindo as variáveis básicas das variáveis não-básicas.

Definição 2.2 (Partição Básica de X). O vetor da variáveis pode ser reescrito como

$$X = \begin{bmatrix} X_{\bar{B}} \\ X_N \end{bmatrix},$$

em que

- $X_{\bar{B}} = \begin{bmatrix} x_{\bar{B}_1} \\ x_{\bar{B}_2} \\ \vdots \\ x_{\bar{B}_m} \end{bmatrix}$, é o vetor das variáveis básicas com m componentes;
- $X_N = \begin{bmatrix} x_{N_1} \\ x_{N_2} \\ \vdots \\ x_{N_{n-m}} \end{bmatrix}$, é o vetor das variáveis não-básicas com $n - m$ componentes.

A partir destas definições, o sistema que define as restrições no modelo algébrico da Programação Linear pode ser reescrito utilizando as matrizes \bar{B} , N e os vetores $X_{\bar{B}}$, X_N

$$AX = B \quad \Longleftrightarrow \quad [\bar{B} \mid N] \begin{bmatrix} X_{\bar{B}} \\ X_N \end{bmatrix} = B.$$

Note que $\bar{B} X_{\bar{B}} + N X_N = B$, logo o vetor $X_{\bar{B}}$ é dado por

$$X_{\bar{B}} = \bar{B}^{-1} B - \bar{B}^{-1} N X_N. \quad (2.1)$$

Definição 2.3 (Solução Básica). Considere uma partição básica da matriz dos coeficientes $A = [\bar{B} \mid N]$. A solução que se obtém ao fixar todas as $n - m$ variáveis de X_N em zero é chamada de solução básica. Esta solução pode ser melhor visualizada como

$$\hat{X} = \begin{cases} \hat{X}_{\bar{B}} = \bar{B}^{-1} B \\ \hat{X}_N = 0 \end{cases}$$

Em particular, a solução básica é uma solução factível para o problema e será o ponto de partida na busca da solução ótima, ou seja, a partir dela procura-se uma solução melhor.

Analogamente aos casos anteriores, obtemos a decomposição abaixo.

Definição 2.4 (Partição Básica de C). O vetor de custos será dado por

$$C = [C_{\bar{B}} \mid C_N],$$

onde,

- $C_{\bar{B}} = [c_{\bar{B}_1}, c_{\bar{B}_2}, \dots, c_{\bar{B}_m}]$, são os coeficientes das variáveis básicas na função objetivo;
- $C_N = [c_{N_1}, c_{N_2}, \dots, c_{N_{n-m}}]$, são coeficientes das variáveis não-básicas na função objetivo.

Veamos como a função objetivo também pode ser expressa utilizando a estratégia da partição básica:

$$Z = C X \quad \Longleftrightarrow \quad [C_{\bar{B}} \mid C_N] \begin{bmatrix} X_{\bar{B}} \\ X_N \end{bmatrix} = C_{\bar{B}} X_{\bar{B}} + C_N X_N = Z. \quad (2.2)$$

Substituindo a equação (2.1) na equação (2.2) tem-se

$$Z = C_{\bar{B}} (\bar{B}^{-1} B - \bar{B}^{-1} N X_N) + C_N X_N = C_{\bar{B}} \bar{B}^{-1} B - C_{\bar{B}} \bar{B}^{-1} N X_N + C_N X_N \quad (2.3)$$

Calculando agora a função objetivo na solução básica teremos

$$\hat{Z} = C_{\bar{B}} \hat{X}_{\bar{B}} + C_N \hat{X}_N = C_{\bar{B}} \bar{B}^{-1} B + C_N 0 = C_{\bar{B}} \bar{B}^{-1} B \quad (2.4)$$

Em seguida, iremos apresentar a definição de um vetor auxiliar que facilitará a notação e os cálculos desenvolvidos a seguir (além disso, é utilizado na construção da dualidade no problema da Programação Linear, consulte Apêndice A).

Definição 2.5 (Vetor Multiplicador do *Simplex*). O vetor $m \times 1$ dado por

$$\lambda^T = C_{\bar{B}} \bar{B}^{-1}$$

é chamado de vetor multiplicador do *Simplex* (ou vetor de variáveis duais).

A partir do vetor multiplicador do *Simplex*, realizando as seguintes manipulações algébricas (detalhadas em (WEBER, 2018)), encontraremos a função objetivo. Primeiramente, substituindo a equação (2.4) na equação (2.3), obteremos

$$Z = \hat{Z} - C_{\bar{B}} \bar{B}^{-1} N X_N + C_N X_N. \quad (2.5)$$

Inserindo o vetor multiplicador do *Simplex*, da Definição 2.5, tem-se

$$Z = \hat{Z} - \lambda^T N X_N + C_N X_N = \hat{Z} + (C_N - \lambda^T N) X_N. \quad (2.6)$$

Ao expor, na equação (2.6), os elementos de C_N e N com seus respectivos índices não-básicos, em conjunto com o vetor das variáveis não-básicas X_N , obteremos a seguinte função objetivo

$$Z = \hat{Z} + (c_{N_1} - \lambda^T a_{N_1})x_{N_1} + (c_{N_2} - \lambda^T a_{N_2})x_{N_2} + \cdots + (c_{N_{n-m}} - \lambda^T a_{N_{n-m}})x_{N_{n-m}}. \quad (2.7)$$

Definição 2.6 (Custos Relativos). Os coeficientes $\hat{c}_{N_j} = (c_{N_j} - \lambda^T a_{N_j})$ são chamados custos relativos (ou custos reduzidos).

Dessa forma, poderemos escrever a equação (2.7) como sendo

$$Z = \hat{Z} + \hat{c}_{N_1}x_{N_1} + \hat{c}_{N_2}x_{N_2} + \cdots + \hat{c}_{N_{n-m}}x_{N_{n-m}}.$$

Com a introdução do conceito de custo relativo, obtém-se todas as ferramentas necessárias para saber se a solução é ótima. Para isso, basta procurar algum custo relativo negativo pois, se este for o caso, tem-se uma solução melhor do que a solução básica. Por outro lado, se nenhum custo relativo for negativo a solução básica será ótima. Isto estabelece um importante critério para obter otimalidade.

Definição 2.7 (Critério de Otimalidade). Dada uma partição básica $A = [\bar{B} \mid N]$, considere a solução básica factível $\hat{X}_{\bar{B}} = \bar{B}^{-1} B \geq 0$ e o vetor multiplicador do *Simplex* $\lambda^T = C_{\bar{B}} \bar{B}^{-1}$. Se $\hat{c}_{N_j} = (c_{N_j} - \lambda^T a_{N_j}) \geq 0$, para todo $j = 1, 2, \dots, n - m$, então a solução básica é ótima.

Esta definição nos mostra claramente que se todos os custos relativos são não-negativos, então a solução básica é ótima, pois ao acrescentar valores positivos à função objetivo obtém-se uma solução pior que a anterior.

De posse do critério de parada para a condição de otimalidade, busca-se o processo para encontrar sempre uma solução factível melhor que a presente; ou seja, caso a solução básica não seja ótima, como determinar uma outra solução básica factível melhor?

Primeiramente, suponha que solução não ótima é não-degenerada, ou seja, não apresenta nenhuma solução básica factível com uma variável básica com valor zero. Assim, existe um custo relativo da variável não-básica x_{N_k} que é negativo, ou seja, existe um k tal que

$$\hat{c}_{N_k} = (c_{N_k} - \lambda^T a_{N_k}) < 0,$$

que irá diminuir o valor da função objetivo e fornecer uma solução factível melhor.

Definição 2.8 (Estratégia *Simplex*). Denomina-se de estratégia *Simplex* a perturbação de uma solução básica factível, que altera as variáveis não-básicas, da seguinte forma

$$\begin{cases} x_{N_k} = \epsilon \leq 0, & \text{(variável com custo relativo negativo)} \\ x_{N_j} = 0, & j = 1, 2, \dots, n - m \text{ e } j \neq k \end{cases}$$

Definição 2.9 (Direção *Simplex*). A direção *Simplex* é o vetor que fornece os coeficientes de como são alteradas as variáveis básicas pela estratégia *Simplex*.

Atente que a Definição 2.8 mostra que a estratégia *Simplex* consiste em pegar uma variável não-básica x_{N_k} , que possui custo relativo negativo, e atribuí-la um valor $\epsilon \geq 0$, diminuindo assim a solução básica factível encontrada anteriormente, ao mesmo tempo que todas as outras variáveis não-básicas x_{N_j} continuam igual a zero. É esta perturbação que produzirá melhor solução de minimização.

Na literatura, denomina-se regra de Dantzig a escolha da variável não-básica que possui o menor custo relativo, pois “quanto menor o valor de \hat{c}_{N_k} , mais rápido a função objetivo decresce”. Desta forma, passa-se a determinar o maior valor possível de ϵ para que a solução se mantenha factível. Esse processo é chamado de Tamanho do Passo Epsilon.

A partir disso, pode-se repetir o processo, pois se a condição de otimalidade não for verificada, busca-se novamente uma nova partição básica, utilizando-se a estratégia *Simplex* e determinada pela direção *Simplex* que melhore a função objetivo. Este processo pode se repetir quantas vezes for necessário até que a condição de otimalidade seja verificada.

Ferramentas Computacionais

O Método *Simplex* basicamente consiste neste procedimento de desenvolvimento das interações sequenciais para otimização em busca da melhor solução. Neste sentido,

existem bibliotecas de linguagem de programação que possuem *softwares* que são capazes de aplicar o Método *Simplex*, ou seja, são ferramentas digitais que já possuem o algoritmo de resolução desse método devidamente implementado. Esse é o caso do *Lindo*, do *Lingo*, e da ferramenta *Solver*, disponível por exemplo, para plataforma *Windows* no *software Microsoft Excel* e, para *Linux*, via *LibreOffice*; dentre essas, e outras possibilidades, nos limitamos a trabalhar com o *Solver* que é um complemento do *Excel*.

Como vimos anteriormente em nosso exemplo motivador, o *Solver* é uma ferramenta de análise hipotética que é capaz de identificar o valor ideal para uma fórmula matemática alocada numa determinada célula, conhecida como célula de destino dentro de uma estrutura de planilha devidamente montada respeitando a formulação específica de um dado Problema de Programação Linear.

3 Teoria do Transporte Ótimo

No presente capítulo, assumiremos a perspectiva da Teoria do Transporte Ótimo. Discutiremos dois tipos de problemas de transporte, o Problema de Monge e o Problema de Kantorovich, realizando um breve levantamento histórico desta teoria. Ao final, apresentamos a Métrica de Wasserstein, um importante conceito geométrico de Teoria de Probabilidade caracterizado a partir do Problema de Kantorovich.

Toda nossa pesquisa estará restrita ao caso da Teoria do Transporte Ótimo entre probabilidades finitas. Os principais resultados desta teoria, podem ser encontrados nas referências (PEYRÉ; CUTURI et al., 2019; THORPE, 2018; VILLANI, 2003; VILLANI, 2009). Além disso, precisaremos de alguns pré-requisitos (Teoria de Probabilidade e Espaços Métricos), os quais serão explorados ao longo do capítulo.

3.1 O Problema de Monge

Gaspard Monge foi um matemático francês que viveu no período de 1746 a 1818. Desde cedo começou se destacando por apresentar uma grande diversidade de técnicas e habilidades como desenhista e inventor. Monge já havia concluído seus estudos de filosofia, física e matemática aos dezesseis anos, quando foi convidado a trabalhar na *Escola Real de Engenharia de Mézières*. Passou vinte anos de sua vida resolvendo um complicado problema de construção de fortificações, inventando um novo método que viria a ser o alicerce da geometria descritiva.

Acredita-se que foi justamente durante a empreitada mencionada anteriormente que Gaspard Monge fez uma primeira abordagem a respeito de um problema que consiste em transportar determinada quantidade de massa de uma região a outra, de tal forma que o custo do trabalho realizado seja mínimo. Em sua versão do problema de transporte, Monge determinou que a massa a ser transportada não pode ser dividida, ou seja, é preciso transportar toda massa de um fornecedor para um único consumidor. Esta formulação vem a ser um caso específico do problema de transporte da Teoria do Transporte Ótimo e, por isso, passou a se chamar de Problema de Monge.

Descrevemos abaixo uma versão discreta e finita do problema de Monge:

- Encontram-se disponíveis em n origens (fornecedores), cada uma com capacidade de fornecimento a_i , para $i = 1, 2, \dots, n$;
- Serão utilizados m destinos (consumidores), que podem absorver as quantidades b_j com $j = 1, 2, \dots, m$;

- Os valores das origens i serão enviados para os destinos j , com um custo de transporte C_{ij} , esgotando-se as disponibilidades de cada origem e satisfazendo-se as necessidades em cada destino;
- Deve-se respeitar o fato de que cada origem (fornecedor) precisa enviar seu produto total a um único destino (consumidor), mas cada destino (consumidor) pode receber produto de mais de uma origem (fornecedor) para completar sua capacidade.

A partir do último item (a condição determinada por Monge), pode-se descrever uma relação de transporte entre as n origens e m destinos dada por uma função

$$T: \{i = 1, 2, 3, \dots, n\} \longrightarrow \{j = 1, 2, \dots, m\}.$$

Para que o problema de transporte de Monge tenha uma solução ótima, primeiramente, é necessário garantir a existência de ao menos uma função T , que realiza o transporte entre as origens e os destinos, condicionada à seguinte relação entre a capacidade do fornecedor (a_i) e a do consumidor (b_j):

$$b_j = \sum_{T(i)=j} a_i \quad \text{para todo } i = 1, 2, \dots, n \text{ e } j = 1, 2, \dots, m. \quad (3.1)$$

Uma função deste tipo recebe a denominação de Mapa de Transporte.

A partir disso, procurar-se minimizar a soma de todos os custos C_{ij} , onde i é uma das origens e j é o destino fornecido por um mapa de transporte T , isto é, $T(i) = j$, levando em consideração todos mapas de transporte T que podem ser encontradas no processo de distribuição. Esta discussão nos remete à formulação matemática dada no Quadro 4.

Quadro 4 – Problema de Monge

<p>Função Objetivo</p> $\text{Minimizar } Z = \sum_{i=0}^n C_{iT(i)}$ <p>Sujeita às Restrições</p> $T: \{i = 1, 2, 3, \dots, n\} \longrightarrow \{j = 1, 2, \dots, m\} \quad \text{tal que}$ $b_j = \sum_{T(i)=j} a_i \quad \text{para todo } i = 1, \dots, n \text{ e } j = 1, \dots, m.$ <p>Onde</p> <p>$\{C_{11}, C_{12}, \dots, C_{21}, C_{22}, \dots, C_{nm}\}$ são custos de transporte;</p> <p>$T: \{i = 1, 2, 3, \dots, n\} \longrightarrow \{j = 1, 2, \dots, m\}$ é um mapa de transporte;</p> <p>$\{a_1, a_2, a_3, \dots, a_n\}$ são as capacidades das origens;</p> <p>$\{b_1, b_2, b_3, \dots, b_m\}$ são as capacidades dos destinos.</p>

Fonte: Elaborado pelo autor.

Observação. O trabalho de Monge estudava uma versão mais complexa de tal problema, onde cada índice i ou j seriam representados por pontos no espaço \mathbb{R}^n , podendo ter uma infinidade de tais pontos, e o custo de transporte considerado era a distância Euclidiana em \mathbb{R}^n . As capacidades seriam funções de \mathbb{R}^n a valores reais e a condição que define o mapa de transporte teria que ser reformulado utilizando integração.

Atente que a condição (3.1) nem sempre se verifica, portanto, este tipo de problema não apresentará uma solução geral para todos os casos.

Exemplo 3.1. Dadas as origens $\{i = 1, 2, 3\}$ e os destinos $\{j = 1, 2\}$, considere as seguintes capacidades do fornecedor (a_i) e a do consumidor (b_j):

$$\begin{array}{lcl} a_1 = 100, & & b_1 = 250, \\ a_2 = 150, & \text{e} & b_2 = 140. \\ a_3 = 170, & & \end{array}$$

É fácil ver que a condição (3.1) não pode ser apresentada para capacidade do fornecedor $j = 2$, isto é,

$$\begin{array}{lcl} a_1 = 100 \neq b_2, & a_1 + a_2 = 250 \neq b_2, & \\ a_2 = 150 \neq b_2, & a_2 + a_3 = 320 \neq b_2, & \text{e} \quad a_1 + a_2 + a_3 = 420 \neq b_2. \\ a_3 = 170 \neq b_2, & a_3 + a_1 = 270 \neq b_2, & \end{array}$$

(Perceba a similaridade deste problema com o exemplo dado na Seção 2.2.)

No século XIX, foi oferecido um prêmio pela solução do problema de Monge, o qual foi faturado por Paul Émile Appell, que se mostrou capaz de apresentar algumas propriedades geométricas acerca do Transporte Ótimo em \mathbb{R}^2 e em \mathbb{R}^3 e ainda conseguiu retificar algumas observações feitas por Monge em sua resenha original. No entanto, resultados de existência de solução para o problema de Monge só apareceram no final do século XX — por meio de Sudakov em 1979, Brenier em 1991, Ambrosio em 2003 e etc — com um tratamento formal do problema de transporte da Teoria do Transporte Ótimo, dado pelo Problema de Kantorovich.

3.2 O Problema de Kantorovich

Leonid Vitaliyevich Kantorovich nasceu em São Petersburgo e desde muito jovem foi reconhecido como um estudante extremamente talentoso. Matriculou na graduação em Matemática da *Universidade Estadual de Leningrado* com quatorze anos, aos vinte já despontava como professor titular do *Instituto de Engenharia de Construção Industrial* desta universidade e três anos depois, concluiu seu doutorado em Ciências Físico-Matemáticas.

Kantorovich contribuiu tanto para o desenvolvimento da Matemática quanto das Ciências Econômicas. No campo do desenvolvimento econômico, Leonid conseguiu introduzir reformas ao design econômico da União Soviética, desenvolvendo flexibilidade na teoria Marxista. Já no campo da Matemática, aprimorou um método linear dos multiplicadores de Lagrange, o qual é bastante semelhante ao utilizado na Programação Linear atualmente. Mas, sua principal contribuição em ambas as áreas foi ter gerado novas ideias em torno do problema de alocação e transporte de recursos, pelas quais recebeu, em 1975, o *Prêmio de Ciências Econômicas em Memória de Alfred Nobel*, em conjunto com Tjalling Koopmans. Além desta honraria, Kantorovich também recebeu o *Prêmio do Estado de Stalin* e o *Prêmio Lenin*, considerados duas das maiores honras para cientistas da *União das Repúblicas Socialistas Soviéticas (URSS)*, e foi eleito postumamente para o *Hall da Fama da Federação Internacional das Sociedades de Pesquisa Operacional* em 2003.

Para que seja possível discutir o Problema de Kantorovich, que consistirá na formulação geral do problema de transporte da Teoria do Transporte Ótimo, precisaremos de alguns pré-requisitos.

Teoria de Probabilidade

Faremos uma abordagem concisa de alguns importantes conceitos da Teoria de Probabilidade. Para uma apresentação mais detalhada das definições abaixo, veja a referência (ROSS, 2009).

Definição 3.2. Dizemos que $p = (p_1, p_2, \dots, p_n)^T \in \mathbb{R}^n$ é um vetor de probabilidade se suas coordenadas p_1, p_2, \dots, p_n satisfazem as condições:

$$(P1) \quad p_i \geq 0, \text{ para todo } i \in \{1, 2, \dots, n\};$$

$$(P2) \quad p_1 + p_2 + \dots + p_n = 1.$$

Para compreendermos melhor a definição de vetor de probabilidade, começaremos ilustrando uma situação-problema, da qual poderemos extrair tais definições.

Exemplo 3.3. Consideremos três caixas denotadas por R_1, R_2 e R_3 . Suponhamos que existam um total de 100 pequenas bolas distribuídas dentro destas caixas. Não estamos interessados em entender o comportamento de cada bola, mas a distribuição delas nas caixas R_1, R_2 e R_3 .

Se temos R_1 com 30 bolas, R_2 com 50 bolas e R_3 com 20 bolas, isto significa que 30% das bolas estão em R_1 , 50% das bolas estão em R_2 e 20% estão em R_3 . Como não temos interesse na quantidade total de bolas, uma forma simples de representar essa distribuição é considerar um vetor coluna com 3 coordenadas indicando os respectivos

percentuais (em notação decimal) de cada região, ou seja,

$$p = \begin{pmatrix} 0,3 \\ 0,5 \\ 0,2 \end{pmatrix}.$$

No caso em que R_1 possui 55 bolas, R_2 com 45 bolas e R_3 nenhuma bola, teríamos

$$q = \begin{pmatrix} 0,55 \\ 0,45 \\ 0 \end{pmatrix}.$$

Com isso, a qualquer distribuição de bolas associa-se um vetor que possui três coordenadas não negativas cuja soma resulta em 1.

Vejamos como representar graficamente os vetores de probabilidade.

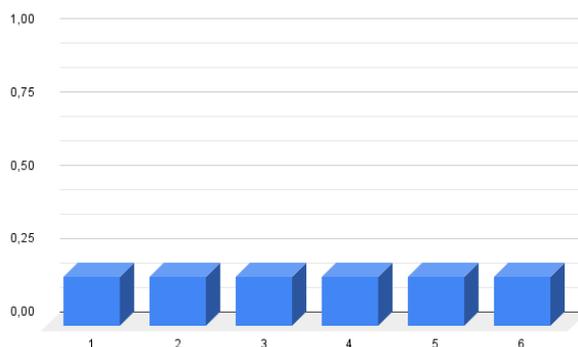
Definição 3.4. O histograma associado a um vetor de probabilidade $p = (p_1, p_2, \dots, p_n)^T$, consiste no gráfico de uma função cujo domínio é dado pelos índices do vetor $\{1, 2, 3, \dots, n\}$ e os valores de cada coordenada p_i fornecem a frequência (ou altura da barra) associada ao índice i .

Exemplo 3.5 (Lançamento de um dado). Sabemos que a chance de cair um certo número no lançamento de um dado (honesto) é de 1 em 6, ou seja, cada face deste dado tem probabilidade $\frac{1}{6}$ de aparecer. Representamos a distribuição da probabilidade de cada face deste dado por um vetor de probabilidade de coluna com 6 coordenadas

$$p = \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6} \right)^T,$$

onde cada entrada de índice i fornece a probabilidade de cair a face i virada para cima. Segundo a Definição 3.4, seu respectivo histograma é apresentado na Figura 6.

Figura 6 – Histograma do Lançamento de um Dado Honesto p



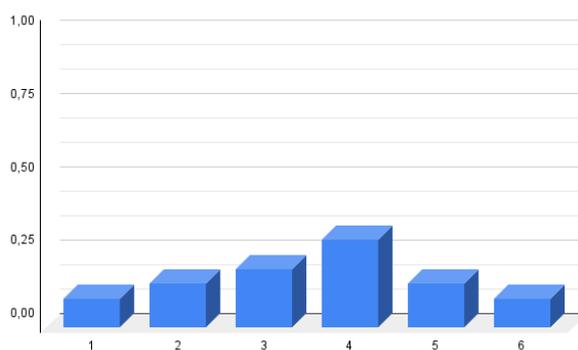
Fonte: Elaborada pelo autor.

Analogamente, se o dado for viciado, o lançamento pode dar preferência para uma ou mais faces fornecendo outros vetores de probabilidades (ou de frequências) e respectivos histogramas:

$$q = (0, 1; 0, 15; 0, 2; 0, 3; 0, 15; 0, 1)^T \quad \text{e} \quad r = (0, 1, 0, 0, 0, 0)^T.$$

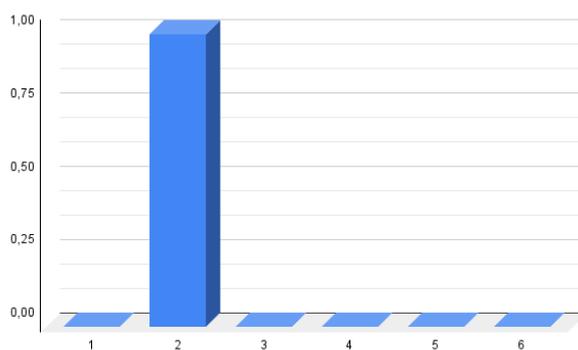
As Figuras 7 e 8 representam os histogramas das distribuições q e r , respectivamente.

Figura 7 – Histograma do Lançamento de um Dado Viciado q



Fonte: Elaborada pelo autor.

Figura 8 – Histograma do Lançamento de um Dado Viciado r



Fonte: Elaborada pelo autor.

O conjunto de todas as possíveis distribuições de probabilidades associadas a um experimento aleatório ou situação-problema é um importante subconjunto de \mathbb{R}^n . Vamos denotar por \mathcal{P}_n o conjunto formado por todos os vetores de probabilidade de tamanho n .

Veremos na próxima subseção que a formulação de Kantorovich para o problema de transporte, utilizará vetores de probabilidade com o intuito de realizar um “transporte” entre suas coordenadas. Antes, vamos definir a matriz de plano de transporte.

Definição 3.6 (Plano de Transporte). Dados vetores de probabilidade

$$a = (a_1, a_2, \dots, a_n)^\top \in \mathcal{P}^n \quad \text{e} \quad b = (b_1, b_2, \dots, b_m)^\top \in \mathcal{P}^m,$$

denominamos uma matriz

$$X = \begin{pmatrix} X_{11} & \cdots & X_{1m} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{nm} \end{pmatrix} \in \mathbb{R}_+^{n \times m}$$

de plano de transporte, se obedece as condições:

(T1) A soma em cada linha da matriz X fornece a respectiva entrada do vetor de probabilidade a , ou seja, $\sum_{j=1}^m X_{ij} = a_i$ para todo $i = 1, 2, \dots, n$;

(T2) A soma em cada coluna da matriz X fornece a respectiva entrada do vetor de probabilidade b , isto é, $\sum_{i=1}^n X_{ij} = b_j$ para todo $j = 1, 2, \dots, m$.

Exemplo 3.7 (Plano de Transporte I). Dadas as probabilidade

$$a = \begin{pmatrix} 0, 24 \\ 0, 36 \\ 0, 40 \end{pmatrix} \in \mathcal{P}_3 \quad \text{e} \quad b = \begin{pmatrix} 0, 64 \\ 0, 36 \end{pmatrix} \in \mathcal{P}_2$$

vamos verificar que a matriz

$$X = \begin{pmatrix} 0, 24 & 0 \\ 0 & 0, 36 \\ 0, 40 & 0 \end{pmatrix}$$

é um plano de transporte de a para b . De fato,

$$(T1) : \begin{cases} 0, 24 + 0 = 0, 24 = a_1 \\ 0 + 0, 36 = 0, 36 = a_2 \\ 0, 40 + 0 = 0, 40 = a_3 \end{cases} \quad \text{e} \quad (T2) : \begin{cases} 0, 24 + 0 + 0, 40 = 0, 64 = b_1 \\ 0 + 0, 36 + 0, 40 = 0, 36 = b_2 \end{cases}.$$

Analogamente, podemos identificar (T1) e (T2) no seguinte caso:

$$\bar{X} = \begin{pmatrix} 0, 16 & 0, 08 \\ 0, 36 & 0 \\ 0, 12 & 0, 28 \end{pmatrix} \quad \begin{pmatrix} 0, 24 \\ 0, 36 \\ 0, 40 \end{pmatrix} = a$$

$$b = (0, 64 \quad 0, 36)^\top$$

Donde,

$$(T1) : \begin{cases} 0, 16 + 0, 08 = 0, 24 = a_1 \\ 0, 36 + 0 = 0, 36 = a_2 \\ 0, 12 + 0, 28 = 0, 40 = a_3 \end{cases} \quad \text{e} \quad (T2) : \begin{cases} 0, 16 + 0, 36 + 0, 12 = 0, 64 = b_1 \\ 0, 08 + 0 + 0, 28 = 0, 36 = b_2 \end{cases}.$$

Também é possível reescrever os sistemas (T1) e (T2) algebricamente, para isso, considere o vetor coluna

$$1_n = \left. \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right\} n \text{ cópias,}$$

e observe que:

$$X 1_n = \begin{pmatrix} X_{11} & \cdots & X_{1m} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{nm} \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} X_{11} + \cdots + X_{1m} \\ \vdots \\ X_{n1} + \cdots + X_{nm} \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_2 \end{pmatrix} = a.$$

$$X^T 1_m = \begin{pmatrix} X_{11} & \cdots & X_{1n} \\ \vdots & \ddots & \vdots \\ X_{m1} & \cdots & X_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} X_{11} + \cdots + X_{1n} \\ \vdots \\ X_{m1} + \cdots + X_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_2 \end{pmatrix} = b.$$

Vejamos outro exemplo de plano de transporte sob este ponto de vista algébrico.

Exemplo 3.8 (Plano de Transporte II). A matriz

$$X = \begin{pmatrix} 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{20} & \frac{1}{4} & 0 & \frac{1}{5} \end{pmatrix}$$

é um plano de transporte entre os vetores de probabilidade $a = \left(\frac{1}{2}, 0, \frac{1}{2}\right)^T$ e $b = \left(0, \frac{3}{10}, \frac{1}{2}, 0, \frac{1}{5}\right)^T$. De fato,

$$X 1_n = \begin{pmatrix} \frac{1}{4} + \frac{1}{4} \\ 0 \\ \frac{1}{20} + \frac{1}{4} + \frac{1}{5} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix} = a \quad \text{e} \quad X^T 1_m = \begin{pmatrix} 0 \\ \frac{1}{4} + \frac{1}{20} \\ \frac{1}{4} + \frac{1}{4} \\ 0 \\ \frac{1}{5} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{3}{10} \\ \frac{1}{2} \\ 0 \\ \frac{1}{5} \end{pmatrix} = b.$$

Apresentamos uma última definição e um ultimo resultado que serão fundamentais para apresentar o problema de transporte da Teoria do Transporte Ótimo.

Definição 3.9. Definimos o conjunto de todos planos de transporte entre os vetores de probabilidade $a \in \mathcal{P}_n$ e $b \in \mathcal{P}_m$ como

$$U(a, b) = \left\{ X \in \mathbb{R}_+^{n \times m} : \sum_{j=1}^m X_{ij} = a_i \quad \text{e} \quad \sum_{i=1}^n X_{ij} = b_j \right\}$$

$$= \left\{ X \in \mathbb{R}_+^{n \times m} : X 1_n = a \quad \text{e} \quad X^T 1_m = b \right\}.$$

Proposição 3.10. *Existe ao menos um plano de transporte entre os vetores de probabilidade $a \in \mathcal{P}_n$ e $b \in \mathcal{P}_m$ (ou seja, $U(a, b) \neq \emptyset$).*

Demonstração. A partir dos vetores

$$a = (a_1, a_2, \dots, a_n)^\top \in \mathcal{P}^n \quad \text{e} \quad b = (b_1, b_2, \dots, b_m)^\top \in \mathcal{P}^m,$$

defina a matriz $n \times m$

$$X = \begin{pmatrix} a_1 b_1 & a_1 b_2 & \cdots & a_1 b_m \\ a_2 b_1 & a_2 b_2 & \cdots & a_2 b_m \\ \vdots & \vdots & \ddots & \vdots \\ a_n b_1 & a_n b_2 & \cdots & a_n b_m \end{pmatrix}.$$

Para garantir que $X = (a_i b_j)_{ij} \in U(a, b)$, basta verificar que tal matriz satisfaz as condições

$$(T1) : \sum_{j=1}^m X_{ij} = \sum_{j=1}^m a_i b_j = a_i \underbrace{\left(\sum_{j=1}^m b_j \right)}_{=1} = a_i$$

$$(T2) : \sum_{i=1}^n X_{ij} = \sum_{i=1}^n a_i b_j = b_j \underbrace{\left(\sum_{i=1}^n a_i \right)}_{=1} = b_j$$

onde as últimas igualdades seguem do fato de que a e b são vetores de probabilidade. \square

Observação. Nos trabalhos originais de Kantorovich, define-se a classe de planos de transporte como sendo o conjunto de todas as medidas de probabilidades sobre o respectivo espaço produto e com distribuições marginais fixadas que representam, respectivamente, as distribuições de massa inicial e final.

Problema de Transporte da Teoria do Transporte Ótimo

Em 1942, Kantorovich, mesmo sem ter conhecimento do trabalho de Monge acerca do problema de transporte, escreveu um artigo no qual apresentou uma nova formulação que se mostrou inovadora e suscetível de solução, quando comparada com o Problema de Monge. A formulação de Kantorovich substituiu o problema de obter um mapa de transporte pelo problema de encontrar um plano de transporte, e assim minimizar o custo total. O grande diferencial foi permitir que a massa a ser transportada de uma origem possa ser fragmentada entre os diversos destinos disponíveis, o que gerou maiores possibilidades de distribuição da massa total.

Descrevemos abaixo uma versão discreta e finita da proposta de Kantorovich:

- Encontram-se disponíveis em n origens, com capacidade a_i , para $i = 1, 2, \dots, n$;

- Serão utilizados m destinos, que podem absorver as quantidades b_j com $j = 1, 2, \dots, m$;
- Assumimos que as capacidades das origens e dos destinos formam vetores de probabilidade (a_1, a_2, \dots, a_n) e (b_1, b_2, \dots, b_m) ;
- Os valores das origens i podem ser enviados para os destinos j , com um custo de transporte (por unidade) C_{ij} ;
- Será permitido particionar a capacidade de cada origem i e transportá-los para os destinos j , por meio de um plano de transporte X_{ij} , esgotando-se as disponibilidades de cada origem e satisfazendo-se as necessidades em cada destino.

Os dois últimos itens serão reescritos em termos matemáticos, isto é

$$\sum_{j=1}^m X_{ij} = a_i \quad \text{e} \quad \sum_{i=1}^n X_{ij} = b_j.$$

Já o terceiro item acima impõe as condições

$$\sum_{i=1}^n a_i = 1 \quad \text{e} \quad \sum_{j=1}^m b_j = 1.$$

O Problema de Kantorovich consiste basicamente em determinar o menor custo da distribuição de um determinado produto, de modo a minimizar o somatório de todos os custos de transporte por quantidades transportadas $C_{ij} X_{ij}$, levando em consideração todos os possíveis planos de transportes $\Pi_{ij} \in U(a, b)$ envolvidos no processo de distribuição. Dessa forma, apresentamos no Quadro 5 a formulação geral do Problema de Kantorovich.

Quadro 5 – Problema de Kantorovich

Função Objetivo

$$\text{Minimizar } Z = \sum_{i=1}^n \sum_{j=1}^m C_{ij} X_{ij}$$

Sujeita às Restrições

$$X \in U(a, b) = \left\{ X \in \mathbb{R}_+^{n \times m} : \sum_{j=1}^m X_{ij} = a_i \quad \text{e} \quad \sum_{i=1}^n X_{ij} = b_j \right\}$$

$$a \in \mathcal{P}_n \quad \left(\text{ou seja, } \sum_{i=1}^n a_i = 1 \right) \quad \text{e} \quad b \in \mathcal{P}_m \quad \left(\text{isto é, } \sum_{j=1}^m b_j = 1 \right).$$

Onde

C denota a matriz de custos de transporte cujas entradas C_{ij} são os custo unitário de transporte da origem i até cada o destino j ;

X denota um plano de transporte, cujas entradas X_{ij} são as quantidades a ser transportada de cada origem i para o destino j ;

$a = (a_1, a_2, \dots, a_n)^\top$ denota o vetor de probabilidade, com coordenadas a_i que são as capacidades das origens;

$b = (b_1, b_2, \dots, b_m)^\top$ denota o vetor de probabilidade, com coordenadas b_j que são as capacidades dos destinos.

Fonte: Elaborado pelo autor.

De forma concisa, podemos reduzir a uma única expressão

$$Z = \min_{X \in U(a,b)} \sum_{i=1}^n \sum_{j=1}^m C_{ij} X_{ij}.$$

A existência de um plano de transporte ótimo é verificada de forma mais simples no formalismo de Kantorovich.

Proposição 3.11. *Sempre existe plano de transporte $\Pi \in U(a, b)$ tal que*

$$Z = \min_{X \in U(a,b)} \sum_{i=1}^n \sum_{j=1}^m C_{ij} X_{ij} = \sum_{i=1}^n \sum_{j=1}^m C_{ij} \Pi_{ij}.$$

Demonstração. Utilizaremos um resultado clássico do cálculo em várias variáveis (LIMA, 2020, Capítulo 8, Proposição 4): toda função (a valores reais) contínua definida em um conjunto fechado e limitado (compacto) possui pontos de máximo e mínimo.

Primeiramente, estamos investigando a função (a valores reais)

$$\begin{aligned} \Phi: \quad U(a, b) &\longrightarrow \mathbb{R} \\ X = (X_{ij}) &\longmapsto \sum_{i=1}^n \sum_{j=1}^m C_{ij} X_{ij}, \end{aligned}$$

Por ser uma função linear nas variáveis X_{ij} , ela é uma função contínua.

Resta argumentar que o conjunto $U(a, b) \subset \mathbb{R}^{n \times m}$ é fechado e limitado (para a definição de tais noções, consulte (LIMA, 2020)). Perceba que as variáveis X_{ij} estão restritas ao intervalo $[0, 1]$, logo $U(a, b)$ é um subconjunto do cubo $n \times m$ -dimensional $[0, 1]^{n \times m}$, o qual é um conjunto limitado. Para mostrar a propriedade fechada, considere as seguintes funções

$$\begin{aligned} \alpha_i: U(a, b) &\longrightarrow \mathbb{R} & \beta_j: U(a, b) &\longrightarrow \mathbb{R} \\ X &\longmapsto \sum_{j=1}^m X_{ij} & X &\longmapsto \sum_{i=1}^n X_{ij} \end{aligned} \quad \text{e}$$

para todo $1 \leq i \leq n$ e $1 \leq j \leq m$. Note que as restrições que definem as matrizes em $U(a, b)$ podem ser escritas através dos conjuntos $\alpha_i^{-1}(a_i)$, para $1 \leq i \leq n$, e $\beta_j^{-1}(b_j)$, para $1 \leq j \leq m$. Em particular, por serem funções (lineares) contínuas, todos estes conjuntos são fechados. Por construção, temos que

$$U(a, b) = \left(\bigcap_{i=1}^n \alpha_i^{-1}(a_i) \right) \cap \left(\bigcap_{j=1}^m \beta_j^{-1}(b_j) \right).$$

Como as interseções de conjuntos fechados também são conjuntos fechados, garantimos que $U(a, b)$ é fechado, donde segue o resultado. \square

Observação. Outras forma equivalente de apresentar o problema de transporte da Teoria de Transporte Ótimo, é conhecida como Dualidade de Kantorovich. Trata-se de um resultado clássico que pode ser encontrado nas referências (VILLANI, 2003; THORPE, 2018; PEYRÉ; CUTURI et al., 2019).

3.3 Elo entre a Teoria de Transporte Ótimo e a Programação Linear

Para compreender a conexão entre as formulações do problema de transporte na Teoria de Transporte Ótimo (Quadro 5) e na Programação Linear (Quadro 2 e Quadro 3), necessitamos traduzir:

- a função objeto $Z = \sum_{i=1}^n \sum_{j=1}^m C_{ij} X_{ij}$ em um produto de um vetor linha por um vetor coluna;
- e as restrições $\sum_{j=1}^m X_{ij} = a_i$ e $\sum_{i=1}^n X_{ij} = b_j$, num sistema linear dado por.

Como teremos que transformar matrizes em vetores, introduzimos o procedimento de vetorização de matrizes

Definição 3.12. (Vetorização de Matrizes) A partir de uma matriz $n \times m$

$$M = \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix}$$

definiremos um vetor coluna $nm \times 1$, empilhando as colunas da matriz M

$$\text{vet}(M) = \begin{pmatrix} x_{1,1} \\ \vdots \\ x_{n,1} \\ x_{1,2} \\ \vdots \\ x_{n,2} \\ \vdots \\ x_{1,m} \\ \vdots \\ x_{n,m} \end{pmatrix}.$$

Para a função objetivo,

$$\begin{aligned} \text{vet}(C)^T \text{vet}(X) &= (C_{11}, \dots, C_{n1} \mid C_{12}, \dots, C_{n2} \mid \cdots \mid C_{1m}, \dots, C_{nm}) \begin{pmatrix} X_{11} \\ \vdots \\ X_{n1} \\ X_{12} \\ \vdots \\ X_{n2} \\ \vdots \\ X_{1m} \\ \vdots \\ X_{nm} \end{pmatrix} \\ &= \sum_{i=1}^n \sum_{j=1}^m C_{ij} X_{ij} = Z. \end{aligned}$$

Por outro lado, para traduzir as restrições, utilizaremos as seguintes notações para a matriz identidade e para o vetor linha com todas as entradas 1 em dimensão n , respectivamente

$$Id_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \quad \text{e} \quad \hat{1}_n = \underbrace{(1, 1, \dots, 1)}_{n \text{ cópias}}.$$

A restrição do tipo $\sum_{j=1}^m X_{ij} = a_i$ realizará a soma de coordenadas de $\text{vet}(X)$ com espaçamento de n . Temos então que

$$\underbrace{\left(\begin{array}{cccc|c|c|c} 1 & 0 & \cdots & 0 & & & \\ 0 & 1 & \cdots & 0 & & & \\ \vdots & \vdots & \ddots & \vdots & & & \\ 0 & 0 & \cdots & 1 & & & \end{array} \right)}_{m \text{ cópias}} \begin{array}{c} \left(\begin{array}{c} X_{11} \\ \vdots \\ X_{n1} \\ X_{12} \\ \vdots \\ X_{n2} \\ \vdots \\ X_{1m} \\ \vdots \\ X_{nm} \end{array} \right) \end{array} = \begin{array}{c} \left(\begin{array}{c} X_{11} + X_{12} + \cdots + X_{1m} \\ X_{21} + X_{22} + \cdots + X_{2m} \\ \vdots \\ X_{n1} + X_{n2} + \cdots + X_{nm} \end{array} \right) \end{array} = \begin{array}{c} \left(\begin{array}{c} a_1 \\ a_2 \\ \vdots \\ a_n \end{array} \right) \end{array} = a.$$

Já a restrição do tipo $\sum_{i=1}^n X_{ij} = b_j$, a soma as m coordenadas de $\text{vet}(X)$ em sequência, donde utilizaremos o vetor $\hat{1}_n$, isto é,

$$\underbrace{\left(\begin{array}{ccc|c|c|c} 1 & \cdots & 1 & 0 & \cdots & 0 \\ \hline 0 & & & \hat{1}_n & \cdots & 0 \\ \hline \vdots & & & 0 & \ddots & \vdots \\ \hline 0 & & & 0 & \cdots & \hat{1}_n \end{array} \right)}_{m \times m \text{ subdivisões}} \begin{array}{c} \left(\begin{array}{c} X_{11} \\ \vdots \\ X_{n1} \\ X_{12} \\ \vdots \\ X_{n2} \\ \vdots \\ X_{1m} \\ \vdots \\ X_{nm} \end{array} \right) \end{array} = \begin{array}{c} \left(\begin{array}{c} X_{11} + X_{21} + \cdots + X_{n1} \\ X_{12} + X_{22} + \cdots + X_{n2} \\ \vdots \\ X_{1m} + X_{2m} + \cdots + X_{nm} \end{array} \right) \end{array} = \begin{array}{c} \left(\begin{array}{c} b_1 \\ b_2 \\ \vdots \\ b_m \end{array} \right) \end{array} = b$$

Podemos agora rerepresenta o problema de Kantorovich em termos do seguinte problema de Programação Linear dado no Quadro 6.

Quadro 6 – Problema de Kantorovich em Termos da Programação Linear

Função Objetivo

$$\text{Minimizar } Z = \text{vet}(C)^\top \text{vet}(X)$$

Sujeita às Restrições

$$A \text{vet}(X) = B, \quad \text{vet}(X) \geq 0 \quad \text{e} \quad B \geq 0$$

Onde

$$\text{vet}(C)^\top = (C_{11}, \dots, C_{n1} \mid C_{12}, \dots, C_{n2} \mid \dots \mid C_{1m}, \dots, C_{nm}),$$

$$\text{vet}(X) = \begin{pmatrix} X_{11} \\ \vdots \\ X_{n1} \\ \hline X_{12} \\ \vdots \\ X_{n2} \\ \vdots \\ \hline X_{1m} \\ \vdots \\ X_{nm} \end{pmatrix}, \quad A = \begin{pmatrix} & Id_n & & Id_n & \cdots & Id_n \\ 1 & \cdots & 1 & 0 & \cdots & 0 \\ \hline & & & \hat{1}_n & \cdots & 0 \\ & & & 0 & \ddots & \vdots \\ \hline & & & 0 & & \hat{1}_n \end{pmatrix}$$

$$\text{e} \quad B = \begin{bmatrix} a \\ b \end{bmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \\ \hline b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Fonte: Elaborado pelo autor.

3.4 Distância de Wasserstein

Concluída a parte de apresentação do problema de transporte da Teoria do Transporte Ótimo, iniciaremos uma discussão sobre as características geométricas e intrínsecas do Problema de Kantorovich. Para que seja possível introduzir a Métrica de Wasserstein, que consistirá na ferramenta que fornece a estrutura geométrica a ser investigada, precisaremos de mais alguns pré-requisitos.

Espaço Métrico

Veremos nesta subseção uma noção bastante simples acerca do conceito de uma distância. Para mais detalhes sobre este tópico, consulte a referência (LIMA, 2020; DEZA;

DEZA, 2009).

Intuitivamente uma distância representa a medida do espaço que separa dois objetos. Uma métrica é um conceito que generaliza essa ideia geométrica de distância e se mostra uma ferramenta de muita importância em vários campos da ciência e tecnologia, tais como: Geometria, Teoria dos Grafos e Redes, Computação Gráfica, Probabilidade, Estatística, Astronomia, Biologia, Física, dentre outros.

Definição 3.13. Uma métrica (ou distância) sobre um conjunto não vazio M qualquer pode ser entendida como uma função $d: M \times M \rightarrow \mathbb{R}_+$, que deve satisfazer às seguintes propriedades:

- (M1) Ser simétrica, ou seja, $d(x, y) = d(y, x)$ para todo $x, y \in M$;
- (M2) Ser nula para pontos coincidentes, isto é, $d(x, y) = 0$ se e somente se $x = y$;
- (M3) Atender à desigualdade triangular, dada por

$$d(x, z) \leq d(x, y) + d(y, z) \quad \text{para todo } x, y, z \in M.$$

Essas propriedades são facilmente verificadas nos seguintes casos.

Exemplo 3.14 (Reta Real \mathbb{R}). A métrica usual sobre o a reta real \mathbb{R} é simplesmente $d: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}_+$ definida por

$$d(x, y) = |x - y|,$$

isto é, o valor absoluto da diferença entre dois números.

Para estes casos particulares, a desigualdade triangular pode ser interpretada, de maneira intuitiva, através do fato que o comprimento de qualquer um dos lados de um triângulo não pode ser maior que a soma dos outros dois lados.

Ter um conjunto munido de uma função que atende às condições mostradas na definição anterior, ou seja, às propriedades de uma métrica, nos garante uma estrutura geométrica sobre este conjunto cujas características podem ser exploradas.

Definição 3.15. Definiremos um espaço métrico como sendo um par ordenado (M, d) no qual M é um conjunto não vazio e d é uma métrica em M .

Vamos ver alguns exemplos já conhecidos de espaços métricos.

Exemplo 3.16 (Plano \mathbb{R}^2). Sejam $x = (x_1, x_2)$ e $y = (y_1, y_2)$ elementos de \mathbb{R}^2 . Definiremos a métrica $d: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}_+$ como sendo

$$d(x, y) = \|x - y\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}.$$

O Plano Euclidiano (\mathbb{R}^2, d) é um espaço métrico.

Exemplo 3.17 (Espaço Euclidiano n -dimensional \mathbb{R}^n). Sejam $x = (x_1, x_2, x_3, x_4, \dots, x_n)$ e $y = (y_1, y_2, y_3, y_4, \dots, y_n)$ elementos de \mathbb{R}^n . Consideramos o espaço métrico Euclidiano n -dimensional (\mathbb{R}^n, d) , cuja a métrica $d: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$ é

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$

(Vale salientar que esta não é a única métrica definida em \mathbb{R}^n , veja (LIMA, 2020).)

Veremos abaixo uma construção mais abstrata de espaço métrico.

Exemplo 3.18 (Conjunto de três elementos). Considere um conjunto M formado por três objetos aleatórios, por exemplo, uma cadeira, uma mesa e um sofá, denotados por c , m e s , respectivamente. Definimos a seguinte função $d: M \times M \rightarrow \mathbb{R}_+$ tal que:

$$d(c, c) = 0, \quad d(m, m) = 0, \quad d(s, s) = 0,$$

$$d(c, m) = 1, \quad d(m, c) = 1, \quad d(m, s) = 1, \quad d(s, m) = 1, \quad d(s, c) = 1 \quad \text{e} \quad d(c, s) = 1.$$

O Lema 3.19 mostra que $M = \{c, m, s\}$ é um espaço métrico.

Lema 3.19. *O conjunto $M = \{c, m, s\}$ munido da função $d: M \times M \rightarrow \mathbb{R}_+$ dada por*

$$d(x, y) = \begin{cases} 0, & \text{se } x = y \\ 1, & \text{se } x \neq y \end{cases}$$

é um espaço métrico.

Demonstração. Temos que verificar a validade das propriedades da Definição 3.13. Através da construção do exemplo, fica notório a verificação da positividade da função d , pois segue do próprio enunciado que $d(x, y) = 0$ ou $d(x, y) = 1 > 0$ para quaisquer que sejam $x, y \in M$.

A partir do enunciado, temos

$$d(x, y) = 1 = d(y, x) \quad \text{e} \quad d(x, x) = 0$$

Segue que, para quaisquer que sejam $x, y \in M$, teremos $d(x, y) = d(y, x)$ e, assim, verificamos a propriedade (M1). Da mesma forma, temos que $d(x, x) = 0$ para qualquer $x \in M$. Logo $d(x, y) = 0$ se, e somente se, $x = y$, donde verificamos a propriedade (M2).

Para mostrarmos a propriedade (M3), que é a desigualdade triangular, iremos dividir a análise em dois casos:

- Caso 1: Se $x = z$, então $d(x, z) = 0$. Logo $0 = d(x, z) \leq d(x, y) + d(y, z)$, pois 0 é menor ou igual que a soma de quaisquer números não negativos, inclusive o próprio 0;

- Caso 2: Se $x \neq z$, então $d(x, z) = 1$. Observe que sem dúvidas teremos ou $x \neq y$ ou $z \neq y$, visto que, se $x = y$ e $y = z$ então $x = z$, contrariando a hipótese. Dessa forma, $d(x, z) = 1 \leq d(x, y) + d(y, z)$.

□

Observação. De maneira geral, qualquer conjunto finito X pode se tornar um espaço métrico usando a construção demonstrada acima. Para tanto, basta que definamos a *métrica zero-um* que satisfaz $d(x, x) = 0$ para qualquer $x \in X$ e $d(x, y) = 1$ para quaisquer que sejam $x, y \in X$, com $x \neq y$.

Exemplo 3.20 (Distância entre Índices). O conjunto finito dos n primeiros números naturais $\{1, 2, 3, \dots, n\}$ se torna um espaço métrico se o munirmos da métrica $d: \{1, 2, \dots, n\} \times \{1, 2, \dots, n\} \rightarrow \mathbb{R}_+$ tal que

$$d(i, j) = |i - j|.$$

Trata-se da restrição da distância Euclidiana em \mathbb{R} a um subconjunto próprio, logo satisfaz todas as propriedades da Definição 3.13.

Propriedades Métricas do Problema de Transporte

Nesta seção, introduziremos mais um exemplo de espaço métrico, cujo conjunto base será o espaço dos vetores de probabilidade \mathcal{P}_n . A função distância estará intimamente relacionada ao problema de transporte da Teoria do Transporte Ótimo e, conseqüentemente, a Programação Linear.

Teorema 3.21. *Considere C uma matriz de custos $n \times n$ que satisfaz:*

(C1) *A matriz C é simétrica, ou seja, $C_{ij} = C_{ji}$;*

(C2) *$C_{ij} = 0$ se, e somente se, $i = j$;*

(C3) *$C_{ik} \leq C_{ij} + C_{jk}$ para quaisquer índices i, j, k .*

Então

$$W(a, b) = \min_{X \in U(a, b)} \sum_{i=1}^n \sum_{j=1}^n C_{ij} X_{ij}$$

é uma métrica, denominada de distância (ou métrica) de Wasserstein, e o espaço dos vetores de probabilidade munido desta distância, (\mathcal{P}_n, W) , é um espaço métrico.

Perceba que as condições sobre a matriz de custos do problema de transporte se assemelham com as propriedades (M1), (M2) e (M3) de uma métrica (da Definição 3.13). Em particular, podemos afirmar que $d(i, j) = C_{ij}$ determina uma distância sobre o conjunto de índices $\{1, 2, \dots, n\}$.

De forma mais geral, o resultado acima estabelece que o valor ótimo do problema de transporte herdará as propriedades de distância de sua matriz de custos, quando esta determina uma métrica entre as localizações investigadas no problema de transporte, ou seja, as origens e os destinos. A demonstração desenvolvida abaixo detalha as linhas gerais da prova apresentada em (PEYRÉ; CUTURI et al., 2019, Proposição 2.2).

Demonstração. Partindo de (C1), (C2), (C3) e das particularidades do problema de Kantorovich, queremos provar as propriedades (M1), (M2) e (M3) da Definição 3.13 para a métrica de Wasserstein $W(a, b)$.

Note que, a positividade de W é consequência dos custos e planos de transportes sempre serem valores não negativos,

$$C_{ij} \geq 0 \quad \text{e} \quad X_{ij} \geq 0 \quad \implies \quad W(a, b) = \min_{X \in U(a, b)} \sum_{i=1}^n \sum_{j=1}^n C_{ij} X_{ij} \geq 0.$$

Verificaremos a primeira propriedade (M1), a simetria. Observe a reescritura

$$\sum_{i=1}^n \sum_{j=1}^n C_{ij} X_{ij} \stackrel{(C1)}{=} \sum_{j=1}^n \sum_{i=1}^n C_{ji} X_{ji}^T.$$

Sabemos que sempre é possível selecionar um plano de transporte ótimo $F \in U(a, b)$, isto é, que atinge o valor mínimo do problema de transporte de a para b , logo

$$W(a, b) = \sum_{i=1}^n \sum_{j=1}^n C_{ij} F_{ij}.$$

Deste plano de transporte, obtemos outro plano F^T que transporta b até a . Veja que

$$W(a, b) = \sum_{i=1}^n \sum_{j=1}^n C_{ij} F_{ij} \stackrel{(C1)}{=} \sum_{j=1}^n \sum_{i=1}^n C_{ji} F_{ji}^T \geq \min_{X \in U(b, a)} \sum_{j=1}^n \sum_{i=1}^n C_{ji} X_{ji}^T = W(b, a).$$

Por outro lado, selecionamos o plano de transporte ótimo $G \in U(b, a)$, tal que

$$W(b, a) = \sum_{j=1}^n \sum_{i=1}^n C_{ji} G_{ji},$$

e observamos que

$$W(b, a) = \sum_{j=1}^n \sum_{i=1}^n C_{ji} G_{ji} \stackrel{(C1)}{=} \sum_{j=1}^n \sum_{i=1}^n C_{ij} G_{ij}^T \geq \min_{X \in U(b, a)} \sum_{i=1}^n \sum_{j=1}^n C_{ij} X_{ij} = W(a, b).$$

Portanto, dois valores são menores e iguais entre si sempre que $W(a, b) = W(b, a)$ e, assim, a distância de Wasserstein é simétrica.

Para obter a segunda propriedade (M2), vamos verificar a nulidade para pontos coincidentes.

- $a = b \implies W(a, b) = W(a, a) = 0$:

Primeiramente, tomamos

$$A = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix} \in U(a, a)$$

e calculamos

$$\sum_{i=1}^n \sum_{j=1}^n C_{ij} A_{ij} = \sum_{i=1}^n C_{ii} a_i \stackrel{(C2)}{=} \sum_{i=1}^n \sum_{j=1}^n 0 \cdot a_i = 0.$$

Como $W(\cdot, \cdot) \geq 0$, temos que

$$0 \leq W(a, a) = \min_{X \in U(a, a)} \sum_{i=1}^n \sum_{j=1}^n C_{ij} X_{ij} \leq \sum_{i=1}^n \sum_{j=1}^n C_{ij} A_{ij} = 0$$

Logo, a matriz A é um plano de transporte ótimo e $W(a, a) = 0$.

- $W(a, b) = 0 \implies a = b$:

Suponha que existe um plano de transporte ótimo $\Pi \in U(a, b)$, tal que

$$\sum_{i=1}^n \sum_{j=1}^n C_{ij} \Pi_{ij} = W(a, b) = 0.$$

Como $C_{ij} \geq 0$ e $\Pi_{ij} \geq 0$, temos que $C_{ij} \Pi_{ij} = 0$ para todo $1 \leq i, j \leq n$. Segue daí que

Para $i = j$, teremos $C_{ij} = C_{ii} \stackrel{(C2)}{=} 0$;

Para $i \neq j$, teremos $C_{ij} \stackrel{(C2)}{\neq} 0$, logo $\Pi_{ij} = 0$.

Mas recorde que Π_{ij} é um plano de transporte, que verifica as condições

$$a_i = \sum_{j=1}^n \Pi_{ij} = \Pi_{ii} \quad \text{e} \quad b_j = \sum_{i=1}^n \Pi_{ij} = \Pi_{jj}$$

para todo $1 \leq i, j \leq n$, portanto

$$\Pi = \begin{pmatrix} \Pi_{11} & 0 & \cdots & 0 \\ 0 & \Pi_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Pi_{nn} \end{pmatrix} = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix} = \begin{pmatrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_n \end{pmatrix}.$$

Em particular, temos $a_k = b_k$, ou seja, $a = b$

Para obter a desigualdade triangular (M3),

$$W(a, c) \leq W(a, b) + W(b, c),$$

consideremos:

- P um plano de transporte ótimo entre a e b , isto é, $W(a, b) = \sum_{i=1}^n \sum_{k=1}^n C_{ik} P_{ik}$ e

$$P \in U(a, b) \iff \begin{cases} P \mathbf{1}_n = a \\ P^\top \mathbf{1}_n = b \end{cases} \quad \text{ou} \quad \begin{cases} \sum_{j=1}^n P_{kj} = a_k \\ \sum_{i=1}^n P_{ik} = b_k \end{cases}.$$

- Q um plano de transporte ótimo entre b e c , ou seja, $W(b, c) = \sum_{k=1}^n \sum_{j=1}^n C_{kj} Q_{kj}$ e

$$Q \in U(b, c) \iff \begin{cases} Q \mathbf{1}_n = b \\ Q^\top \mathbf{1}_n = c \end{cases} \quad \text{ou} \quad \begin{cases} \sum_{j=1}^n Q_{kj} = b_k \\ \sum_{j=1}^n Q_{jk} = c_k \end{cases}.$$

Assumiremos que o vetor de probabilidade b não tenha coordenada nula, ou seja $b_i \neq 0$ para todo $1 \leq i \leq n$. Caso contrário, $b_k = 0$ para algum k , observe que não há nada o que ser transportado das entradas de a para b_k , logo a k -ésima coluna do plano P será toda zero, e também nada será transportado de b_k para as entradas de c , donde a k -ésima linha do plano Q será toda nula. Sendo assim, os produtos envolvendo os termos $P_{ik} = Q_{kj} = 0$ seriam ignorados nas contas que se seguem.

Definiremos a matriz S da seguinte forma

$$S = P \begin{pmatrix} \frac{1}{b_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{b_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{b_n} \end{pmatrix} Q.$$

Inicialmente, verificaremos que S é um plano de transporte entre a e c , isto é,

$$S \in U(a, c) \iff \begin{cases} S \mathbf{1}_n = a \\ S^\top \mathbf{1}_n = c \end{cases}.$$

Multiplicando a equação que define S por $\mathbf{1}_n$, em ambos os lados, teremos:

$$S \mathbf{1}_n = P \begin{pmatrix} \frac{1}{b_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{b_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{b_n} \end{pmatrix} \underbrace{Q \mathbf{1}_n}_{=b} = P \begin{pmatrix} \frac{1}{b_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{b_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{b_n} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = P \mathbf{1}_n = a.$$

Analogamente, segue que

$$S^\top \mathbf{1}_n = Q^\top \begin{pmatrix} \frac{1}{b_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{b_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{b_n} \end{pmatrix}^\top \underbrace{P^\top \mathbf{1}_n}_{=c} = Q^\top \begin{pmatrix} \frac{1}{b_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{b_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{b_n} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = Q^\top \mathbf{1}_n = c.$$

Agora, vamos obter a desigualdade triangular para a métrica de Wasserstein, utilizando o plano de transporte S como valor intermediador. isto é, vamos mostrar que

$$W(a, c) \leq \sum_{i=1}^n \sum_{j=1}^n C_{ij} S_{ij} \leq W(a, b) + W(b, c).$$

Pela definição da matriz S , temos que

$$S_{ij} = \sum_{k=1}^n P_{ik} \frac{1}{b_k} Q_{kj}.$$

Com isso,

$$\begin{aligned} W(a, c) &= \min_{X \in U(a, c)} \sum_{i=1}^n \sum_{j=1}^n C_{ij} X_{ij} \stackrel{S \in U(a, c)}{\leq} \sum_{i=1}^n \sum_{j=1}^n C_{ij} S_{ij} \\ &= \sum_{i=1}^n \sum_{j=1}^n C_{ij} \left(\sum_{k=1}^n P_{ik} \frac{1}{b_k} Q_{kj} \right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n C_{ij} \left(P_{ik} \frac{1}{b_k} Q_{kj} \right) \\ &\stackrel{(C3)}{\leq} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n (C_{ik} + C_{kj}) \left(P_{ik} \frac{1}{b_k} Q_{kj} \right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n C_{ik} P_{ik} \frac{1}{b_k} Q_{kj} + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n C_{kj} P_{ik} \frac{1}{b_k} Q_{kj} \\ &= \sum_{i=1}^n \sum_{k=1}^n C_{ik} P_{ik} \frac{1}{b_k} \underbrace{\sum_{j=1}^n Q_{kj}}_{=b_k} + \sum_{k=1}^n \sum_{j=1}^n C_{kj} Q_{kj} \frac{1}{b_k} \underbrace{\sum_{i=1}^n P_{ik}}_{=b_k} \\ &= \sum_{i=1}^n \sum_{k=1}^n C_{ik} P_{ik} \left(\frac{1}{b_k} b_k \right) + \sum_{k=1}^n \sum_{j=1}^n C_{kj} Q_{kj} \left(\frac{1}{b_k} b_k \right) \\ &= \sum_{i=1}^n \sum_{k=1}^n C_{ik} P_{ik} + \sum_{k=1}^n \sum_{j=1}^n C_{kj} Q_{kj} \\ &= W(a, b) + W(b, c). \end{aligned}$$

Por fim, $W(a, c) \leq W(a, b) + W(b, c)$ como desejado. \square

Comparação de Histogramas

Finalizamos este capítulo com alguns exemplos simples, onde a métrica de Wasserstein pode ser facilmente calculada (ou estimada), como forma de compreender as características dos vetores de probabilidades, ou dos histogramas, que esta distância se propõem a quantificar.

A primeira definição é sobre a matriz de custos que devemos levar em consideração. Segundo o enunciado do Teorema 3.21, uma tal matriz deve satisfazer as condições (C1),

(C2) e (C3). Uma escolha natural, será considerar o módulo da diferença entre os índices da matriz, ou seja, $C_{ij} = |i - j|$, que já foi introduzida no Exemplo 3.20.

Exemplo 3.22 (Deslocamento de Histograma I). Desejamos calcular a distância entre os seguintes vetores de probabilidade em \mathcal{P}_8 , dados por

$$a = (0, 0, 1, 0, 0, 0, 0, 0)^T, \quad b = (0, 0, 0, 1, 0, 0, 0, 0)^T \quad \text{e} \quad c = (0, 0, 0, 0, 0, 0, 0, 1)^T.$$

Seus respectivos histogramas são as Figuras 9, 10 e 11, que são similares a menos de deslocamento.

Figura 9 – Histograma Associado ao Vetor a



Fonte: Elaborada pelo autor.

Figura 10 – Histograma Associado ao Vetor b



Fonte: Elaborada pelo autor.

Figura 11 – Histograma Associado ao Vetor c



Fonte: Elaborada pelo autor.

Abaixo exibimos os cálculos da métrica de Wasserstein envolvendo tais vetores/histogramas, para isto, analisaremos o respectivo Problema de Kantorovich cuja matriz de custos é

$$C = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 0 & 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 1 & 0 & 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 & 0 & 1 & 2 & 3 \\ 5 & 4 & 3 & 2 & 1 & 0 & 1 & 2 \\ 6 & 5 & 4 & 3 & 2 & 1 & 0 & 1 \\ 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 \end{pmatrix}$$

onde, por exemplo, para se fazer o transporte do índice 3 para o 4 teremos um custo de transporte de $|3 - 4| = 1$ e para o transporte do índice 4 para o 8 teremos um custo de $|4 - 8| = 4$.

- Distância de Wasserstein entre a e b ($W(a, b) = W(b, a)$): É fácil ver que existe único plano de transporte, $U(a, b) = \{\Pi\}$, dado pela matriz

$$\Pi = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

que também deverá ser o plano de transporte ótimo. Portanto,

$$W(a, b) = \min_{X \in U(a, b)} \sum_{i=1}^n \sum_{j=1}^n C_{ij} X_{ij} = \sum_{i=1}^n \sum_{j=1}^n C_{ij} \Pi_{ij} = C_{3,4} \Pi_{3,4} = |3 - 4| \times 1 = 1;$$

- Distância de Wasserstein entre b e c ($W(b, c) = W(c, b)$): Analogamente, $U(b, c) = \{\Pi^*\}$ para

$$\Pi^* = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

e a distancia é dada por

$$W(b, c) = \min_{X \in U(b, c)} \sum_{i=1}^n \sum_{j=1}^n C_{ij} X_{ij} = \sum_{i=1}^n \sum_{j=1}^n C_{ij} \Pi_{ij}^* = C_{4,8} \Pi_{4,8}^* = |4 - 8| \times 1 = 4;$$

- Distância de Wasserstein entre c e a ($W(c, a) = W(a, c)$): Por fim, $U(b, c) = \{\Pi^{**}\}$ com

$$\Pi^{**} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

cuja distância é

$$W(c, a) = \min_{X \in U(c, a)} \sum_{i=1}^n \sum_{j=1}^n C_{ij} X_{ij} = \sum_{i=1}^n \sum_{j=1}^n C_{ij} \Pi_{ij}^{**} = C_{8,3} \Pi_{8,3}^{**} = |8 - 3| \times 1 = 5.$$

A partir dos valores acima, observe que a métrica de Wasserstein é capaz de identificar e quantificar o deslocamento de massa entre índices distintos para os vetores de probabilidade mencionados anteriormente.

Por outro lado, este comportamento de deslocamento não é reconhecido pela distância Euclidiana (veja Exemplo 3.17) entre os vetores a , b e c como calculada abaixo.

- Distância Euclidiana entre a e b ($d(a, b) = d(b, a) = \sqrt{\sum_{i=1}^8 (a_i - b_i)^2}$):

$$\begin{aligned} d(a, b) &= \sqrt{(0-0)^2 + (0-0)^2 + (1-0)^2 + (0-1)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2} \\ &= \sqrt{2} \end{aligned}$$

- Distância Euclidiana entre b e c ($d(b, c) = d(c, b) = \sqrt{\sum_{i=1}^8 (b_i - c_i)^2}$):

$$\begin{aligned} d(b, c) &= \sqrt{(0-0)^2 + (0-0)^2 + (0-0)^2 + (1-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-1)^2} \\ &= \sqrt{2} \end{aligned}$$

- Distância Euclidiana entre c e a ($d(c, a) = d(a, c) = \sqrt{\sum_{i=1}^8 (c_i - a_i)^2}$):

$$\begin{aligned} d(c, a) &= \sqrt{(0-0)^2 + (0-0)^2 + (0-1)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (0-0)^2 + (1-0)^2} \\ &= \sqrt{2} \end{aligned}$$

Exemplo 3.23 (Comparação de Histogramas II). Considere os vetores de probabilidade em \mathcal{P}_7

$$a = \left(0, 0, 0, \frac{1}{10}, \frac{2}{10}, \frac{4}{10}, \frac{3}{10}\right)^T \quad \text{e} \quad b = \left(\frac{1}{10}, \frac{2}{10}, \frac{4}{10}, \frac{3}{10}, 0, 0, 0\right)^T$$

e seus respectivos histogramas dados nas Figuras 12 e 13.

Figura 12 – Histograma Associado ao Vetor a



Fonte: Elaborada pelo autor.

Figura 13 – Histograma Associado ao Vetor b



Fonte: Elaborada pelo autor.

Dada a matriz de custos

$$C = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 0 & 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 0 & 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 0 & 1 & 2 & 3 \\ 4 & 3 & 2 & 1 & 0 & 1 & 2 \\ 5 & 4 & 3 & 2 & 1 & 0 & 1 \\ 6 & 5 & 4 & 3 & 2 & 1 & 0 \end{pmatrix},$$

apresentamos um plano de transporte $\Pi \in U(a, b)$, dado pela matriz

$$\Pi = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{10} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{2}{10} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{4}{10} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{3}{10} & 0 & 0 & 0 \end{pmatrix},$$

que não necessariamente será único ou ótimo. Em particular, temos que

$$W(a, b) \leq \sum_{i=1}^n \sum_{j=1}^n C_{ij} \Pi_{ij} = |1 - 4| \frac{1}{10} + |2 - 5| \frac{2}{10} + |3 - 6| \frac{4}{10} + |4 - 7| \frac{3}{10} = 3.$$

Pelos resultados anteriores, a distância de Wasserstein consegue assimilar e quantificar o deslocamento de vetores de probabilidade que são similares. Na realidade, em vista do problema de transporte, a Métrica de Wasserstein permite comparar as entradas dos vetores de probabilidade em índices distintos i e j , mas para isso, prescreve um custo que é proporcional a separação destes índices. Destacamos que tal comportamento pode ser útil em diversas aplicações que tentaremos explorar no próximo capítulo.

4 Aplicações

Neste capítulo, exploraremos algumas aplicações relacionadas à comparação de histogramas, utilizando a métrica de Wasserstein como principal ferramenta de estudo.

4.1 Comparação de Histogramas da COVID-19

Etapa 1. Compreensão da Situação-Problema:

Atualmente, os meios de comunicações veiculam frequentemente a seguinte frase: “[...] *estamos vivendo uma nova onda na pandemia causada pela COVID-19* [...]”. Tal afirmação nos leva a questionar o que é (e como reconhecer) uma “onda” no contexto da pandemia.

No século XIX, pesquisadores (Biólogos, Epidemiologistas, etc) reconheceram padrões regulares de um comportamento ondulatório em seus dados sobre a propagação de uma doença/vírus, o que levou a disseminação do termo “onda” epidemiológica. As tentativas de definição formal para este fenômeno, muitas vezes não conseguem abranger a diversidade da doença/vírus existentes e, por este motivo, não existe um consenso do que de fato caracteriza uma “onda” epidemiológica.

Nossa principal motivação ao desenvolver esta aplicação é o fato de que, na Matemática, pode-se descrever diversos tipos de ondas do mundo real por meio de funções conhecidas. Por exemplo:

- A vibração de uma corda de um violão é caracterizada pelas ondas harmônicas (unidimensionais), cuja representação emprega a função cosseno, $y = \cos(x)$ (ou a função seno, $y = \sin(x)$);
- As frentes de onda (unidimensional), observada em ondas oceânicas, podem ser descritas pela função cúspide cúbica, dada por $y = 1 - |x|^{\frac{2}{3}}$.

Nesta seção, iremos comparar por similaridade o histogramas obtidos a partir de gráficos da propagação da COVID-19 no Brasil com histogramas associados às funções matemáticas que modelam este fenômeno de “onda” (função cosseno $\cos(x)$ e função cúspide cúbica $1 - |x|^{\frac{2}{3}}$).

Etapa 2. Construção de um Modelo Matemático:

Inicialmente, fomos em busca dos dados da evolução da COVID-19 no Brasil, para que pudéssemos definir como abordaríamos estes dados. Consultamos o banco de

dados do *Conselho Nacional de Secretários de Saúde (CONASS)* — única plataforma atualizada disponível no período da análise — e extraímos as informações necessárias para nosso trabalho. Coletamos os dados, referentes ao período do dia 15/03/2020 até o dia 25/09/2021, organizados por semanas epidemiológicas da COVID-19. Note que a primeira semana de incidência do vírus no Brasil, a partir do dia 15/03/2020 até 21/03/2020, coincide com os seguintes acontecimentos: registro da primeira morte pela doença (em 18/03/2020) e declaração de transmissão comunitária no país pelo *Ministério da Saúde* (no dia 21/03/2020). Vale salientar que esta primeira semana do vírus no Brasil, na realidade, corresponde à 12ª semana na escala internacional indicada pela *World Health Organization (WHO)*¹.

Encontramos dois conjuntos de dados que julgamos pertinentes para nossa análise: o primeiro é referente à evolução de números de casos da doença e o segundo diz respeito ao número de óbitos ocasionados pela COVID-19. Para melhorar o entendimento e simplificar a apresentação das comparações que serão realizadas, agrupamos os dados a cada 4 semanas, ou seja, o primeiro dado apresentado vai da semana 12 até a semana 15, o segundo contabiliza das semanas 16 até 19 e assim por diante até o último dado que compreende da semana 89 a semana 91. Essa divisão propõem a apresentação dos casos de óbitos em 20 intervalos, abrangendo um total de 79 semanas de incidência da COVID-19 no Brasil.

O primeiro conjunto de dados referente à propagação de números de casos da COVID-19 no Brasil está apresentado na Tabela 2, onde também apresentamos uma normalização dos mesmos, que se justifica pelo fato de que temos a intenção de trabalhar com vetores de probabilidade. Desta forma, realizamos a construção do histograma que representa esses dados percentuais dos números de casos na Figura 14.

Analogamente, exibimos a Tabela 3 referente ao segundo conjunto de dados sobre a evolução dos números de óbitos no Brasil e respectivo histograma dado pela Figura 15.

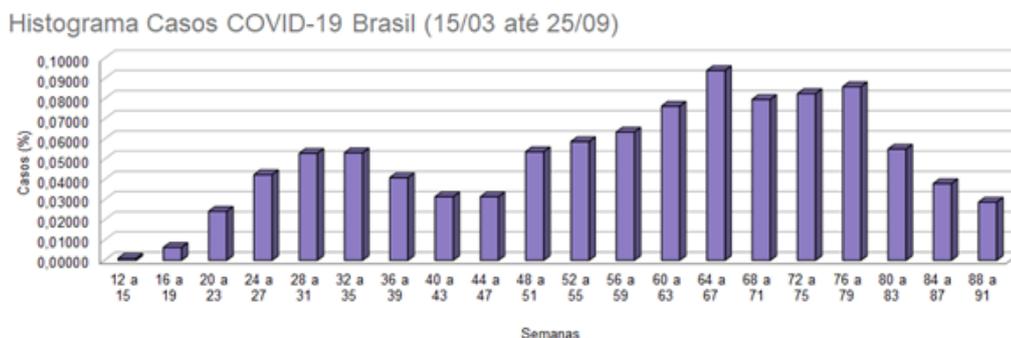
¹ Tradução livre do autor: Organização Mundial de Saúde (OMS)

Tabela 2 – Números de Casos da COVID-19

Intervalo	Semanas	Número de Casos	Porcentagem dos Casos
1	12 a 15	20606	0,00097
2	16 a 19	135212	0,00633
3	20 a 23	516907	0,02421
4	24 a 27	904158	0,04234
5	28 a 31	1130873	0,05296
6	32 a 35	1136930	0,05325
7	36 a 39	873184	0,04089
8	40 a 43	671648	0,03146
9	44 a 47	672151	0,03148
10	48 a 51	1147922	0,05376
11	52 a 55	1254351	0,05875
12	56 a 59	1354695	0,06345
13	60 a 63	1629181	0,07630
14	64 a 67	2006071	0,09395
15	68 a 71	1700873	0,07966
16	72 a 75	1761546	0,08250
17	76 a 79	1834600	0,08592
18	80 a 83	1175830	0,05507
19	84 a 87	810750	0,03797
20	88 a 91	614699	0,02879
	Total	21352187	1,00000

Fonte: Elaborada pelo autor.

Figura 14 – Histograma dos Números de Casos da COVID-19



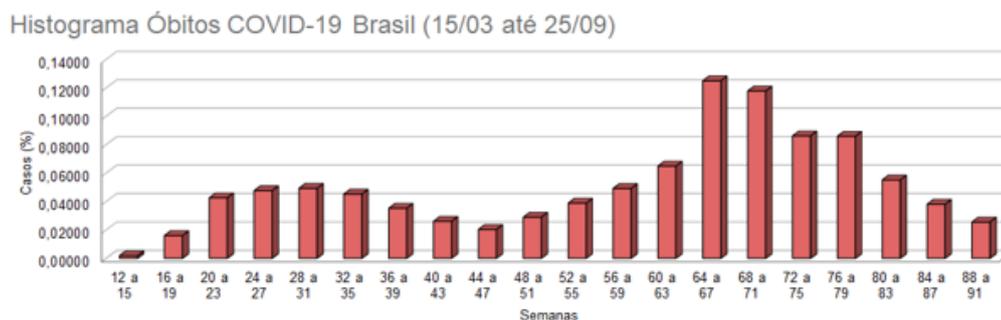
Fonte: Elaborado pelo autor.

Tabela 3 – Números de Óbitos da COVID-19

Intervalo	Semanas	Número de Óbitos	Porcentagem dos Óbitos
1	12 a 15	1124	0,00189
2	16 a 19	9503	0,01599
3	20 a 23	25303	0,04258
4	24 a 27	28335	0,04769
5	28 a 31	29298	0,04931
6	32 a 35	26889	0,04525
7	36 a 39	20954	0,03526
8	40 a 43	15497	0,02608
9	44 a 47	12086	0,02034
10	48 a 51	17216	0,02897
11	52 a 55	23091	0,03886
12	56 a 59	29236	0,04920
13	60 a 63	38559	0,06489
14	64 a 67	74243	0,12495
15	68 a 71	69982	0,11778
16	72 a 75	51215	0,08619
17	76 a 79	51056	0,08592
18	80 a 83	32783	0,05517
19	84 a 87	22640	0,03810
20	88 a 91	15190	0,02556
	Total	594200	1,00000

Fonte: Elaborada pelo autor.

Figura 15 – Histograma dos Números de Óbitos da COVID-19



Fonte: Elaborado pelo autor.

A partir de breve observação visual do comportamento dos dois histogramas apresentados anteriormente, percebemos duas regiões com características muito próximas de uma “onda”. A primeira região é formada pelos primeiros dez intervalos, ou seja, do intervalo 1 (semanas de 12 a 15) até o intervalo 10 (semanas de 48 a 51). Vimos em ambos os histogramas (Figura 15 e Figura 14) a possibilidade de similaridade desta região com uma função que classicamente é utilizada para modelar ondas, deste modo, escolhemos a função Cosseno

$$y = \cos(x)$$

como um parâmetro de comparação. Já a segunda região vai do intervalo 11 (semanas de 52 a 55) até o intervalo 20 (semanas de 88 a 91), ou melhor, os dez últimos intervalos. Principalmente, para o histograma que representa o número de óbitos (Figura 15), identificamos um comportamento de “cume” ou frente de onda e, por conta disso, escolhemos função Cúspide Cúbica

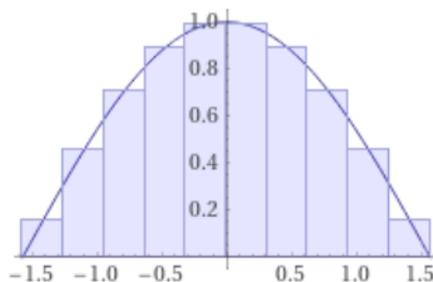
$$y = \text{cusp}(x) = 1 - |x|^{\frac{2}{3}}$$

para servir de parâmetro de nossas comparações de histogramas.

Com esta observação nos dez primeiros e dez últimos intervalos, resolvemos modelar os histogramas associados à função cosseno e à função cúspide cúbica com dez barras como detalhado a seguir:

- Construímos o gráfico da função cosseno $y = \cos(x)$ no intervalo de $-\frac{\pi}{2}$ a $\frac{\pi}{2}$ e dividimos esse intervalo em dez partes iguais. Desejamos construir um histograma sobre o gráfico, onde cada barra incide sob o valor da função no ponto médio em cada um destes intervalos. Daí, encontramos o valor de cada um desses pontos médios, que denotamos por $x_i = -\frac{\pi}{2} + \left[(2 \times i - 1) \times \frac{\pi}{20}\right]$, e em seguida calculamos o valor do cosseno de cada x_i , no *Excel*. Por fim, normalizamos esses valores para encontrar um vetor de probabilidade com 10 entradas e assim apresentar o histograma da função cosseno. Veja a Figura 16, a Tabela 4 e a Figura 17.

Figura 16 – Gráfico da Função Cosseno



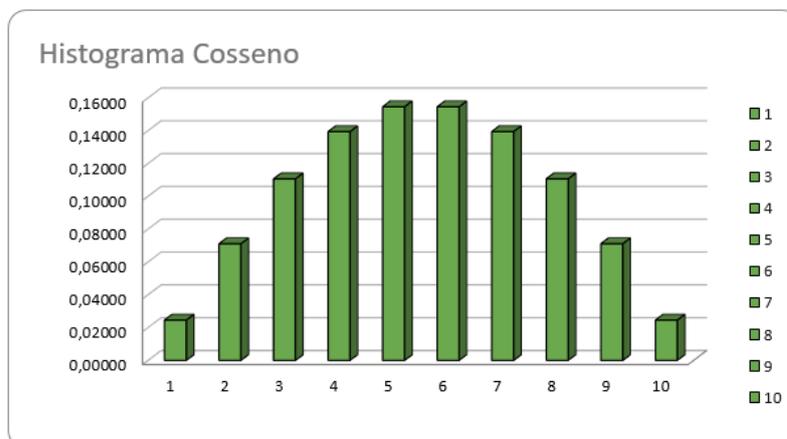
Fonte: Elaborada no ([WOLFRAM|ALPHA](#), 2022).

Tabela 4 – Valores para Função Cosseno

Índice i	Fórmula x_i	x_i	$\cos(x_i)$	Porcentagem $\cos(x_i)$
1	$-\frac{\pi}{2} + (1 \times \frac{\pi}{20})$	-1,413717	0,156	0,02447
2	$-\frac{\pi}{2} + (3 \times \frac{\pi}{20})$	-1,099557	0,454	0,07102
3	$-\frac{\pi}{2} + (5 \times \frac{\pi}{20})$	-0,785398	0,707	0,11062
4	$-\frac{\pi}{2} + (7 \times \frac{\pi}{20})$	-0,471239	0,891	0,13938
5	$-\frac{\pi}{2} + (9 \times \frac{\pi}{20})$	-0,157080	0,988	0,15451
6	$-\frac{\pi}{2} + (11 \times \frac{\pi}{20})$	0,157080	0,988	0,15451
7	$-\frac{\pi}{2} + (13 \times \frac{\pi}{20})$	0,471239	0,891	0,13938
8	$-\frac{\pi}{2} + (15 \times \frac{\pi}{20})$	0,785398	0,707	0,11062
9	$-\frac{\pi}{2} + (17 \times \frac{\pi}{20})$	1,099557	0,454	0,07102
10	$-\frac{\pi}{2} + (19 \times \frac{\pi}{20})$	1,413717	0,156	0,02447
		Total	3,972	1,00000

Fonte: Elaborada pelo autor.

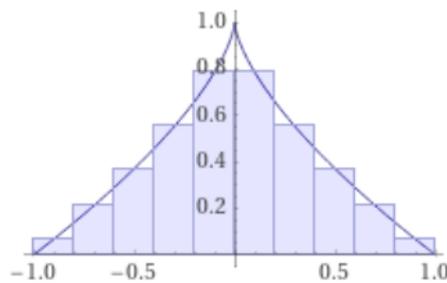
Figura 17 – Histograma da Função Cosseno



Fonte: Elaborada pelo autor.

- Consideramos o intervalo de -1 a 1 para análise do gráfico da função cúspide cúbica $y = \text{cusp}(x) = 1 - |x|^{\frac{2}{3}}$ e, em seguida, dividimos esse intervalo em dez partes iguais. Assim como na função cosseno, o gráfico da cúspide cúbica intercepta a linha superior das barras do histograma nos respectivos pontos médios dos intervalos em questão. Chamamos de $x_i = -1 + [(2 \times n - 1) \times 0,1]$ os pontos médios dos intervalos e, então, calculamos os valores da função para cada um dos x_i encontrados. Ao fim, normalizamos esses valores de x_i para obter o vetor de probabilidade e o histograma associado a função cúspide cúbica. Elaboramos a Figura 18, a Tabela 5 e a Figura 19 dessa função conforme abaixo.

Figura 18 – Gráfico da Função Cúspide Cúbica



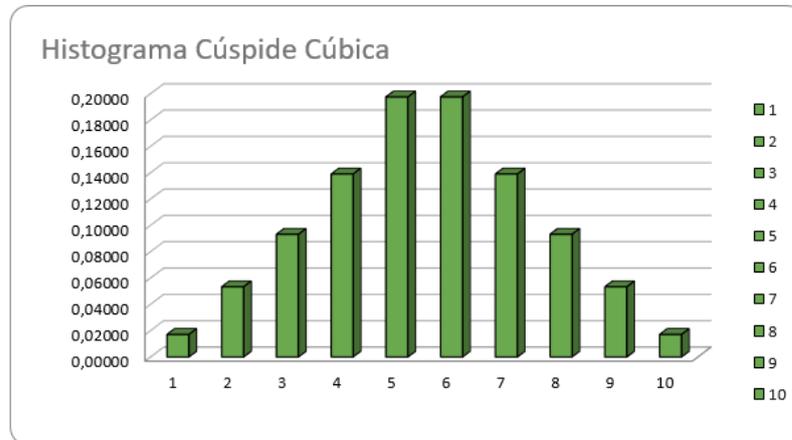
Fonte: Elaborada no ([WOLFRAM|ALPHA, 2022](#))

Tabela 5 – Valores para Função Cúspide Cúbica

Índice i	Fórmula x_i	x_i	$\text{cusp}(x_i)$	Porcentagem $\text{cusp}(x_i)$
1	$-1 + (1 \times 0,1)$	-0,9	0,068	0,01708
2	$-1 + (3 \times 0,1)$	-0,7	0,212	0,05328
3	$-1 + (5 \times 0,1)$	-0,5	0,370	0,09317
4	$-1 + (7 \times 0,1)$	-0,3	0,552	0,13894
5	$-1 + (9 \times 0,1)$	-0,1	0,785	0,19753
6	$-1 + (11 \times 0,1)$	0,1	0,785	0,19753
7	$-1 + (13 \times 0,1)$	0,3	0,552	0,13894
8	$-1 + (15 \times 0,1)$	0,5	0,370	0,09317
9	$-1 + (17 \times 0,1)$	0,7	0,212	0,05328
10	$-1 + (19 \times 0,1)$	0,9	0,068	0,01708
		Total	3,972	1,00000

Fonte: Elaborada pelo autor.

Figura 19 – Histograma da Função Cúspide Cúbica



Fonte: Elaborada pelo autor.

Note que os dois histogramas que obtivemos referente à COVID-19 possuem 20 barras (logo os vetores de probabilidade possuem 20 entradas) e os histogramas das funções cosseno e cúspide cúbica tem 10 barras (vetores de probabilidade com 10 entradas). Para que seja possível compará-los, decidimos dividir cada histograma da COVID-19 em 11 novos histogramas, cada um com 10 barras (10 entradas).

Inicialmente, tomamos os dados referentes aos intervalos de 1 até 10 do Histograma de Casos da COVID-19 (Tabela 2), normalizamos e geramos um vetor de probabilidade a_{1-10} . Em seguida, para os dados de 2 até 11, realizamos novamente a normalização e obtemos o vetor a_{2-11} . Repetimos esse processo até obter o vetor de probabilidade a_{11-20} . Analogamente, para o Histograma de Óbitos da COVID-19 (Tabela 3) construímos os vetores d_{1-10} , d_{2-11} até d_{11-20} .

Os respectivos vetores de probabilidade que decompõem o Histograma de Casos da COVID-19 e o Histograma de Óbitos da COVID-19 são apresentados no Apêndice B. Além disso, a correspondência entre o índice que referencia estes vetores e o período que ele compreende, é dado na Tabela 6

Tabela 6 – Divisão Intervalos dos Histogramas para Cálculo de Distâncias

Intervalo de Índices	Semanas Iniciais	Semanas Finais
1-10	12 a 15	48 a 51
2-11	16 a 19	52 a 55
3-12	20 a 23	56 a 59
4-13	24 a 27	60 a 63
5-14	28 a 31	64 a 67
6-15	32 a 35	68 a 71
7-16	36 a 39	72 a 75
8-17	40 a 43	76 a 79
9-18	44 a 47	80 a 83
10-19	48 a 51	84 a 87
11-20	52 a 55	88 a 91

Fonte: Elaborada pelo autor.

Etapa 3. Resolução do Modelo (via *Solver*):

Definido todos os vetores de probabilidade e respectivos histogramas que serão utilizados como parâmetros para a comparação, vamos começar a construir as planilhas para os cálculos de distâncias por meio da métrica de Wasserstein.

A primeira definição é a função custo unitário, que será definida como sendo a distância de um índice a outro, ou seja, o módulo da diferença entre os índices que variam de 1 até 10. Assim, nossos coeficientes das variáveis de decisão serão representados por $C_{ij} = |i - j|$ ou, mais especificamente, pela matriz 10×10

$$C = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 2 & 1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 3 & 2 & 1 & 0 & 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 & 0 & 1 & 2 & 3 & 4 \\ 6 & 5 & 4 & 3 & 2 & 1 & 0 & 1 & 2 & 3 \\ 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 & 1 & 2 \\ 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 & 1 \\ 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 \end{pmatrix}.$$

Em particular, sua respectiva vetorização é dada pelo vetor 1×100

$$\text{vet}(C)^T = (0, 1, 2, 3, 4, 5, 6, 7, 8, 9 \mid 1, 0, 1, 2, 3, 4, 5, 6, 7, 8 \mid \dots \mid 9, 8, 7, 6, 5, 4, 3, 2, 1, 0).$$

Também apresentamos o plano de transporte (uma matriz 10×10) e sua vetorização

O primeiro passo é alimentar o *Excel* com a matriz A que define as restrições para um problema de Programação Linear, digitando todas as entradas dessa matriz na planilha do *Excel*. Abaixo desta matriz colocamos uma linha com 100 colunas onde inserimos os respectivos custos para cada uma das colunas. Logo abaixo desta linha, criamos outra linha com 100 colunas onde o *Solver* através da fórmula $[SOMARPRODUTO]$ irá preencher com os valores ótimos das variáveis de decisão, ou melhor, com as entradas de um plano de transporte ótimo. Posteriormente, criamos três colunas cada uma com 20 linhas: a da esquerda, chamamos de **(Totais)** e será preenchida automaticamente pelo *Solver* com a fórmula $[SOMARPRODUTO]$; a coluna do meio, possuirá a representação simbólica do sinal das restrições (neste caso, sempre será o símbolo =); e na última coluna, denominada de **(Restrições)**, iremos preencher com a concatenação dos vetores de probabilidades dados pelos histogramas. Por fim, criamos uma célula, logo abaixo da coluna dos **(Totais)**, onde o *Solver* irá preencher com a distância total ótima de nossa comparação, também em função da fórmula $[SOMARPRODUTO]$.

Devido ao tamanho de nossa matriz, compartilhamos apenas um recorte da parte final da tela quando configuramos nossa planilha para o cálculo com o *Solver*, conforme Figura 20.

Figura 20 – Recorte da Planilha de Preparação para o *Solver*

CJ	CK	CL	CM	CN	CO	CP	CQ	CR	CS	CT	CU	CV	CW	CX	CY	CZ	DA
9:7	9:8	9:9	9:10	10:1	10:2	10:3	10:4	10:5	10:6	10:7	10:8	10:9	10:10		Totais		Restrições
0	0	0	0	1	0	0	0	0	0	0	0	0	0		=SOMARPRODUTO(B3:CW3;\$B\$26:\$CW\$26)		
0	0	0	0	0	1	0	0	0	0	0	0	0	0		0.00000	="	0.00000
0	0	0	0	0	0	1	0	0	0	0	0	0	0		0.00000	="	0.00000
0	0	0	0	0	0	0	1	0	0	0	0	0	0		0.00000	="	0.00000
0	0	0	0	0	0	0	0	1	0	0	0	0	0		0.00000	="	0.00000
1	0	0	0	0	0	0	0	0	1	0	0	0	0		0.00000	="	0.00000
0	1	0	0	0	0	0	0	0	0	1	0	0	0		0.00000	="	0.00000
0	0	1	0	0	0	0	0	0	0	0	1	0	0		0.00000	="	0.00000
0	0	0	1	0	0	0	0	0	0	0	0	1	0		0.00000	="	0.00000
0	0	0	0	0	0	0	0	0	0	0	0	0	1		0.00000	="	0.00000
0	0	0	0	0	0	0	0	0	0	0	0	0	0		0.00000	="	0.00000
0	0	0	0	0	0	0	0	0	0	0	0	0	0		0.00000	="	0.00000
0	0	0	0	0	0	0	0	0	0	0	0	0	0		0.00000	="	0.00000
0	0	0	0	0	0	0	0	0	0	0	0	0	0		0.00000	="	0.00000
0	0	0	0	0	0	0	0	0	0	0	0	0	0		0.00000	="	0.00000
0	0	0	0	0	0	0	0	0	0	0	0	0	0		0.00000	="	0.00000
1	1	1	1	0	0	0	0	0	0	0	0	0	0		0.00000	="	0.00000
0	0	0	0	1	1	1	1	1	1	1	1	1	1		0.00000	="	0.00000
2	1	0	1	9	8	7	6	5	4	3	2	1	0		0.00000		
0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000				

Fonte: Elaborada pelo autor.

Para nossas comparações, iremos alimentar a coluna **(Restrições)** da planilha do *Excel* com os dados dos vetores de probabilidade (histogramas) para os quais desejamos calcular a distância de Wasserstein. Como os vetores possuem 10 entradas e esta coluna possui 20 linhas, devemos proceder da seguinte forma:

- As dez primeiras linhas serão alimentadas com os dados de um dos vetores de

probabilidade referente aos histogramas de caso ou óbitos da COVID-19, neste casos, os vetores $a_{(\cdot)}$ ou $d_{(\cdot)}$.

- As dez últimas possuirão dados do vetor de probabilidade associado aos histogramas da função que utilizaremos como parâmetro de comparação, em nosso caso, ou a função cosseno b_{\cos} ou a função cúspide cúbica b_{cusp} .

Lembrando que todos os vetores de probabilidade mencionados acima podem ser encontrados no Apêndice B.

A partir daqui executamos 44 instâncias do *Solver* no *Excel*, para realizar o cálculo de: 11 distâncias $W(a_{(\cdot)}, b_{\cos})$; 11 distâncias $W(a_{(\cdot)}, b_{\text{cusp}})$; 11 distâncias $W(d_{(\cdot)}, b_{\cos})$; 11 distâncias $W(d_{(\cdot)}, b_{\text{cusp}})$. Não vamos apresentar todas as telas de cálculo destas distâncias, principalmente, por conta dos tamanhos da matriz e dos vetores envolvidos. Caso o leitor tenha interesse em vê-los, pode consultar a planilha no seguinte link disponibilizado pelos autores: (GOMES; SILVA, 2022).

Etapa 4. Análise da Solução:

Realizaremos tal comparação de similaridade entre os histogramas por meio de distâncias entre os respectivos vetores de probabilidades, ou seja, quanto mais próximo do valor zero for o valor da distância, mais similares serão os vetores de probabilidades e os histogramas. Para nossas análises utilizaremos a métrica de Wasserstein e a distância Euclidiana em \mathbb{R}^{10} . Intuitivamente, a distância Euclidiana em \mathbb{R}^{10} compara e quantifica a diferença de entrada de mesmo índice i dos vetores de probabilidade (barras em mesma posição dos histogramas). Recorde que a métrica de Wasserstein também permite comparar e quantificar a diferença de entrada em índices distintos i e j , colocando um custo C_{ij} proporcional a separação destes índices.

Etapa 4.1. Análise da Solução: Histograma de Casos da COVID-19

Devido a decomposição dos dados do Histograma de Casos da COVID-19 (Tabela 2 e Figura 14) em 11 novos vetores de probabilidade ($a_{(\cdot)}$), teremos que fazer 11 comparações em relação ao vetor de probabilidade da função cosseno b_{\cos} (Tabela 4 e Figura 17) e mais outras 11 em relação ao vetor de probabilidade da função cúspide cúbica b_{cusp} (Tabela 5 e Figura 19).

Apresentaremos abaixo as tabelas contendo todos os 22 valores de distâncias que apuramos para os Casos da COVID-19 nas distâncias de Wasserstein e Euclidiana.

Tabela 7 – Distâncias de Wasserstein para o Histogramas de Casos da COVID-19

Hist. Casos	Período em Semanas	Cosseno b_{\cos}	Cúspide b_{cusp}
a_{1-10}	12 a 15 - 48 a 51	1,01192	1,01192
a_{2-11}	16 a 19 - 52 a 55	0,69220	0,77904
a_{3-12}	20 a 23 - 56 a 59	0,78854	1,02478
a_{4-13}	24 a 27 - 60 a 63	0,98304	1,21928
a_{5-14}	28 a 31 - 64 a 67	1,12071	1,30801
a_{6-15}	32 a 35 - 68 a 71	1,08575	1,23595
a_{7-16}	36 a 39 - 72 a 75	1,02563	1,10281
a_{8-17}	40 a 43 - 76 a 79	0,88056	0,97366
a_{9-18}	44 a 47 - 80 a 83	0,57082	0,72544
a_{10-19}	48 a 51 - 84 a 87	0,35687	0,59311
a_{11-20}	52 a 55 - 88 a 91	0,48288	0,7149
	Variação Total	0,76384	0,69693

Fonte: Elaborada pelo autor.

Tabela 8 – Distâncias Euclidianas para o Histogramas de Casos da COVID-19

Hist. Casos	Período em Semanas	Cosseno b_{\cos}	Cúspide b_{cusp}
a_{1-10}	12 a 15 - 48 a 51	0,15561	0,16552
a_{2-11}	16 a 19 - 52 a 55	0,16544	0,20231
a_{3-12}	20 a 23 - 56 a 59	0,18707	0,24152
a_{4-13}	24 a 27 - 60 a 63	0,21275	0,26835
a_{5-14}	28 a 31 - 64 a 67	0,23060	0,27705
a_{6-15}	32 a 35 - 68 a 71	0,20927	0,25153
a_{7-16}	36 a 39 - 72 a 75	0,17907	0,22262
a_{8-17}	40 a 43 - 76 a 79	0,15374	0,19833
a_{9-18}	44 a 47 - 80 a 83	0,11035	0,15860
a_{10-19}	48 a 51 - 84 a 87	0,08751	0,13953
a_{11-20}	52 a 55 - 88 a 91	0,09073	0,14008
	Variação Total	0,14309	0,13752

Fonte: Elaborada pelo autor.

Em relação à função cosseno b_{\cos} :

Primeiramente, buscamos regiões que podem caracterizar uma “onda” epidemiológica. Podemos observar que a porção do Histograma de Casos da COVID-19 que mais se mostrou similar a esta função foi no intervalo de **10 - 19**, visto que a métrica de Wasserstein obteve o menor valor $W(a_{10-19}, b_{\cos}) = 0,35687$. Este intervalo correspondente às semanas de (48 a 51) até (84 a 87), onde reconhecemos a “segunda onda” da pandemia. Além disso, o segundo e o terceiro menor valor da métrica de Wasserstein encontram-se na vizinhança deste período. Ao verificar o quarto menor valor $W(a_{2-11}, b_{\cos}) = 0,69220$, identificamos a “primeira onda” da pandemia no intervalo de **2 - 12**, associado às semanas de (16 a 19) até (52 a 55).

O intervalo do Histograma de Casos da COVID-19 que mais se mostrou diferente da função cosseno foi a porção de **5 - 14**, pois a distância de Wasserstein obteve o maior valor $W(a_{5-14}, b_{\cos}) = 1,12071$. De fato, neste intervalo verificamos uma redução das taxas de casos que correspondem às semanas de (28 a 31) até (64 a 67).

Note que com a distância Euclidiana, os valores máximo e mínimo coincidem com a análise da métrica de Wasserstein. Por outro lado, a identificação da “primeira onda” da pandemia só ocorre no sexto menor valor $\|a_{1-10} - b_{\cos}\| = 0,15561$, isto é, no intervalo de **1 - 10** das semanas de (12 a 15) até (48 a 51).

Outra observação importante é a variação total entre as distâncias calculadas. Veja que há uma diferença entre a maior e a menor distância de Wasserstein de $1,12071 - 0,35687 = 0,76384$, que é uma variabilidade muito maior que a da distância Euclidiana de $0,23060 - 0,08751 = 0,14309$.

Em relação à função cúspide cúbica b_{cusp} :

Observamos um comportamento muito parecido ao da função cosseno: o menor valor foi de $W(a_{10-19}, b_{\text{cusp}}) = 0,59311$, que ocorreu no intervalo de **10 - 19**; o quarto menor foi de $W(a_{2-11}, b_{\text{cusp}}) = 0,77904$, correspondente à porção de **2 - 11**; e o maior valor foi de $W(a_{5-14}, b_{\text{cusp}}) = 1,30801$, correspondente à porção de **5 - 14**. Por sua vez, a variação total da métrica de Wasserstein é dada por uma diferença de $1,30801 - 0,59311 = 0,71490$, valor um pouco menor que a variação total desta mesma distância para função cosseno, $0,76384$, e muito maior que a da distância Euclidiana para função cúspide que foi de $0,27705 - 0,13953 = 0,13752$.

Tais números nos mostram que a função cosseno tem uma melhor aproximação do Histograma de Casos da COVID-19 do que a função cúspide cúbica, pois os valores da distância de Wasserstein listados na primeira coluna das Tabelas 7 e 8 são sempre menores que os da segunda coluna.

Por fim, ambas as distâncias associam o comportamento no “cume”, presente na função cúspide cúbica, no período **10 - 19** (semanas de (48 a 51) até (84 a 87)), identificado como a “segunda onda” da pandemia.

Etapa 4.2. Análise da Solução: Histograma de Óbitos da COVID-19

Da mesma forma, fizemos a decomposição dos dados do Histograma de Óbitos da COVID-19 (Tabela 3 e Figura 15) em 11 vetores de probabilidade ($d_{(\cdot)}$), logo encontramos 11 distâncias em relação ao vetor de probabilidade da função cosseno b_{\cos} (Tabela 4 e Figura 17) e outras 11 distâncias com respeito à probabilidade da função cúspide cúbica b_{cusp} (Tabela 5 e Figura 19). Ao total, apresentaremos duas tabelas, cada uma com 22 distâncias, com respeito as distância de Wasserstein (Tabela 9) e Euclidiana (Tabela 10).

Em relação à função cosseno b_{\cos} :

Inicialmente procuramos o fenômeno de “onda” epidemiológica. Observando os valores apurados, vemos que a porção do Histograma de Óbitos da COVID-19 que mais se assemelhou a função cosseno é o intervalo de **1 - 10**, onde foi encontrada a menor distância de Wasserstein $W(d_{1-10}, b_{\cos}) = 0,24478$. Claramente, tivemos números altos de óbitos no período correspondente às semanas de (12 a 15) até (48 a 51) que pode ser identificada como a “primeira onda” da pandemia. Além disso, para o segundo o menor valor da métrica de Wasserstein $W(d_{11-20}, b_{\cos}) = 0,31289$, já identificamos a “segunda onda” no intervalo de **11 - 20**, associado às semanas de (52 a 55) até (88 a 91).

Já a maior distância de Wasserstein foi de $W(d_{6-15}, b_{\cos}) = 1,64019$, donde temos que a porção que menos se assemelhou com a função cosseno foi o intervalo de **6 - 15**, correspondente às semanas de (32 a 35) até (68 a 71). De fato, notamos uma baixa nos números de óbitos para este período.

Encontramos um intervalo de $1,64019 - 0,24478 = 1,39541$ entre os valores extremos da métrica de Wasserstein para a função cosseno, o que denota uma variabilidade muito maior que a distância Euclidiana, $0,31496 - 0,08751 = 0,22674$

Para a distância Euclidiana, a menor distância corresponde aos intervalos **10 - 19**, que é vizinho ao período **11 - 20** da segunda menor distância de Wasserstein. Já o período associado ao valor de maior distância é o intervalo **5 - 14**, também vizinho ao da maior distância de Wasserstein. Por fim, a segunda menor distância Euclidiana se verifica em **1 - 10** que concorda com o período de menor valor da métrica de Wasserstein.

Tabela 9 – Distâncias Wasserstein para o Histogramas de Óbitos da COVID-19

Hist. Óbitos	Período em Semanas	Cosseno b_{\cos}	Cúspide b_{cusp}
d_{1-10}	12 a 15 - 48 a 51	0,24478	0,38896
d_{2-11}	16 a 19 - 52 a 55	0,61139	0,80889
d_{3-12}	20 a 23 - 56 a 59	0,94404	1,18028
d_{4-13}	24 a 27 - 60 a 63	1,06892	1,30516
d_{5-14}	28 a 31 - 64 a 67	1,44842	1,63908
d_{6-15}	32 a 35 - 68 a 71	1,64019	1,76467
d_{7-16}	36 a 39 - 72 a 75	1,55553	1,62057
d_{8-17}	40 a 43 - 76 a 79	1,36731	1,38571
d_{9-18}	44 a 47 - 80 a 83	0,91489	0,93567
d_{10-19}	48 a 51 - 84 a 87	0,35708	0,43428
d_{11-20}	52 a 55 - 88 a 91	0,31289	0,39147
	Variação Total	1,39541	1,37571

Fonte: Elaborada pelo autor.

Tabela 10 – Distâncias Euclidianas para o Histogramas de Óbitos da COVID-19

Hist. Óbitos	Período em Semanas	Cosseno b_{\cos}	Cúspide b_{cusp}
d_{1-10}	12 a 15 - 48 a 51	0,08822	0,11480
d_{2-11}	16 a 19 - 52 a 55	0,14760	0,19091
d_{3-12}	20 a 23 - 56 a 59	0,20711	0,25663
d_{4-13}	24 a 27 - 60 a 63	0,24229	0,29471
d_{5-14}	28 a 31 - 64 a 67	0,31496	0,35773
d_{6-15}	32 a 35 - 68 a 71	0,31279	0,35455
d_{7-16}	36 a 39 - 72 a 75	0,26624	0,31039
d_{8-17}	40 a 43 - 76 a 79	0,21318	0,25442
d_{9-18}	44 a 47 - 80 a 83	0,14305	0,16885
d_{10-19}	48 a 51 - 84 a 87	0,08751	0,13953
d_{11-20}	52 a 55 - 88 a 91	0,09073	0,14008
	Variação Total	0,22674	0,24293

Fonte: Elaborada pelo autor.

Em relação à função cúspide cúbica b_{cusp} :

Mais um vez obtemos o mesmo comportamento da função cosseno: o menor valor foi de $W(a_{1-10}, b_{\text{cusp}}) = 0,38896$, que ocorreu no intervalo de **1 - 10**; o segundo menor foi de $W(d_{11-20}, b_{\text{cusp}}) = 0,39147$, correspondente à porção de **11 - 20**; e o maior valor foi de $W(d_{6-15}, b_{\text{cusp}}) = 1,76467$, correspondente à porção de **6 - 15**. Novamente, todas as distâncias calculadas com a função cúspide cúbica foram maiores de que os respectivos valores apurados para função cosseno. Além disso, a variação total da métrica de Wasserstein, $1,76467 - 0,38896 = 1,37571$, é muito maior que a da distância Euclidiana para função cúspide, $0,11480 - 0,35773 = 0,24293$.

Um fato curioso é que nenhuma das distâncias associou o comportamento de “cume”, presente na função cúspide cúbica, à “segunda onda” da pandemia. Em particular, a distância Euclidiana aproximou a “segunda onda” à função cosseno e a “primeira onda” à função cúspide cúbica, fornecendo assim uma resposta contrária à intuição visual que tivemos no início.

Etapa 4.3 Análise da Solução: Comentários finais

Em linhas gerais, consideramos que obtivemos bons resultados no reconhecimento de “ondas” epidemiológicas com ambas as distâncias Wasserstein e Euclidiana. Acreditamos que a métrica de Wasserstein, devido a sua variabilidade, pode ser mais sensível a pequenas diferenças no histograma.

A função cosseno foi a melhor escolha para a comparação, uma vez que, sempre obteve valores de distância menores. Ainda assim, a função cúspide cúbica é uma boa escolha já que os valores apurados foram próximos do da função cosseno. Por outro lado, o comportamento de “cume”, presente na função cúspide cúbica, não foi precisamente reconhecido no período da “segunda onda” da pandemia.

4.2 Outras Aplicações

Uma das principais aplicações relacionadas à comparação de histogramas é a possibilidade de comparar imagens usando suas características visuais intrínsecas (consulte (SAPPELLI; VERBERNE; KRAAIJ, 2017; WILSON; MARTINEZ, 1997)). Por exemplo, quando um médico radiologista manipula e interpreta imagens radiológicas com o objetivo de ter auxílio no diagnóstico e tomada de decisão para determinados casos (ver (AKGÜL et al., 2011)). Outra importante utilização é a detecção e o reconhecimento de sinais de Libras utilizando a visão computacional para processar imagens e vídeos.

Segundo Liu et al. (2007), ao executar uma consulta por similaridade em imagens, dois tipos de abordagens são amplamente consideradas na literatura:

- As Comparações por Contexto em meta-descrições como *tags*², palavras-chave, rótulos ou ontologias que são associados aos dados por processos externos, tais como, por meio de especialistas humanos ou por estratégias automatizadas. Aplicações de recomendação de conteúdos na web e busca em dados não estruturados, são exemplos de aplicações que utilizam essa abordagem;
- As Comparações por Conteúdo são capazes de expressar a similaridade (ou a dissimilaridade) entre pares de elementos, representados em um mesmo conjunto, utilizando a extração de características representadas internamente nos próprios objetos complexos.

Em ambos os tipos, são amplamente usadas funções para medir a semelhança (ou diferença) entre os elementos. Do ponto de vista formal, este o problema de comparações pode ser adequadamente modelado pelas métricas introduzidas na Seção 3.4. Uma grande diversidade de métricas podem ser encontradas na literatura (consulte (DEZA; DEZA, 2009)), e essa variedade se dá pelo fato de que diversas áreas do conhecimento fazem uso destas funções, sendo algumas mais adequadas em determinados tópicos do que outras. Isto se justifica, pois diferentes distâncias podem refletir a similaridades entre dois objetos de maneiras distintas, dependendo muitas vezes do contexto, dos dados e da aplicação em questão.

Finalizaremos este capítulo apresentando a ideia resumida de tradução de imagens através de vetores de histogramas ou probabilidade. A base dessa representação de imagens será por meio de suas características intrínsecas, concentrando-se em um recurso valioso dado pela sua distribuição de cores.

Primeiramente, introduziremos os elementos fundamentais que modelam uma imagem digital:

- Um *pixel* é o menor elemento em uma imagem onde é possível atribuir uma cor;
- Uma cor é dada por um trio ordenado (r, g, b) , onde cada entrada fornece a tonalidade, respectivamente, das cores vermelho (r), verde (g) e azul (b);
- A tonalidade de uma cor é dada por um número inteiro que varia de 0 até 255, isto é, tons mais claros são dados por valores próximos ao 255, enquanto os tons mais escuros são valores próximos de 0.

Deste modo, uma imagem digital pode ser apresentada por uma malha de *pixels*, onde cada *pixel* contém a informação de uma cor, ou seja, um vetor de três tonalidades, ou números inteiros entre 0 e 255.

² Tradução livre do autor: Etiquetas

Por exemplo, uma imagem *Full HD*³ é uma malha de 1920×1080 vetores tridimensionais com entrada inteiras. Já as imagens em preto e branco, tem sua tradução simplificada, pois os vetores podem ser substituídos por um único número inteiro que varia de 0 (cor branca) até 255 (cor preta).

A partir do modelo matemático de uma imagem poderá ser construído uma representação gráfica da forma como os tons de cores são distribuídos em uma imagem digital. De acordo com Antonello (2018), um histograma de uma imagem é um gráfico de colunas que representa a distribuição dos valores dos *pixels*, ou seja, a porcentagem do total de *pixels* associada a cada valor/vetor tonal. Desta forma, temos um vetor de probabilidade de tamanho dado pela quantidade de cores (vetores) que existem na imagem cujas entradas fornecem a frequência desta cor (vetor) dentre todos os *pixels* da imagem.

Uma vez que traduzimos imagens em histogramas/vetores de probabilidade, é possível reproduzir os cálculos da seção para comparar tais imagens/histogramas/probabilidade com a distância Euclidiana, a métrica de Wasserstein e/ou outras distâncias.

³ Tradução livre do autor: Alta Definição Total

5 Conclusão

Neste Trabalho, trouxemos duas propostas (equivalentes) para abordagem do problema de transporte. A primeira análise ocorreu à luz de um importante e vasto ramo da Matemática, a Programação Linear, onde analisamos os custos de transportes associados à distribuição de produtos. Na segunda abordagem, lançamos mão da Teoria do Transporte Ótimo, que se mostrou uma ferramenta muito útil para nossa proposta de analisarmos propriedades geométricas entre histogramas (probabilidades) intrinsecamente relacionadas ao problema de transporte.

Do ponto de vista histórico, fomos capazes de evidenciar que o surgimento da Pesquisa Operacional veio de necessidades concretas em aplicações militares e que o seu desenvolvimento atual e permanente é fruto de problemas colocados pelas mais diversas áreas: Logística, Indústria, Economia, Computação, Matemática, etc. A análise da evolução histórica também nos revelou que a Programação Linear e a Teoria do Transporte Ótimo surgiram a partir da sagacidade de estudiosos como Monge, Kantorovich, Koopmans, Dantzing entre outros, que fizeram abordagens e métodos de resolução inovadores nessas teorias.

Vale reforçar que nossa abordagem da Programação Linear foi no âmbito do ensino da Matemática, fazendo referência às aplicações que esta tem e a importância em utilizá-las para estudar conceitos matemáticos. Em nosso Exemplo Motivador (Seção 2.2), ficou muito evidente a necessidade de utilizarmos a modelagem matemática para resolvermos problemas que envolvam deslocamentos de objetos e que tenhamos o objetivo de minimizar custos ou maximizar lucros da atividade de transporte concreta de um determinado produto. Além disso, notamos a grande importância do desenvolvimento da lógica e do raciocínio matemático durante a interpretação e resolução desses problemas. Ainda sobre este, foi muito significativo a aplicação da ferramenta *Solver*, através do *Microsoft Excel*, que se demonstrou ser uma alternativa eficaz para o tratamento dos dados na geração de soluções para o problema de transporte e que foi a principal ferramenta para o desenvolvimento de nossas aplicações.

Da mesma forma, trabalhamos com a Teoria de Transporte Ótimo (para probabilidades finitas) com a finalidade de desenvolver a capacidade de usar a Matemática como instrumento de interpretação e intervenção no mundo real, buscando uma multiplicidade de novas aplicações para o problema de transporte, ainda fazendo uso dos métodos de resolução da Programação Linear. Em nossa aplicação à Comparação de Histogramas da COVID-19 (Seção 4.1) foi possível identificar a clara similaridade existente entre as características das ondas, que podem ser modeladas por meio das funções matemáticas, e

comportamentos epidemiológicos. É evidente que não estudamos características inerentes das ondas como, frequência, comprimento, velocidade de propagação ou amplitude, mas comparamos o comportamento de dois histogramas de evolução da COVID-19, com funções que tem as características principais para a modelagem das ondas. Dessa forma, podemos dizer que o comportamento tanto da evolução do número de casos quanto do número de óbitos da COVID-19, no período que abordamos, possuem uma grande similaridade com o comportamento das ondas que são entendidas como sendo uma perturbação que se propaga num meio.

Salientamos a dificuldade inerente que muitas vezes há na própria formalização dos problemas, a qual advém da complexidade dos enunciados e suas interpretações. Mesmo assim, continuações naturais deste trabalho seriam:

- Estender os cálculos da Seção 4.1 para histogramas em outros países, outras doenças e outras funções de onda para comparação;
- Aplicar noções de *Data Clustering*¹ aos diversos dados da pandemia utilizando a métrica de Wasserstein, a qual acreditamos que possa fornecer importantes resultados.
- Investigar a Teoria de Transporte Ótimo para probabilidades na reta real \mathbb{R} (caso contínuo), cenário onde a Programação Linear não se aplica.

Ressaltamos que as nossas aplicações são exemplos simples e ingênuos da área de *Machine Learning*². Apesar de usualmente esta área explorar outros métodos (regressão linear ou logística, gradiente descendente, etc), a Teoria de Transporte Ótimo tem começado a despontar como ferramenta fundamental para o desenvolvimento de novas técnicas.

Em suma, atendendo a todos os aspectos aqui mencionados, acreditamos que este texto serve como uma preciosa ferramenta de complementação aos diversos trabalhos em Programação Linear já desenvolvidos no âmbito do PROFMAT e uma das poucas referências na literatura acadêmica em língua portuguesa que investiga o paralelismo entre a Programação Linear e a Teoria de Transporte Ótimo.

¹ Tradução livre do autor: Agrupamento de Dados

² Tradução livre do autor: Aprendizagem de Máquina

Referências

- AKGÜL, C. B. et al. Recuperação de imagem baseada em conteúdo em radiologia: status atual e direções futuras. *Jornal de Imagem Digital*, Springer, v. 24, n. 2, p. 208–222, 2011.
- ANTONELLO, R. *Introdução a Visão Computacional com Python e OpenCV*. 1. ed. Santa Catarina, 2018.
- BALLOU, R. H. *Logística empresarial: transportes, administração de materiais e distribuição física*. São Paulo: Atlas, 1993. ISBN 978-8522408740.
- BASSANEZI, R. C. *Ensino-aprendizagem com modelagem matemática: uma nova estratégia*. São Paulo: Editora Contexto, 2002. ISBN 978-8572442077.
- BAZARAA, M. S.; JARVIS, J. J.; SHERALI, H. D. *Linear programming and network flows*. New York: John Wiley & Sons, 2008. ISBN 978-0470462720.
- BORBA, M. C.; MENEGHETTI, R. C.; HERMINI, H. A. Estabelecendo critérios para avaliação do uso de modelagem em sala de aula: estudo de um caso em um curso de ciências biológicas. BORBA, MC *Calculadoras gráficas e educação matemática. Rio de Janeiro: USU, Ed. Bureau ou In: FAINGUELERNT, E. K.; GOTTLIEB, F. C. (Org.). Calculadoras Gráficas e Educação Matemática. Rio de Janeiro: Art Bureau, p. 95–113, 1999.*
- CAIXETA FILHO, J. V. *Pesquisa operacional: técnicas de otimização aplicadas a sistemas agroindustriais*. São Paulo: Atlas, 2001. ISBN 978-8522437344.
- DANTZIG, G. B. *Linear Programming and Extensions*. Santa Monica: Princeton university press, 1963.
- DANTZIG, G. B.; THAPA, M. N. *Linear programming 1: introduction*. New York: Springer Science & Business Media, 2006. ISBN 978-0387948331.
- DEZA, M. M.; DEZA, E. Encyclopedia of distances. In: *Encyclopedia of distances*. [S.l.]: Springer, 2009. p. 1–583.
- DUALIDADE. In: DICIONÁRIO Online de Português. Porto: 7Graus, 2022. Disponível em: <<https://www.dicio.com.br/dualidade/>>. Acesso em: 20 fev. 2022.
- GOLDBARG, M.; LUNA, H. *Otimização combinatória e programação linear: modelos e algoritmos*. Rio de Janeiro: Elsevier Brasil, 2005. v. 2. ISBN 978-8535215205.
- GOMES, J. T. A.; SILVA, J. d. L. *Histograma-COVID-19*. Repositório - GitHub, 2022. Disponível em: <<https://github.com/jtagomes/Histograma-COVID-19>>. Acesso em: 20 fev. 2022.
- GOMES JÚNIOR, A. d. C.; SOUZA, M. J. F. *Solver (Excel): Manual de referência*. Ouro Preto, 2004. Disponível em: <http://www.decom.ufop.br/marcone/Disciplinas/OtimizacaoCombinatoria/solver_p.pdf>. Acesso em: 12 set. 2021.

- GRANJA, D. I.; RUIZ, J. J. *História da Pesquisa Operacional*. 2006. Disponível em: <<http://www.phpsimplex.com/pt/historia.htm>>. Acesso em: 21 fev. 2022.
- HILLIER, F. S.; LIEBERMAN, G. J. *Introdução à pesquisa operacional*. Porto Alegre: McGraw Hill Brasil, 2013. ISBN 978-8580551181.
- KARMAKAR, N. A new polynomial-time algorithm for linear programming. *Combinatorica*, n. 4, p. 373–395, 1984.
- LIMA, E. L. *Espaços métricos*. Rio de Janeiro: IMPA, 2020. ISBN 978-6599052873.
- LIU, Y. et al. A survey of content-based image retrieval with high-level semantics. *Pattern recognition*, Elsevier, v. 40, n. 1, p. 262–282, 2007.
- LÓSS, Z. E. O desenvolvimento da pesquisa operacional no brasil. *PTS*, v. 10, p. 81, 1981.
- MARINS, F. A. S. Introdução à pesquisa operacional. *Cultura Acadêmica*, São Paulo, 2011.
- PAIVA, S. M. d. A. *A programação linear no ensino secundário*. Dissertação (Mestrado) — Universidade Portucalense Infante Don Henrique, Porto, 2008.
- PASSOS, E. J. P. F. d. *Programação linear: como instrumento da pesquisa operacional*. São Paulo: Editora Atlas, 2008. ISBN 978-8522448395.
- PEYRÉ, G.; CUTURI, M. et al. Computational optimal transport: with applications to data science. *Foundations and Trends® in Machine Learning*, Now Publishers, Inc., v. 11, n. 5-6, p. 355–607, 2019.
- PLOSKAS, N.; SAMARAS, N. Gpu acelerado pivot regras para o algoritmo simplex. v. 96, p. 1–9, 2014.
- POLYA, G. *A arte de resolver problemas*. Rio de Janeiro: Interciência, 1978. ISBN 978-8571931367.
- ROSS, S. *Probabilidade: um curso moderno com aplicações*. Porto Alegre: Bookman, 2009. ISBN 978-8577806218.
- SAPPELLI, M.; VERBERNE, S.; KRAAIJ, W. Avaliação de sistemas de recomendação cientes do contexto para reencontro de informações. *Jornal da Associação de Ciência e Tecnologia da Informação*, Wiley Online Library, v. 68, n. 4, p. 895–910, 2017.
- THORPE, M. *Introduction to optimal transport*. Cambridge, 2018. Disponível em: <http://www.damtp.cam.ac.uk/research/cia/files/teaching/Optimal_Transport_Notes.pdf>. Acesso em: 12 set. 2021.
- VILLANI, C. *Topics in optimal transportation*. Providence: American Mathematical Society, 2003. ISBN 978-0821833124.
- VILLANI, C. *Optimal transport: old and new*. Berlin: Springer, 2009. ISBN 978-3662501801.

WEBER, R. F. F. *Programação Linear no Ensino Médio: um estudo dos modelos de transporte com uma proposta para deficientes visuais*. Dissertação (Mestrado) — Universidade Estadual de Campinas, Campinas, 2018.

WILSON, D. R.; MARTINEZ, T. R. Funções de distância heterogênea aprimoradas. *Jornal de Pesquisa de Inteligência Artificial*, v. 6, p. 1–34, 1997.

WOLFRAM|ALPHA. Champaign: Wolfram Research, Inc., 2022. Disponível em: <<https://www.wolframalpha.com/>>. Acesso em: 20 fev. 2022.

Apêndices

APÊNDICE A – Dualidade

Etimologicamente, a palavra dualidade deriva do latim “dualitas, atis”, com o mesmo sentido “qualidade do que é duplo”. Em (DUALIDADE, 2022), esta é definida como sendo a qualidade do que contém em sua essência duas substâncias, dois princípios, duas naturezas. Esse significado amplo, gerou várias aplicações em diversos campos de conhecimento e ganhou ocupações diferentes até dentro de uma mesma área. Vejamos abaixo algumas aplicações desse termo:

- Na psicologia, pode ser entendida como sendo o princípio ideológico segundo o qual é possível haver a coexistência simultânea de coisas antagônicas dentro de um mesmo objeto (a matéria e o espírito; o corpo e a alma; o bem e o mal). Uma bem conhecida versão é a dualidade mente–corpo, atribuída a René Descartes (1641), segundo a qual a mente é uma substância não física identificada com a consciência, distinguindo-a do cérebro, considerado como o assento da inteligência.
- Para a linguística, dualidade expressa a particularidade da categoria número que, quando presente em determinadas línguas, representa duas entidades isoláveis, por oposição a singularidade e a pluralidade;
- Em física, existe a dualidade onda-partícula, propriedade dos entes físicos possuírem tanto comportamento de partículas como comportamento de ondas, e também é conhecida como dualidade matéria-energia.

Este termo ocupa várias facetas em quase todos os diferentes tópicos do conhecimento matemático. Como de origem do próprio nome, as dualidades em teorias matemáticas estabelecem algum tipo de relação entre objetos que correspondem a pares, por exemplo:

- Para Lógica Proposicional, os conectivos lógicos (\wedge) e (\vee) são duais. Dadas P e Q sentenças lógicas, destacamos as equivalências

$$\begin{aligned}\neg(P \wedge Q) &\iff \neg P \vee \neg Q \\ \neg(P \vee Q) &\iff \neg P \wedge \neg Q;\end{aligned}$$

- Na Álgebra Booleana, o princípio da dualidade diz que cada expressão ou identidade algébrica dedutível a partir dos postulados em uma Álgebra Booleana, continua válida se todas as ocorrências dos operadores $+$ e \cdot e dos elementos 0 e 1 são trocados um pelo outro;

- A dualidade de Álgebra Linear diz respeito a análise específica da relação entre um espaço vetorial e o seu espaço dual. A saber, se X é um espaço vetorial (sobre o corpo \mathbb{K}) de dimensão finita, consideremos o conjunto dos funcionais lineares $l: X \rightarrow \mathbb{K}$ que tem uma estrutura de espaço vetorial para as operações de soma de funções e multiplicação por um escalar. Denotamos por $X' = \{l: X \rightarrow \mathbb{K}: l \text{ é linear}\}$ o espaço dual de X ;
- No contexto da Topologia, as noções de conjunto aberto e fechado são duais, pois vale a seguinte equivalência: A é aberto de X , se e somente, se seu conjunto complementar A^c é fechado em X .

Programação Linear

A noção de dualidade na programação linear significa a existência de um segundo modelo, chamamos de problema dual, associado ao problema original, que será chamado de problema primal.

No problema primal, de cada ponto sub-ótimo que satisfaça todas as restrições, há uma direção ou subespaço de direções que aumentará o valor a função objetivo. Ao mover-se em qualquer uma dessas direções, remove-se a folga entre a solução candidata e uma ou mais restrições. Um valor inviável da solução candidata é aquele que excede uma ou mais das restrições.

No problema dual, o vetor dual multiplica as restrições que determinam as posições das restrições no primal. Variar o vetor dual no problema dual é equivalente a revisar os limitantes superiores do problema primal. O limitante superior mais baixo é procurado, ou seja, o vetor dual é minimizado para remover a folga entre as posições candidatas das restrições e o ótimo real. Um valor inviável do vetor dual é aquele que é muito baixo, o qual define as posições candidatas de uma ou mais restrições em uma posição que exclui o ótimo real.

Dessa forma, a transição de um problema primal para um problema dual, ou vice-versa, deverá seguir as seguintes regras:

- Se o primal for um problema de maximização (ou minimização), então o problema dual será um problema de minimização (ou maximização);
- Cada variável do primal está associada a uma restrição no dual e cada restrição do primal está associada a uma variável do dual;
- Os coeficientes de custo da função objetivo do primal correspondem aos termos independentes das restrições do dual;

- Os termos independentes das restrições do primal correspondem aos coeficientes de custo da função objetivo do dual;
- A transposta da matriz dos coeficientes tecnológicos do primal é a matriz dos coeficientes tecnológicos do dual.

Abaixo, rerepresentamos o problema primal (recorte do Quadro 3) e, de acordo com as regras descritas acima, exibimos o problema dual correspondente:

Quadro 7 – Problema Primal

<p>Função Objetivo</p> <p style="text-align: center;">Maximizar (ou Minimizar) $Z = CX$</p> <p>Sujeita às Restrições</p> <p style="text-align: center;">$AX \leq B, \quad X \geq 0 \quad \text{e} \quad B \geq 0$</p> <p>Onde</p> $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}, \quad C = (c_1, \cdots, c_m)$

Fonte: Elaborado pelo autor.

Quadro 8 – Problema Dual

<p>Função Objetivo</p> <p style="text-align: center;">Minimizar (ou Maximizar) $D = B^T Y$</p> <p>Sujeita às Restrições</p> <p style="text-align: center;">$A^T Y \geq C, \quad Y \geq 0 \quad \text{e} \quad C \geq 0$</p> <p>Onde</p> $A^T = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}, \quad B^T = (b_1, \cdots, b_n), \quad C = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}$

Fonte: Elaborado pelo autor.

APÊNDICE B – Vetores de Probabilidade da Seção 4.1

Vetor de probabilidade associado ao Histograma de Casos da COVID- 19
(Tabela 2 e Figura 14):

$$h_{\text{Casos}} = \begin{pmatrix} 0,00097 \\ 0,00633 \\ 0,02421 \\ 0,04234 \\ 0,05296 \\ 0,05325 \\ 0,04089 \\ 0,03146 \\ 0,03148 \\ 0,05376 \\ 0,05875 \\ 0,06345 \\ 0,07630 \\ 0,09395 \\ 0,07966 \\ 0,08250 \\ 0,08592 \\ 0,05507 \\ 0,03797 \\ 0,02879 \end{pmatrix}$$

Vetores de probabilidade associados a decomposição do Histograma de Casos da COVID-19 (Figura 6):

$$a_{1-10} = \begin{pmatrix} 0,00286 \\ 0,01875 \\ 0,0717 \\ 0,12541 \\ 0,15686 \\ 0,1577 \\ 0,12111 \\ 0,09316 \\ 0,09323 \\ 0,15922 \end{pmatrix} \quad a_{2-11} = \begin{pmatrix} 0,01600 \\ 0,06122 \\ 0,10709 \\ 0,13394 \\ 0,13465 \\ 0,10342 \\ 0,07955 \\ 0,07961 \\ 0,13596 \\ 0,14856 \end{pmatrix} \quad a_{3-12} = \begin{pmatrix} 0,05349 \\ 0,09357 \\ 0,11703 \\ 0,11766 \\ 0,09037 \\ 0,06951 \\ 0,06956 \\ 0,1188 \\ 0,12981 \\ 0,1402 \end{pmatrix}$$

$$a_{4-13} = \begin{pmatrix} 0,08392 \\ 0,10495 \\ 0,10551 \\ 0,08104 \\ 0,06233 \\ 0,06238 \\ 0,10653 \\ 0,11641 \\ 0,12572 \\ 0,1513 \end{pmatrix} \quad a_{5-14} = \begin{pmatrix} 0,09522 \\ 0,09573 \\ 0,07352 \\ 0,05655 \\ 0,05659 \\ 0,09665 \\ 0,10561 \\ 0,11406 \\ 0,13717 \\ 0,1689 \end{pmatrix} \quad a_{6-15} = \begin{pmatrix} 0,09134 \\ 0,07015 \\ 0,05396 \\ 0,054 \\ 0,09222 \\ 0,10078 \\ 0,10884 \\ 0,13089 \\ 0,16117 \\ 0,13665 \end{pmatrix}$$

$$a_{7-16} = \begin{pmatrix} 0,0668 \\ 0,05138 \\ 0,05142 \\ 0,08782 \\ 0,09596 \\ 0,10364 \\ 0,12463 \\ 0,15347 \\ 0,13012 \\ 0,13476 \end{pmatrix} \quad a_{8-17} = \begin{pmatrix} 0,04786 \\ 0,0479 \\ 0,0818 \\ 0,08939 \\ 0,09654 \\ 0,1161 \\ 0,14295 \\ 0,1212 \\ 0,12553 \\ 0,13073 \end{pmatrix} \quad a_{9-18} = \begin{pmatrix} 0,04624 \\ 0,07896 \\ 0,08629 \\ 0,09319 \\ 0,11207 \\ 0,138 \\ 0,117 \\ 0,12117 \\ 0,1262 \\ 0,08088 \end{pmatrix}$$

$$a_{10-19} = \begin{pmatrix} 0,07822 \\ 0,08547 \\ 0,09231 \\ 0,11101 \\ 0,13669 \\ 0,1159 \\ 0,12003 \\ 0,12501 \\ 0,08012 \\ 0,05524 \end{pmatrix} \quad a_{11-20} = \begin{pmatrix} 0,08868 \\ 0,09579 \\ 0,1152 \\ 0,14185 \\ 0,12027 \\ 0,12456 \\ 0,12972 \\ 0,08314 \\ 0,05733 \\ 0,04346 \end{pmatrix}$$

Vetor de probabilidade associado ao Histograma de Óbitos da COVID-19 (Tabela 3 e Figura 15):

$$h_{\text{Óbitos}} = \begin{pmatrix} 0,00189 \\ 0,01599 \\ 0,04258 \\ 0,04769 \\ 0,04931 \\ 0,04525 \\ 0,03526 \\ 0,02608 \\ 0,02034 \\ 0,02897 \\ 0,03886 \\ 0,04920 \\ 0,06489 \\ 0,12495 \\ 0,11778 \\ 0,08619 \\ 0,08592 \\ 0,05517 \\ 0,03810 \\ 0,02556 \end{pmatrix}$$

Vetores de probabilidade associados a decomposição do Histograma de Óbitos da COVID-19 (Figura 6):

$$d_{1-10} = \begin{pmatrix} 0,00603 \\ 0,05104 \\ 0,13589 \\ 0,15217 \\ 0,15734 \\ 0,14441 \\ 0,11253 \\ 0,08323 \\ 0,06491 \\ 0,09245 \end{pmatrix} \quad d_{2-11} = \begin{pmatrix} 0,04565 \\ 0,12155 \\ 0,13611 \\ 0,14074 \\ 0,12917 \\ 0,10066 \\ 0,07444 \\ 0,05806 \\ 0,0827 \\ 0,11092 \end{pmatrix} \quad d_{3-12} = \begin{pmatrix} 0,11103 \\ 0,12433 \\ 0,12855 \\ 0,11798 \\ 0,09194 \\ 0,068 \\ 0,05303 \\ 0,07554 \\ 0,10132 \\ 0,12828 \end{pmatrix}$$

$$d_{4-13} = \begin{pmatrix} 0,11749 \\ 0,12149 \\ 0,1115 \\ 0,08689 \\ 0,06426 \\ 0,05012 \\ 0,07139 \\ 0,09575 \\ 0,12123 \\ 0,15988 \end{pmatrix} \quad d_{5-14} = \begin{pmatrix} 0,10206 \\ 0,09367 \\ 0,07299 \\ 0,05398 \\ 0,0421 \\ 0,05997 \\ 0,08044 \\ 0,10184 \\ 0,13432 \\ 0,25863 \end{pmatrix} \quad d_{6-15} = \begin{pmatrix} 0,08204 \\ 0,06393 \\ 0,04728 \\ 0,03688 \\ 0,05253 \\ 0,07045 \\ 0,0892 \\ 0,11765 \\ 0,22652 \\ 0,21352 \end{pmatrix}$$

$$d_{7-16} = \begin{pmatrix} 0,05952 \\ 0,04402 \\ 0,03433 \\ 0,0489 \\ 0,06558 \\ 0,08304 \\ 0,10952 \\ 0,21087 \\ 0,19877 \\ 0,14545 \end{pmatrix} \quad d_{8-17} = \begin{pmatrix} 0,04055 \\ 0,03162 \\ 0,04505 \\ 0,06042 \\ 0,0765 \\ 0,10089 \\ 0,19426 \\ 0,18311 \\ 0,13401 \\ 0,13359 \end{pmatrix} \quad d_{9-18} = \begin{pmatrix} 0,03025 \\ 0,0431 \\ 0,0578 \\ 0,07319 \\ 0,09653 \\ 0,18586 \\ 0,17519 \\ 0,12821 \\ 0,12781 \\ 0,08206 \end{pmatrix}$$

$$d_{10-19} = \begin{pmatrix} 0,04199 \\ 0,05632 \\ 0,0713 \\ 0,09404 \\ 0,18107 \\ 0,17068 \\ 0,12491 \\ 0,12452 \\ 0,07995 \\ 0,05522 \end{pmatrix} \quad d_{11-20} = \begin{pmatrix} 0,0566 \\ 0,07166 \\ 0,09451 \\ 0,18197 \\ 0,17153 \\ 0,12553 \\ 0,12514 \\ 0,08035 \\ 0,05549 \\ 0,03722 \end{pmatrix}$$

Vetor de probabilidade associado ao Histograma do Cosseno (Figura 17):

$$b_{\cos} = \begin{pmatrix} 0,02447 \\ 0,07102 \\ 0,11062 \\ 0,13938 \\ 0,15451 \\ 0,15451 \\ 0,13938 \\ 0,11062 \\ 0,07102 \\ 0,02447 \end{pmatrix}$$

Vetor de probabilidade associado ao Histograma da Cúspide Cúbica (Figura 19):

$$b_{\text{cusp}} = \begin{pmatrix} 0,01708 \\ 0,05328 \\ 0,09317 \\ 0,13894 \\ 0,19753 \\ 0,19753 \\ 0,13894 \\ 0,09317 \\ 0,05328 \\ 0,01708 \end{pmatrix}$$

Observação. Com o objetivo de manter o padrão de 5 casas decimais e tendo vista o arredondamento das entradas dos vetores acima, algumas entradas tiveram que ser ajustadas para que todos os vetores sejam sempre vetores de probabilidade. Observamos que tais ajustes foram pontuais e ocorreram na quinta casa decimal, desta forma, não geram erros muito grandes nos cálculos realizados neste trabalho.