

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

VICTOR GAROFANI NASIMOTO

**INTRODUÇÃO AOS CONCEITOS DA MATEMÁTICA SUPERIOR EM
DISCIPLINAS DO ENSINO MÉDIO**

CURITIBA

2025

VICTOR GAROFANI NASIMOTO

**INTRODUÇÃO AOS CONCEITOS DA MATEMÁTICA SUPERIOR EM
DISCIPLINAS DO ENSINO MÉDIO**

Introdution to the Concepts of Higher Mathmatics in High School Disciplines

Trabalho de Conclusão de Curso de Graduação
apresentado (a) como requisito para obtenção
do título de Licenciado em do Curso
de Matemática da Universidade Tecnológica
Federal do Paraná.
Orientador(a): Dr. José Carlos Pereira Coninck

CURITIBA

2025



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Esta licença permite download e compartilhamento do trabalho desde que sejam atribuídos créditos ao(s) autor(es), sem a possibilidade de alterá-lo ou utilizá-lo para fins comerciais. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

VICTOR GAROFANI NASIMOTO

**INTRODUÇÃO AOS CONCEITOS DA MATEMÁTICA SUPERIOR EM
DISCIPLINAS DO ENSINO MÉDIO**

Trabalho de Conclusão de Curso de Graduação
apresentado (a) como requisito para obtenção
do título de Licenciado em do Curso
de Matemática da Universidade Tecnológica
Federal do Paraná.

Data de aprovação: 24 / junho / 2025

Jose Carlos Pereira Coninck
Dr.
Universidade Tecnológica Federal do Paraná

Patricia Sáñez Pacheco)
Dra.
Universidade Tecnológica Federal do Paraná

Vanessa Ferreira Sehaber)
Dra.
Universidade Tecnológica Federal do Paraná

**CURITIBA
2025**

Dedico este trabalho à minha família, pelos
momentos de ausência.

“A biblioteca é um jardim onde as ideias
florescem e os frutos são colhidos pela
eternidade.” (Candido, 2002).

RESUMO

Este trabalho tem o propósito de apresentar parte da teoria das regressões múltiplas como mecanismo de utilização na introdução dos conceitos da matemática superior aos estudantes do Ensino Técnico Profissionalizante como um material para o professor destes cursos. O propósito visa expandir o conhecimento da regressão linear simples para a regressão linear múltipla no primeiro plano. Em segundo, ampliar o alcance teórico e prático, tanto para o professor que ministrará a aula, quanto ao estudante que necessita da ferramenta com teor mais profundo e formal sobre a teoria da regressão estatística. O trabalho foi desenvolvido para atender aos cursos técnicos e profissionalizantes que necessitam de uma base mais profunda e técnica, das quais envolvem ajustes, leituras e interpretações de medidas e dados coletados em laboratório de testes ou em campo.

Palavras-chave: estatística; análise de regressão múltipla; avaliações educacionais; ensino técnico profissionalizante.

ABSTRACT

This paper aims to present part of the theory of multiple regressions as a mechanism to be used in introducing higher mathematics concepts to technical education students, as reading material for teachers of these courses. The purpose is to expand the knowledge of simple linear regression to multiple linear regression in the foreground. Secondly, to broaden the theoretical and practical scope, both for the teacher who will teach the class and for the student who needs the tool with a deeper and more formal content on the theory of statistical regression. The work was developed to meet the needs of technical and professional courses that refer to a deeper and more technical basis, which involve adjustments, readings and interpretations of measurements and data found in the test laboratory or in the field.

Keywords: statistics; multiple regression analysis; educational assessments; vocational technical education.

LISTA DE TABELAS

Tabela 1 – Análise de Variância (ANOVA).....	36
Tabela 2 – ANOVA para o modelo completo	39
Tabela 3 – ANOVA para o modelo reduzido	39
Tabela 4 – ANOVA para o modelo completo - SQ da regressão particionada	39
Tabela 5 – A forma experimental.	48
Tabela 6 – Tabela auxiliar para cálculo.....	54
Tabela 7 – Tabela taxa de fertilidade por países da América do Sul.....	55
Tabela 8 – Tabela experimentação 2^k	60
Tabela 9 – Tabela ANOVA para 2^k	63
Tabela 10 – Tabela avaliação dos coeficientes β_i	64

LISTA DE ABREVIATURAS E SIGLAS

CDI	Cálculo Diferencial e Integral
EDO	Equações Diferenciais Ordinárias
IDEB	Índice de Desenvolvimento da Educação Básica
IBGE	Instituto Brasileiro de Geografia e Estatística
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira

SUMÁRIO

1	INTRODUÇÃO	10
2	OBJETIVO	12
3	JUSTIFICATIVA	13
4	METODOLOGIA	14
5	FUNDAMENTAÇÃO TEÓRICA	15
5.1	Perspectivas do novo ensino médio	15
5.2	Exigências do mercado de trabalho	17
5.3	DISCIPLINAS DO ENSINO SUPERIOR NO ENSINO MÉDIO.....	18
5.3.1	Cálculo Diferencial e Integral	18
5.3.2	Probabilidade e Estatística	18
5.3.3	Álgebra Linear e Geometria Analítica	18
5.3.4	Pensamento Lógico Computacional e Programação	19
5.3.5	Educação Financeira e Princípios de Contabilidade Geral	19
6	RESULTADOS E DISCUSSÃO	20
6.1	Detalhamento matemático nas regressões estatísticas	20
6.1.1	Regressão Linear vetorial.....	20
6.1.2	Método dos Mínimos Quadrados Ordinários (na forma de somatório)	21
6.1.2.1	Utilização no ensino médio: o efeito da otimização.	24
6.1.2.2	Método dos Mínimos Quadrados Ordinários (Forma Vetorial).....	24
6.1.3	Encontrando o Ponto de Mínimo da Função $y = \beta_0 + \beta_1 x + \epsilon$	25
6.1.3.1	Utilização no ensino médio: o determinante da matriz	26
6.1.3.2	Soma dos Quadrados	27
6.1.3.3	Independência Linear	29
6.1.3.4	Teorema de Pitágoras	30
6.1.3.5	Utilização no ensino médio: o impacto do teorema de Pitágoras dentro do ajuste	31
6.1.4	Desenvolvimento MMQO	31
6.1.4.1	Utilização no ensino médio: a importância da inversa de uma matriz	32
6.1.5	Projeções e Mínimos Quadrados	33
6.1.6	A Esperança Matemática	33
6.2	Intervalos de confiança	35
6.2.1	Análise dos resíduos	41
6.2.2	Validação dos modelos múltiplos	44
6.2.2.1	Teste de hipótese para a inclinação β_i :	44
6.2.2.2	Teste de hipótese para o intercepto β_0 :	45
6.3	Método dos Mínimos Quadrados Generalizado	46
6.4	Planejamento de Experimentos: fatorial 2^k	48
6.4.1	Experimento 2^k	48
6.4.2	Relação com modelos de regressão	49
6.5	Aplicações.....	53
6.5.1	Exemplos de aplicações dos MLMs	53
7	CONCLUSÃO.....	66
	REFERÊNCIAS.....	67
	ANEXO A – Lei N.º 9.610, de 19 de Fevereiro de 1998.....	69

1 INTRODUÇÃO

Dos desafios já enfrentados na educação, um dos mais recentes se trata das rápidas mudanças de paradigmas tecnológicos. As exigências não estão mais vinculadas ao conhecimento de um segundo idioma e/ou conhecimentos básicos na simples operação de um sistema operacional ou na utilização de um banco de dados específico. Além desses requisitos, o novo conceito de analfabetismo digital (Santana *et al.*, 2023) está agora vinculado à forma da utilização das tecnologias IOT (Internet Of Things) e o uso adequado das inteligências artificiais. Cabe ressaltar ao leitor que, historicamente, a revolução industrial na segunda metade do século XVIII, não acabou com os empregos, mas o modificou com novas práticas e métodos de produção. Que a popularização dos computadores pessoais, a internet e o acesso à informação na década de 90, do século XX, não acabou o trabalho do terceiro setor, mas o transformou na forma de gerar, transferir e trabalhar com as informações. Outrora os jornais eram impressos e hoje são lidos digitalmente em plataformas, o dinheiro de papel não sumiu, mas está menos usual sendo gradativamente substituído por pagamentos on-line, serviços públicos estão dentro de plataformas digitais como SouGov ou equivalentes. A própria mídia, antes dominada por televisão na sala ou o rádio na cozinha da casa, atualmente perde espaço para a tecnologia streaming por empresas como Netflix, Amazon Prime e Google, sem mencionar os assistentes virtuais alimentado por inteligência artificial, tais como Alexa que ajudam a administrar a casa. Enquanto isso, os trabalhadores do setor secundário e de transformação que operam máquinas e equipamentos de forma manual, tais como tornos, empilhadeiras, esteiras em linhas de produção, estão sendo "empurrados" aos poucos por atividades atualizadas para sistemas autômatas em que a máquina toma a própria decisão, ficando a cargo do operador o serviço de vigilância e responsabilidade do controle. As atividades não desaparecerão, mas serão seriamente modificadas no seu modo operante. Por exemplo, técnicos em solda seguem com a atividade manual sendo gradativamente adaptada para equipamentos controlados roboticamente por humano ou I.A. quando necessárias atividades de larga escala, precisão, risco ou complexidade (*Remote Control Welding Machines*). Nesse exemplo, o operador não deixará de ser um soldador manual, pois além de obter o conhecimento da técnica de soldagem para a devida operação e objetivo final da robotização, deverá também estar preparado para ocasionais manutenções ou ajustes. Isto é, maior será a carga de conhecimento sobre o profissional, mais conhecimento necessitará, ampliação do eixo de trabalho, mais informação, mais responsabilidade. O exemplo da programação robótica ou da inspeção de máquinas autoprogramáveis via linguagem de máquina é uma realidade sem volta. De qualquer forma, os novos postos de trabalho que geram essas matrizes, com técnicas, processos e serviços, serão incorporados por novos conceitos dentro e fora do mercado de trabalho, independente do ramo, com mais ou menos inserção da inteligência artificial tanto em equipamentos profissionais que usam sistemas embarcados quanto a simples utilização de um site ou serviço pessoal ou profissional de comunicação, como o WhatsApp da empresa Meta. Não o bastante, a quantidade de novos métodos, tecnologias, técnicas, leis e avanços tecnológicos disponíveis nesse século é avassa-

ladora. E nesse sentido, o que esse Trabalho de Conclusão de Curso apresenta é uma forma mais enfática e imersiva de abordagem no ensino da matemática dentro das escolas técnicas e profissionalizantes, pois mais necessários serão os conceitos do ensino superior incorporados dentro dos cursos técnicos de nível médio, dado a necessidade do avanço tecnológico, por intermédio do professor.

2 OBJETIVO

O objetivo principal é usar técnicas estatísticas para introduzir conceitos do ensino superior nos cursos técnicos e profissionalizantes, apresentando elementos relevantes da regressão múltipla e aplicando conceitos da Matemática superior utilizando regressão múltipla.

Objetivos específicos

- Mostrar visualmente a minimização de funções de duas variáveis.
- Demonstrar aplicações de multiplicação de matrizes.
- Desenvolver o Método dos Mínimos Quadrados.

3 JUSTIFICATIVA

Com o rápido avanço das tecnologias, muitos aspectos da análise de dados surgiram quase que imperceptivelmente no cotidiano das pessoas, principalmente nas exigências dos novos profissionais, que passaram a exigir um novo idioma para além do idioma, isto é, conhecimentos em estatística, banco de dados, programação e inteligência artificial. Esse trabalho apresentamos uma proposta de ensino que viabilize o ensino e aprendizado de conceitos da matemática superior no ensino profissionalizante e técnico, deixando as tecnologias utilizadas pelos futuros profissionais mais acessíveis teoricamente e não apenas mecanizadas no cotidiano do técnico.

4 METODOLOGIA

O presente estudo contou com uma metodologia quantitativa, aplicada, descritiva. A revisão feita aqui de caráter narrativo buscou elucidar o contexto educacional em que o Ensino Profissionalizante se encontra e o desenvolvimento teórico da Matemática envolvida.

Primeiramente desenvolvemos a teoria da regressão múltipla incluindo conceitos matemáticos em algumas etapas cruciais. Após a incursão desses conceitos exploramos alguns exemplos práticos podendo ser utilizado em sala de aula. O conceito de regressão não precisa ser utilizado apenas como uma técnica de ajustes de curvas, mas também como ferramenta para apresentar conceitos do nível superior aos estudantes dos cursos técnicos e profissionalizantes. Dessa forma, além da praticidade visual, os resultados auxiliam na explicação dos conceitos mais avançados.

5 FUNDAMENTAÇÃO TEÓRICA

5.1 Perspectivas do novo ensino médio

Alguns trabalhos discutem o motivo da exclusão do conteúdo de Cálculo Diferencial e Integral (CDI) no ensino médio (Borgo, 2019). Apesar de não presente tanto na Base Nacional Comum Curricular (Brasil, 2018) e não ser contemplado diretamente para o projeto do Novo Ensino Médio até o momento, diversas discussões sobre o assunto já foram colocadas em pauta sobre seu ensino no Ensino Médio (Rocha, 2018) e no Ensino Técnico e Profissionalizante (Pagani, 2014). Ao longo desses estudos, percebeu-se que, principalmente no Ensino Técnico e Profissionalizante, não apenas os professores e alunos compreendem a importância do CDI no seu processo de ensino-aprendizagem, como são capazes de identificar objetivamente os aspectos do CDI que são efetivamente utilizados nos cursos (Pagani, 2014).

No que diz respeito ao Ensino Médio, a utilização e ensino do cálculo encontram mais percalços. Além de o currículo da Matemática ser extenso, autores já discutiram a maturidade dos estudantes (Rocha, 2018) para abstrair conteúdos complexos como limites, sem ter aplicações diretas. Para lidar com essas dificuldades, sem abrir mão da compreensão preliminar do assunto, pesquisadores buscaram utilizar a própria curiosidade dos estudantes para levantar questões que os motivem a buscar respostas no CDI (Dreyfus anatoli kouropatov, 2021). É possível que, mesmo sem abordá-los formalmente, sejam ensinados conceitos como limites, derivadas e integrais, explorando ideias intuitivas dos alunos e levantando dúvidas pertinentes à maturidade matemática deles. Dessa forma, os alunos podem ser capazes de entenderem conceitos como derivada por meio de questões como pontos de máximo ou mínimo de funções.

A Estatística, por sua vez, está prevista no ensino brasileiro (Brasil, 2018) dentro da unidade temática *Probabilidade e Estatística*. Em comparação ao CDI, a Estatística já apresenta mais possibilidades de ser incorporada, não só puramente na Matemática, mas na abordagem dos temas contemporâneos transversais (Cazorla, 2021), não ficando restrita à apresentação de dados aos alunos, mas para que consigam elaborar seus próprios gráficos, fazer suas próprias previsões e formular tomadas de decisão baseadas em dados. Entretanto, apenas saber observar dados em gráficos e tabelas não é o suficiente para uma educação estatística de qualidade. Existem diversas falácias e equívocos (Filho; Leão, 2023) que podem levar a tomadas erradas de decisão em diversas esferas da sociedade. Apesar de a coleta e interpretação equivocada dos dados não ser sempre intencional, o conhecimento mais completo possível de probabilidade e estatística é a forma mais consistente de um cidadão contornar esses problemas. Portanto, a formação sólida do professor e dos estudantes é necessária para que a análise de dados seja uma ferramenta de auxílio em todas as áreas do conhecimento.

Na proposta do Novo Ensino Médio (Silva; Pasqualli; Spessatto, Recuperado de <https://periodicos.ufff.br/index.php/edufoco/article/view/39210>, 2023), cujo ponto central é a flexibilização do currículo escolar, diversas disciplinas foram abertas na maioria das escolas do Brasil. Ao observar o catálogo de disciplinas eletivas da SEEDF (Federal, 2025). pode-se per-

ceber rapidamente uma boa quantidade de disciplinas que teriam seu plano de ensino enriquecido com educação estatística, tanto dentro da área da Matemática quanto das Ciências da Natureza, como "Agroecologia", "A Incrível Máquina Humana", "Química Ambiental", entre outras.

Outra interface de interação direta entre o estudante e a estatística é a inteligência artificial. Apesar de inteligência artificial poder ser definida como "Um sistema computacional que simula a capacidade do ser humano em resolver problemas" (Casagrande, 2019), comumente chamamos de inteligência artificial apenas algoritmos de redes neurais. Os algoritmos de redes neurais utilizam técnicas estatísticas em grandes volumes de dados (Haykin, 2001), o que torna sua compreensão indissociável da compreensão da estatística. Com todas essas interações entre estudantes e estatística, é razoável considerar que o sistema de educação deva não só contemplar a educação estatística, como constantemente atualizar os tópicos mais pertinentes às tecnologias atuais.

Segundo o relatório da OCDE, "*PISA Vocational Education and Training (VET): Assessment and analytical framework*" (Ocde, 2024), além da avaliação tradicional internacional de aprendizagem *Programme for International Student Assessment - PISA*, aplicada aos eixos da matemática, leitura e ciências aos alunos de 15 anos, a nova modalidade está próxima de ser implementada: a modalidade da qualidade de aprendizagem em cursos técnicos e profissionalizantes. Nesse documento, inicia-se experimentalmente com três categorias profissionais: técnicos em eletricidade, técnicos automotivos e técnicos em hotelaria. Nos três casos, a estatística está presente, tanto no contexto teórico como prático. Segundo o documento VET (Ocde, 2024), podemos citar os técnicos em eletricidade, os quais necessitam compreender o conhecimento da teoria e dos fundamentos elétricos. No entanto, o aluno precisa ter um profundo conhecimento de aritmética, álgebra, geometria, cálculo, estatística e suas aplicações. No serviço de hotéis, o técnico em hotelaria coletam e interpretam informações estatísticas sobre os planos comerciais de produtos e serviços do estabelecimento hoteleiro, uso de planilhas e interpretação estatística que envolve logística e administração hoteleira. Os técnicos automotivos necessitam em trabalhar com medidas mecânicas que envolvem variações e calibrações. Em todos os casos o documento enfatiza a necessidade do conhecimento em estatística para a devida aplicação do trabalho ou serviço. O projeto VET é uma iniciativa do Programa Internacional de Avaliação de Estudantes (PISA) para Educação e Formação Profissional (EFP), desenvolvida pela OCDE, visa mensurar as competências dos alunos que estão prestes a concluir os seus programas de educação e formação profissional em áreas ocupacionais selecionadas. Esse exemplo mostra-nos o quanto a análise estatística impactará nas transições das novas carreiras, das quais dependerão mais profundamente do tratamento dos dados. Isso demonstra que a educação deverá se adaptar o mais breve possível às mudanças significativas na forma, conceitos e paradigmas do mercado de trabalho global. Quando pensamos no estudante do ensino técnico e profissional, o conceito do PISA/VET é importante como avaliação. De fato, não é a obrigatoriedade da empresa em ensiná-los, mas sim selecioná-los para as atividades empresariais, sejam elas quais forem.

5.2 Exigências do mercado de trabalho

Segundo o relatório *The Future of Jobs Report* ((Forum, 2020)), os três trabalhos com maior tendência de crescimento até 2025 seriam especialistas em Big Data, inteligência artificial e cientistas de dados, todas sendo profissões indissociáveis do aprendizado de métodos estatísticos. O relatório de 2025 trouxe novamente tendências em inteligência artificial e Big Data, prevendo o crescimento do uso dessas tecnologias no mundo entre 2025 e 2030, o que pode se manter nas próximas décadas também. Grandes volumes de dados observacionais são gerados e armazenados todos os dias e devem ter seu volume diário aumentado nos próximos anos (Statista, 2020). O gerenciamento desses dados por empresas, governos, instituições e pesquisadores é feito por métodos estatísticos. Os dados são capazes de gerar, após a manipulação, grandezas quantitativas que auxiliam os pesquisadores a tomar decisões, além de realizar descobertas que dependam de enxergar relações minuciosas entre os dados. Revistas científicas como a *Remote Sensing* publicam constantemente estudos com manipulação de dados espaciais, normalmente gerados por satélites, que criam diversos modelos matemáticos para prever diversas características de um sistema terrestre como impactos ambientais (Ozigis *et al.*, 2025), prever comportamento térmico do ar em cidades (Kubinić; Moffett; Blount, 2025) ou avaliar a emissão de carbono de vulcões (Romaniello; Ganci, 2025).

Não restrita à área de sensoriamento remoto, áreas interdisciplinares como a medicina se beneficiam diretamente da análise minuciosa de dados como em diagnóstico e tratamento de câncer (Wu, 2024) com a análise de dados em grande escala, ou em menor escala para geração de modelos preditivos para auxiliar no tratamento de cada caso, por exemplo com relação à lesões por pressão na pele (Voziniak *et al.*, 2023). A matemática pode vir não só em pontos específicos como uma equação, mas em conjunto com diversos conceitos, ampliando a perspectiva pela qual observam-se os problemas, como por exemplo com o uso de equações diferenciais junto de testes de hipótese, dentro de redes neurais, etc. A demanda por matemática como base para novas descobertas, visto que as relações observadas nas situações-problema de pesquisa envolvem muitas grandezas ou interações complexas a ponto de leis determinísticas não serem trivialmente notáveis. Além da pesquisa científica, os setores de serviços e indústria se beneficiam de pessoas que possuem formação e habilidades matemáticas. Cada vez mais oportunidades de trabalho requerem conhecimento em manipulação de dados como *SQL*, e manipulação de computação em nuvem como *Amazon Web Service*, *Google Cloud*, entre outras ferramentas que gerenciam grandes volumes de dados. Entre os trabalhos de cientista de dados, analistas de dados, estatísticos, *business intelligence*, e outras, a tendência é que as empresas utilizem mão-de-obra qualificada para identificar, selecionar e limpar os dados obtidos por diversos meios para prever comportamentos de máquinas, mercados, cadeias logísticas, entre quaisquer outras aplicações relevantes para as operações das empresas.

5.3 DISCIPLINAS DO ENSINO SUPERIOR NO ENSINO MÉDIO

Não somente a área das exatas que os alunos devem aprender. É inequívoco que as disciplinas de Língua Portuguesa e Matemática norteiam o ensino das escolas básicas e fundamentais. De fato, ambas as disciplinas permeiam e costuram as demais áreas que necessitam da leitura e raciocínio lógico e abstrato. O problema está na saída dos alunos, ou seja, o quanto preparado estará para seguir com uma profissão ou somente como um cidadão. Abaixo algumas das disciplinas do ensino superior que poderiam ser melhor incorporadas no ensino médio.

5.3.1 Cálculo Diferencial e Integral

Alguns trabalhos já preveem o ensino do Cálculo Diferencial (CDI) e Integral no Ensino Médio (Dias, 2019), inclusive com o auxílio de ferramentas de visualização. Diversos autores questionam a retirada do CDI dos currículos (Spina, 2002). Como já destacado, com o avanço de tecnologias computacionais, a Matemática se torna cada vez mais indissociável da formação para o mercado de trabalho, e mais complexos os problemas a serem resolvidos se tornam. Isso torna ponderável o ensino de tópicos matemáticos mais avançados para os estudantes. O CDI em específico, mesmo que aprendido de maneira introdutória, pode ajudar em problemas de otimização. Funções de várias variáveis com seus conceitos de curva de nível, podem ser exploradas com dados e a utilização de regressão linear múltipla, pois além do aluno ter o contato visual, o professor pode trabalhar com os dados aprimorando os conceitos que envolvem o Cálculo, tais como limite e derivadas.

5.3.2 Probabilidade e Estatística

Apesar de a Estatística já estar prevista como parte integrante do componente curricular da Matemática no Ensino Básico (Brasil, 2018), os tópicos abordados como a Estatística Descritiva e Exploratória dos Dados já são insuficientes, necessitando de análises mais profundas das quais necessitam da Inferência Estatística, tais como a regressão, pelo menos nos cursos técnicos e profissionalizantes. Uma alternativa seria através do uso da regressão linear simples ou múltipla, explicitando o conceito da inferência estatística dentro das técnicas de ajuste de modelos.

5.3.3 Álgebra Linear e Geometria Analítica

Machine Learning, Deep Learning, Redes Neurais, Sistemas Autônomos, Inteligência Artificial, IoT, entre outras técnicas que atualmente estão sendo incorporadas no dia a dia, tem um nome: Álgebra Linear. Todas essas técnicas tem alta dependência desse conceito. No en-

tanto, a dúvida mais pertinente e corriqueira dos alunos do ensino médio é, sem dúvidas, o uso, manipulação e interpretação das matrizes, o significado e interpretação do determinante, das inversas e demais operações. A regressão linear, no seu formato matricial pode ser um forte aliado para o entendimento das operações matriciais, principalmente com relação ao ajuste de modelos regressivos múltiplos.

5.3.4 Pensamento Lógico Computacional e Programação

Escolas do ensino fundamental já estão ensinando conceitos de programação para crianças abaixo dos 12 anos, no ensino fundamental. São aprendidas conceitos e lógica de programação com o uso do Python, com aplicações do robot LEGO (vide em: <https://spike.legoeducation.com/essential/lobby/>). Esses novos rumos da educação prometem uma revolução significativa no tratamento das informações pelos professores e alunos. Programação e Python e o software R são grandes aliados para introduzir conceitos de programação. Mais além, ao usar as linhas de código para ler, organizar e interpretar dados, a regressão linear simples é um bom exemplar para os alunos.

5.3.5 Educação Financeira e Princípios de Contabilidade Geral

O Projeto de Lei nº 5950, de 2023, propôs alterar a Lei nº 9.394, de 20 de dezembro de 1996, que estabelece as diretrizes e bases da educação nacional, para incluir educação e administração financeira nos currículos da educação básica. Acredita-se que essa lei é importante para a inclusão da educação financeira nas escolas podendo contribuir para o aumento do empreendedorismo, a diminuição das dívidas e inadimplência dos cidadãos e aumentar a capacidade econômica local e, por conseguinte, atrair investimentos nacionais e internacionais. Dessa forma, o aluno se condicionará a sua relação com o dinheiro de forma responsável, com diminuição do stress financeiro para si e a sua família. Na sociedade, essa troca de relação com o dinheiro começando no pessoal e expandindo para o grupo de moradores e cidadãos, poderá melhorar a qualidade de vida de todos. Não o bastante, muitos alunos ao terminar o ensino médio não sabem o básico, como preencher uma declaração de imposto de renda que vincula o seu próprio CPF, ou seja, a sua identidade. Atualmente, com a educação financeira muitas escolas incentivam o empreendedorismo como parte da formação. Previsibilidade, esse é o grande "segredo" de um bom investidor. No mundo dos negócios, saber investir é o principal passo para o bom empreendedorismo. A análise de séries temporais, nada mais são do que ajustes de modelos que, de uma forma ou de outra, seguem o padrão do ajuste de regressão linear. De forma mais simples, ajustar um modelo de custo vs despesas por regressão linear auxilia o aluno na compreensão dos modelos básicos dos princípios de finanças e contabilidade.

6 RESULTADOS E DISCUSSÃO

Embora o assunto seja do ramo da Estatística, convém que os estudantes aprendam sobre a estatística básica conforme o regulamento da OCDE, caso desejam melhorar o índice do PISA. O assunto da regressão múltipla engloba muitos assuntos das ferramentas da matemática superior, tais como: funções de várias variáveis, álgebra linear (operação com matrizes), curvas de nível, vetores gradientes e etc. Nesse contexto, usamos a teoria da criação da regressão múltipla para que o professor compreenda a dinâmica do ajuste dos dados, incorporando conhecimentos da matemática superior no ensino médio.

6.1 Detalhamento matemático nas regressões estatísticas

Para que o professor possa utilizar a técnica com seus alunos, é importante que ele possua um entendimento integral da regressão múltipla e sua constituição matemática. Assim, é importante que façamos a caracterização detalhada da Matemática envolvida na técnica a seguir.

6.1.1 Regressão Linear vetorial

A regressão linear é uma forma de modelagem que busca modelar a relação entre duas variáveis, supondo que exista uma relação de correlação entre as duas. Na natureza, podemos observar diversos fenômenos cujas variáveis interagem linearmente, ou seja, proporcionalmente. Por exemplo, velocidade e a distância percorrida por um corpo, volume de um corpo e o empuxo de um líquido em que esse corpo se submerge, corrente elétrica em uma resistência e o calor dissipado por ela, entre outros tantos outros exemplos.

Entretanto, ao se fazer medições em experimentos, percebe-se uma diferença entre o que se espera teoricamente e os dados medidos. Isso se dá por diversos fatores aleatórios, como o desconhecimento de todas as variáveis que interferem no sistema, imprecisões nas medições e até a própria incerteza obrigatória ao se interagir com um sistema para medir suas grandezas. Dessa forma, é de se esperar que um conjunto de dados observáveis podem estar relacionados entre si por diversos modelos, sendo o mais simples deles, o linear.

Assim, para se determinar o modelo que mais se aproxime do comportamento real das variáveis a partir apenas dos dados observados, busca-se um modelo cuja reta se aproxime ao máximo dos pontos observados. Seja cada ponto um par ordenado (x, y) onde x é a variável independente e y a variável dependente de x e seja a forma de um modelo linear qualquer como $y = \beta_0 + \beta_1 x$ onde a resposta y é resultado de uma entrada x multiplicada por uma inclinação β_1 , somada então a um intercepto β_0 . Para muitas covariáveis a forma vetorial é a mais indicada e será descrita por

$$\mathbf{Y} = \mathbf{XB} + \epsilon$$

Supondo então que exista um modelo linear real que descreva a relação entre x e y , cada medição possui então um erro associado de natureza aleatória ϵ , sendo ele a distância entre o y_i observado e o real y . Ou seja, o modelo que prevê os valores de $\hat{y}$ deve contemplar esse erro da forma $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ para

$$\epsilon \sim N(0, \sigma_\epsilon^2)$$

A qual segue uma distribuição normal de probabilidade. Pois, caso contrário, o modelo deverá ser avaliado por modelos de regressão generalizados. Nesse trabalho de conclusão de curso apresentaremos a matemática de base para o modo normal.

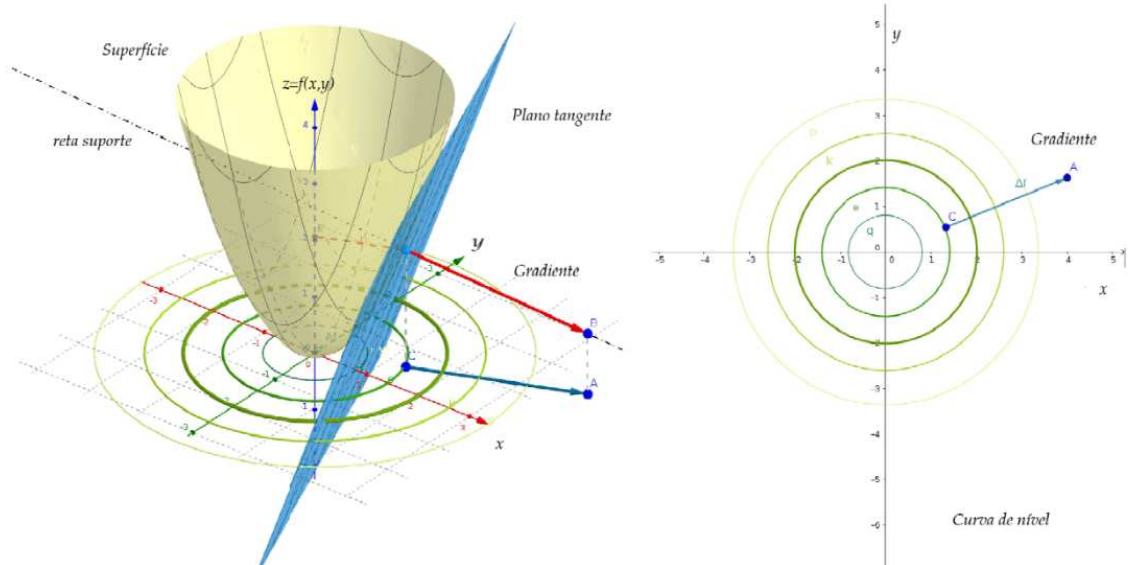
6.1.2 Método dos Mínimos Quadrados Ordinários (na forma de somatório)

Tendo em vista que toda observação possui um erro associado a ela em relação à predição, procura-se diminuir ao máximo os erros. Para isso, seja a soma dos quadrados dos erros uma função f da forma

$$f = \sum_{i=1}^n [y_i - (\beta_0 + \beta_1 x_i)]^2 \quad (1)$$

Para que a função seja minimizada, então, deve-se encontrar as derivadas parciais em relação a β_0 e β_1 , igualando-as a zero para localizar o ponto crítico de f . Vale verificar a Figura 1 do mecanismo para encontrar o ponto de mínimo da função quadrática para a função 1

Figura 1 – Técnica para encontrar o ponto de mínimo para os coeficientes da regressora



Fonte: Autoria própria.

Para isso, torna-se necessário que o gradiente da função seja nulo.

$$\vec{\nabla} f(\beta_0, \beta_1) = 0$$

Desde que o determinante da matriz Hessiana seja verificado.

$$\begin{aligned} \frac{\partial}{\partial \beta_0} \sum_{i=1}^n [y_i - (\beta_0 + \beta_1 x_i)]^2 &= 0 \\ \frac{\partial}{\partial \beta_1} \sum_{i=1}^n [y_i - (\beta_0 + \beta_1 x_i)]^2 &= 0 \end{aligned}$$

Sejam $\hat{\beta}_0$ e $\hat{\beta}_1$ os valores que minimizam a função, temos, assim, o sistema de equações normais

$$\begin{aligned} -2 \sum_{i=1}^n [y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)] &= 0 \\ -2 \sum_{i=1}^n [y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)] x_i &= 0 \end{aligned}$$

Realizando a devida manipulação algébrica para isolar $\hat{\beta}_0$ na primeira equação, tem-se

$$\hat{\beta}_0 = \frac{1}{n} \sum_{i=1}^n y_i - \hat{\beta}_1 \frac{1}{n} \sum_{i=1}^n x_i$$

Substituindo na segunda equação, tem-se

$$-2 \sum_{i=1}^n \left[y_i - \left(\frac{1}{n} \sum_{i=1}^n y_i - \hat{\beta}_1 \frac{1}{n} \sum_{i=1}^n x_i \right) + \hat{\beta}_1 x_i \right] x_i = 0$$

Realizando também a manipulação algébrica para isolar $\hat{\beta}_0$ temos as soluções para a minimização como

$$\begin{aligned} \hat{\beta}_0 &= \frac{1}{n} \sum_{i=1}^n y_i - \hat{\beta}_1 \frac{1}{n} \sum_{i=1}^n x_i \\ \hat{\beta}_1 &= \frac{\sum_{i=1}^n x_i y_i - \frac{1}{n} \sum_{i=1}^n y_i \sum_{i=1}^n x_i}{\sum_{i=1}^n x_i^2 - \frac{1}{n} (\sum_{i=1}^n x_i)^2} \\ \hat{y}_i &= \hat{\beta}_0 + \hat{\beta}_1 x_i \end{aligned}$$

onde os resíduos e_i da regressora são dados por $y_i - \hat{y}_i$

O detalhamento matemático aqui está no conceito de que o denominador deverá ser diferente de zero, implicando em

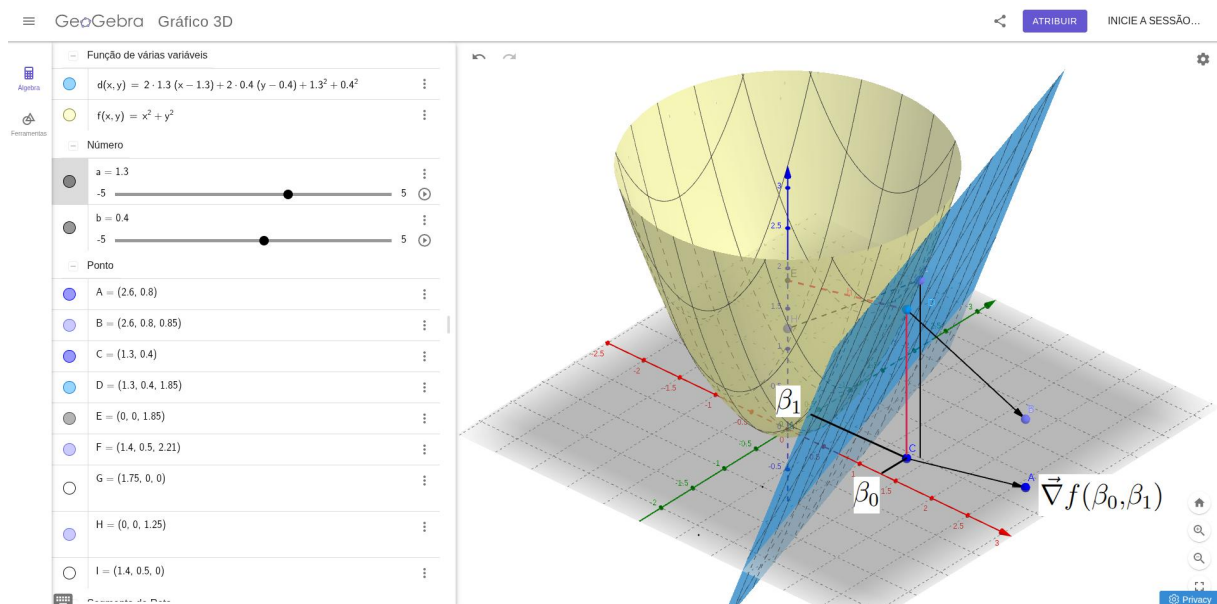
$$\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 > 0 \quad \text{ou} \quad \bar{x}_{RMS} > \bar{x}$$

Esse resultado implica que o valor eficaz da média é superior à média aritmética.

6.1.2.1 Utilização no ensino médio: o efeito da otimização.

O professor pode mostrar a regressão linear simples aplicando os parâmetros otimizados ajustando uma reta a um conjunto de pontos gerados aleatoriamente ou obtidos por uma base de dados ou até por experimentação. Depois os alunos podem propor quaisquer outros valores **não otimizados** β_0 e β_1 , para o caso $\vec{\nabla} f \neq 0$ incluindo as retas na curva observada. Como observado na figura 2, utilizando-se o software Geogebra, pode-se observar o ajuste da curva com base em valores arbitrários para os coeficientes da reta, cujo comportamento pode ser observado no gráfico 3d à direita na figura. O resultado esperado é a falta de um bom ajuste.

Figura 2 – Escolhendo um ponto com $\vec{\nabla} f \neq 0$, caso não ótimo.



Fonte: Autoria própria.

6.1.2.2 Método dos Mínimos Quadrados Ordinários (Forma Vetorial)

Para compactar funções de muitas variáveis, torna-se prático reduzir a notação na forma vetorial. Por exemplo: quando tratamos de duas variáveis, a forma padrão (forma algébrica linear) pode ser adequada. No entanto, para grandes quantidades de variáveis e incluindo a interação entre as variáveis, a notação matricial é a melhor escolha.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \beta_p x_p = \beta_0 + \sum_{i=1}^p \beta_i x_i$$

Para a seguinte forma

$$\mathbf{Y} = \mathbf{X}\beta$$

A partir deste momento, quando necessário, vamos utilizar esta notação.

6.1.3 Encontrando o Ponto de Mínimo da Função $y = \beta_0 + \beta_1 x + \epsilon$

A matemática que envolve o estudo dos pontos de máximo/mínimo para várias variáveis relaciona-se ao gradiente nulo da função, como ponto(s) candidato(s):

$$\vec{\nabla} f(\hat{\beta}_0, \hat{\beta}_1) = \vec{0}$$

cuja existência já foi verificada. Seja, em seguida, a matriz Hessiana formada pelas derivadas de segunda ordem de f dada por

$$H = \begin{bmatrix} f_{\beta_0\beta_0} & f_{\beta_0\beta_1} \\ f_{\beta_1\beta_0} & f_{\beta_1\beta_1} \end{bmatrix}$$

Pode-se utilizar seu determinante $\det(H)$ para determinar a concavidade da curva f . Pelo Teorema da Hessiana temos

$$\det(H) \begin{cases} > 0 \text{ e } f_{\beta_0\beta_0}(\hat{\beta}_0, \hat{\beta}_1) > 0, \text{ então } (\hat{\beta}_0, \hat{\beta}_1) \text{ é um ponto de mínimo local.} \\ > 0 \text{ e } f_{\beta_0\beta_0}(\hat{\beta}_0, \hat{\beta}_1) < 0, \text{ então } (\hat{\beta}_0, \hat{\beta}_1) \text{ é um ponto de máximo local.} \\ < 0, \text{ então } (\hat{\beta}_0, \hat{\beta}_1) \text{ é um ponto de sela.} \\ = 0, \text{ nada se conclui sobre a concavidade da curva.} \end{cases}$$

Assim, temos as derivadas de 2ª ordem dispostas na matriz hessiana como

$$H = \begin{bmatrix} 2n & 2 \sum_{i=1}^n x_i \\ 2 \sum_{i=1}^n x_i & 2 \sum_{i=1}^n x_i^2 \end{bmatrix}$$

Assim, o determinante de H é dado por

$$\begin{aligned} \det(H) &= 4n \sum_{i=1}^n x_i^2 - 4 \left(\sum_{i=1}^n x_i \right)^2 \\ \det(H) &= 4n^2 \left(\frac{\sum_{i=1}^n x_i^2}{n} - \left(\frac{\sum_{i=1}^n x_i}{n} \right)^2 \right) \end{aligned}$$

Pela desigualdade de Cauchy-Schwarz, temos que o termo

$$\frac{\sum_{i=1}^n x_i^2}{n} - \left(\frac{\sum_{i=1}^n x_i}{n} \right)^2 > 0$$

Ou seja, o ponto $(\hat{\beta}_0, \hat{\beta}_1)$ é um ponto de mínimo da função f . De novo, podemos observar a importância do resultado $\bar{x}_{RMS} > \bar{x}$. Então se a raiz do $\det(H)$

$$\sqrt{\det(H)} = a + bi$$

sendo b a parte complexa, para

$$\sqrt{\det(H)} = 2n(\bar{x}_{RMS} - \bar{x}) + 0i$$

é determinada apenas pela parte real. Se ocorrer do $\det(H) = 0$, não podemos encontrar o coeficiente b_1 , pois será inexistente. Isso ocorre quando na análise exploratória encontramos $\bar{x}_{RMS} = \bar{x}$. Ou seja, no âmbito do Teorema da Hessiana, nada podemos concluir sobre a concavidade da curva e, por sua vez, a regressora inexistente!

Para o caso com muitas covariáveis, temos

$$\epsilon = \mathbf{Y} - \mathbf{XB}$$

onde $\mathbf{Y}$ é a matriz de resultados observáveis $y_1, y_2, \dots, y_n$; β é o vetor dos coeficientes relevantes $\beta_0, \beta_1, \beta_2, \dots, \beta_n$; $\mathbf{X}$ é a matriz de variáveis independentes observáveis, cujo número de linhas é o tamanho da amostra, de 1 a n , e o número de colunas é o número de coeficientes referentes às variáveis envolvidas na modelagem mais um termo independente, ou seja o tamanho do vetor β . ϵ segue uma distribuição normal multivariada (NM_n) com média 0 e desvio-padrão $\sigma^2 I$.

6.1.3.1 Utilização no ensino médio: o determinante da matriz

Para que serve o determinante na prática? Essa pergunta é muito comum dos alunos do ensino médio. Aqui podemos ver um caso. Se o determinante da matriz Hessiana resultar em um valor nulo, então não podemos encontrar uma regressora. Isto ocorre se:

$$\begin{aligned}
\det(H) &= 0 \\
\det(H) &= 4n \sum_{i=1}^n x_i^2 - 4 \left(\sum_{i=1}^n x_i \right)^2 = 0 \\
n \sum_{i=1}^n x_i^2 &= \left(\sum_{i=1}^n x_i \right)^2, \quad \left(* \frac{1}{n^2} \right) \\
\frac{\sum_{i=1}^n x_i^2}{n} &= \left(\frac{\sum_{i=1}^n x_i}{n} \right)^2 \\
\sqrt{\frac{\sum_{i=1}^n x_i^2}{n}} &= \frac{\sum_{i=1}^n x_i}{n} \\
\bar{x}_{rms} &= \bar{x}
\end{aligned}$$

Ou seja, se a média quadrática for igual à média aritmética, o determinante da matriz Hessiana resultará em zero e, por sua vez, não encontramos um par de valores que determine uma reta efetiva. Isto é, se $\bar{x}_{rms} = \bar{x}$, não há reta! O professor pode comentar sobre a necessidade da matriz Hessiana existir, explicando que é um característico matemático que define a possibilidade ou não em ajustar uma reta.

6.1.3.2 Soma dos Quadrados

Ao se observar conjuntos de dados, existem diversas maneiras de se explorá-los matematicamente para se obter informações. Podem, ser observadas médias, medidas de dispersão, distância dos dados até pontos de referência, etc. Dos métodos de análise de dados quantitativos, o passo mais comum a se tomar para começar qualquer análise seria a determinação da média aritmética dos valores.

Seja y_i o conjunto de quaisquer valores observados com $i \in \mathbb{N}$, a média aritmética entre os valores de y para n observações é calculada por $\frac{\sum_{i=1}^n y_i}{n}$.

A regressão dos dados pode ser feita para modelar sua tendência e prever valores futuros. Entretanto, caso a média $\bar{y}$ seja suficiente para modelar os dados, não haveria necessidade de se modelar uma regressora. Para que a regressão seja válida, é necessário que haja diferença significativa, então, entre a regressora $\hat{y}$ e a média $\bar{y}$.

Já determinamos que a regressão linear prediz valores de y com base na variável x , Seja $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$. Assim, possuímos três valores distintos no eixo das ordenadas: y que é o valor observado para um determinado x , $\hat{y}$ o valor predito pela regressão de y , e $\bar{y}$ o valor médio de y .

A diferença entre o y observado e o $\hat{y}$ predito pela regressora, denotamos por $d_e = y - \hat{y}$ a diferença entre o y observado e o da regressora. Essa diferença pode ser observada por conta de fatores aleatórios que interferem tanto na observação quanto na medição de qualquer

fenômeno. No mundo real, é pouco provável que encontremos o valor de d_e como 0, visto que a maioria dos processos naturais está sujeita a flutuações e incertezas sobre as quais não possuímos poder ou tempo hábil a modelar.

A outra diferença notável é observada entre os valores de $\hat{y}$ e $\bar{y}$, denotada aqui por $d_{reg} = \hat{y} - \bar{y}$, sendo a diferença entre a regressora e a média dos valores observados de y . O que se busca na análise de regressão, é se existe uma diferença relevante, ou seja, estatisticamente significativa, entre $\hat{y}$ e $\bar{y}$, de forma que a regressora modele melhor os dados com maior precisão do que sua média.

Além disso, nota-se que podemos somar as duas diferenças

$$\begin{aligned}d_e + d_{reg} &= (y - \hat{y}) + (\hat{y} - \bar{y}) \\d_e + d_{reg} &= y - \hat{y} + \hat{y} - \bar{y} \\d_e + d_{reg} &= y - \bar{y}\end{aligned}$$

resultando na nova diferença que chamaremos de desvio total $d_t = y - \bar{y}$.

A cada uma das diferenças, podemos também observar seus respectivos quadrados. Denominaremos então SQT como a soma dos quadrados das diferenças totais, ou seja $SQT = \sum_{i=1}^n (y_i - \bar{y})^2$ o que pode ser expandido adicionando-se $(\hat{y} - \bar{y})$ ao interior do termo quadrático, visto que o termo adicionado é nulo. Assim, temos

$$\begin{aligned}SQT &= \sum_{i=1}^n (y_i - \bar{y} + \hat{y} - \bar{y})^2 \\SQT &= \sum_{i=1}^n [(y_i - \hat{y})^2 + (\hat{y} - \bar{y})^2 + 2(y_i - \hat{y})(\hat{y} - \bar{y})] \\SQT &= \sum_{i=1}^n (y_i - \hat{y})^2 + \sum_{i=1}^n (\hat{y} - \bar{y})^2 - 2 \sum_{i=1}^n (y_i - \hat{y})(\hat{y} - \bar{y})\end{aligned}$$

Agora vamos analisar o termo $-2 \sum_{i=1}^n (y_i - \hat{y})(\hat{y} - \bar{y})$ e suas diferenças internas. Primeiramente observa-se que

$$\sum_{i=1}^n (y_i - \hat{y})(\hat{y} - \bar{y}) = \sum_{i=1}^n e_i(\hat{y} - \bar{y})$$

Voltando ao sistema de equação normais, a partir da primeira equação, vemos que

$$\begin{aligned}\sum_{i=1}^n e_i(\hat{y} - \bar{y}) &= \sum_{i=1}^n y_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum_{i=1}^n x_i = 0 \\ \sum_{i=1}^n y_i - \sum_{i=1}^n (\hat{\beta}_0 + \hat{\beta}_1 x_i) &= 0\end{aligned}$$

Como $\hat{\beta}_0 + \hat{\beta}_1 x_i = \hat{y}_i$ temos então

$$\begin{aligned}\sum_{i=1}^n y_i &= \sum_{i=1}^n \hat{y}_i \\ \sum_{i=1}^n (y_i - \hat{y}_i) &= 0 \\ \sum_{i=1}^n d_e &= 0\end{aligned}$$

Portanto, o termo cruzado é nulo,

$$\sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{y}) = 0$$

ou de forma vetorial

$$(\mathbf{y} - \hat{\mathbf{y}})^T (\hat{\mathbf{y}} - \bar{\mathbf{y}}) = 0$$

O produto escalar acima será zero, se e somente se, $(\mathbf{y} - \hat{\mathbf{y}})$ e $(\hat{\mathbf{y}} - \bar{\mathbf{y}})$ forem vetores ortogonais, tal que o ângulo interno entre eles é de exatos 90° , caso contrário o produto será maior do que zero se o ângulo for menor do que 90° e menor do que zero se o ângulo for maior do que 90°

6.1.3.3 Independência Linear

Em matemática a independência linear existirá entre dois vetores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ não nulos se forem ortogonais.

- *Prova:* Seja $c_1 v_1 + c_2 v_2 + \dots + c_k v_k = 0$.
Vamos multiplicar ambos os lados por v_1^T ;

$$v_1^T (c_1 v_1 + c_2 v_2 + \dots + c_k v_k) = 0$$

Observe que $v_1^T v_2 = 0$ por serem ortogonais, então o resultado será

$$c v_1^T v_1 = 0$$

O único resultado para a nulidade é que $c_1 = 0$. Portanto, o mesmo resultado é válido para quaisquer c_i , levando aos casos em que $c_i = 0, \forall i$.

6.1.3.4 Teorema de Pitágoras

O teorema de Pitágoras informa-nos que a soma dos quadrados dos catetos é igual ao quadrado da hipotenusa:

$$SQT = \sum_{i=1}^n (y_i - \hat{y})^2 + \sum_{i=1}^n (\hat{y} - \bar{y})^2 \equiv \|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

para os vetores desvios do erro e da regressora

$$\mathbf{u} = (\mathbf{y} - \hat{\mathbf{y}}) \quad \text{e} \quad \mathbf{v} = (\hat{\mathbf{y}} - \bar{\mathbf{y}})$$

De forma mais compacta, temos;

$$SQT = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 0$$

implicando na independência entre os vetores $(\mathbf{y} - \hat{\mathbf{y}})$ e $(\hat{\mathbf{y}} - \bar{\mathbf{y}})$.

$$SQT = SQ_e + SQ_{reg}$$

Isso leva-nos a um resultado importante matematicamente sob a ortogonalidade entre SQ_e e SQ_{reg} .

Teorema de Pitágoras: Sejam dois vetores $\mathbf{u}$ e $\mathbf{v}$, então

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

Os vetores $\|\mathbf{u}\|$ e $\|\mathbf{v}\|$ deverão ser ortogonais entre si, e essa é uma propriedade importante para a existência da regressora, $\mathbf{u}^T \mathbf{v} = 0$.

Equanto a quantidade SQT é a soma entre SQ_e e a regressora SQ_{reg} , obtemos o seguinte resultado:

$$\begin{aligned} SQT &= SQ_e + SQ_{reg} \\ 1 &= \frac{SQ_e}{SQT} + \frac{SQ_{reg}}{SQT} \end{aligned}$$

O interesse é a proporção explicada da variação usando o coeficiente de ajuste R^2 :

$$R^2 = \frac{SQ_{reg}}{SQT}$$

o que significa o quanto a variação é explicada pela regressora.

6.1.3.5 Utilização no ensino médio: o impacto do teorema de Pitágoras dentro do ajuste

Embora o teorema de Pitágoras é simples, do ponto de vista geométrico, não podemos dizer o mesmo no aspecto algébrico. No exemplo da regressão múltipla, está vinculado à perpendicularidade entre o desvio dos erros e o desvio da regressora. Em tese, queremos ajustar mais pontos com melhor aproximação dos resultados. Isso se dá quando o quadrado SQ_{reg} é superior ao SQ_e . O professor pode mostrar o efeito do teorema de Pitágoras usando o resultado do coeficiente R^2 . Pois, é a parcela explicativa da área do cateto calculado pelo quadrado da regressora em razão da área total da hipotenusa calculada por SQT .

6.1.4 Desenvolvimento MMQO

O modelo consiste em escrever os termos da regressora em uma forma matricial a fim de encontrarmos os coeficientes β apenas com a informação das covariáveis e a variável resposta Y. Vamos verificar que o valor esperado para os termos da matriz de coeficientes ajustados $\hat{\beta}$

são exatamente os termos dos coeficientes da regressora otimizada. Sejam $y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$,

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & . & . & . & x_{1p} \\ 1 & x_{21} & x_{22} & . & . & . & x_{2p} \\ 1 & x_{31} & x_{32} & . & . & . & x_{3p} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 1 & x_{n1} & x_{n2} & . & . & . & x_{np} \end{bmatrix} \quad \text{e} \quad b = \begin{bmatrix} a \\ b_1 \\ b_2 \\ \vdots \\ b_p \end{bmatrix}$$

Temos que $y - Xb$ é igual a $\begin{bmatrix} y_1 - (a + b_1x_{11} + b_2x_{12} + \dots + b_px_{1p}) \\ y_2 - (a + b_1x_{21} + b_2x_{22} + \dots + b_px_{2p}) \\ \vdots \\ y_n - (a + b_1x_{n1} + b_2x_{n2} + \dots + b_px_{np}) \end{bmatrix}$.

Como o produto de um vetor V pelo seu transposto V' é igual à soma dos quadrados dos seus elementos, temos que

$$\begin{aligned}\sum_{i=1}^n [y_i - (a + b_1x_{i1} + b_2x_{i2} + \dots + b_px_{ip})]^2 &= (y - Xb)'(y - Xb) \\ &= (y' - X'b')(y - Xb) = y'y - 2y'Xb + b'X'Xb\end{aligned}$$

Nota-se que cada termo da expressão $y'y - 2y'Xb + b'X'Xb$ é um escalar. Então derivando a expressão com relação a b , temos

$$\begin{aligned}\frac{\partial y'y}{\partial b} - 2\frac{\partial y'Xb}{\partial b} + \frac{\partial b'X'Xb}{\partial b} &= 0 - 2(y'X)' + 2X'Xb \\ &= -2(X'y) + 2X'Xb\end{aligned}$$

Da mesma forma que fizemos com o caso unidimensional, chamamos de $\hat{b}$ o vetor que torna nula a derivada acima. Então temos

$$\begin{aligned}-2X'y + 2X'X\hat{b} &= 0 \\ X'X\hat{b} &= X'y \\ (X'X)^{-1}X'X\hat{b} &= (X'X)^{-1}X'y \\ \hat{b} &= (X'X)^{-1}X'y\end{aligned}$$

Ou seja, dados $y = (y_1, y_2, \dots, y_n)$ um vetor de dados-resposta a uma matriz X cujas linhas são os valores fixos das variáveis regressoras, o estimador de mínimos quadrados para um vetor de coeficientes β é dado por $\hat{\beta} = (X'X)^{-1}X'y$ desde que a inversa de $X'X$ exista. Obviamente o produto $(X'X) = ||X||^2$ refere-se ao módulo em R^n como sendo uma raiz quadrada positiva de $X'X$.

6.1.4.1 Utilização no ensino médio: a importância da inversa de uma matriz

Dentro das regressões múltiplas, a inversa de uma matriz têm o papel fundamental de dar sequência nos cálculos matemáticos. Mais importante ainda é fornecer o resultado dos coeficientes da regressora. O professor pode usar o resultado

$$\hat{b} = (X'X)^{-1}X'y$$

mostrando claramente que a quantidade $(X'X)^{-1}$ deverá existir para que sejam encontrados os coeficientes.

6.1.5 Projeções e Mínimos Quadrados

Uma outra forma de avaliarmos é através das projeções. Podemos analisar o que ocorre matematicamente usando critérios de projeção.

6.1.6 A Esperança Matemática

Assim, a esperança de $\hat{\beta}$ é dada por

$$E[\hat{\beta}] = E[(X'X)^{-1}X'y]$$

Como X é uma matriz de valores já observados, X é constante, então

$$\begin{aligned} E[\hat{\beta}] &= (X'X)^{-1}X'E[y] \\ &= (X'X)^{-1}X'X\beta \\ &= I\beta \\ &= \beta \end{aligned}$$

Da mesma forma, a variância de $\hat{\beta}$ é dada por

$$\begin{aligned} Var[\hat{\beta}] &= Var[(X'X)^{-1}X'y] \\ &= (X'X)^{-1}X'Var[y](X'X)^{-1}X' \\ &= \sigma^2(X'X)^{-1}X'[(X'X)^{-1}X']'. \text{ visto que } Var(x) = \sigma^2 \\ &= \sigma^2(X'X)^{-1}X'X(X'X)^{-1} \\ &= \sigma^2(X'X)^{-1} \end{aligned}$$

Por questões de notação, adotamos como padrão $(X'X)^{-1}$ como C .

Então, para $\hat{\beta}_j$ um elemento de $\hat{\beta}$, temos

$$\hat{\beta}_j \sim N(\beta_j; \sigma^2 C_{j+1,j+1})$$

sendo $j = 0, 1, \dots, p$ e $C_{j+1,j+1}$ o $(j+1)$ -ésimo elemento da diagonal principal de C , visto que $\hat{\beta}_0$ é o termo independente. A covariância entre $\hat{\beta}_k$ e $\hat{\beta}_j$, dois elementos de $\hat{\beta}$, então, é dada por $\sigma^2 C_{k+1,j+1}$.

Então, seja x'_i a i -ésima linha de X , ou seja, um vetor $[1, x_{i1}, x_{i2}, \dots, x_{ip}]$, o estimador de quadrados mínimos da esperança de Y , correspondente aos valores de x'_i , é

$$\hat{y}_i = x'_i \hat{\beta}$$

Com distribuição normal univariada, visto que possui apenas uma dimensão. A esperança e variância do estimador então são dadas por

$$\begin{aligned} E[x'_i \hat{\beta}] &= x'_i E[\hat{\beta}] = x'_i \beta \\ \text{Var}[x'_i \hat{\beta}] &= x'_i \text{Var}[\hat{\beta}] x_i \\ &= x'_i \sigma^2 (X'X)^{-1} x_i \\ &= \sigma^2 x'_i (X'X)^{-1} x_i \end{aligned}$$

Ou seja:

$$\hat{y}_i = x'_i \hat{\beta} \sim N(x'_i \beta; \sigma^2 x'_i (X'X)^{-1} x_i)$$

Dessa forma, agora que temos a esperança e a média para os estimadores, podemos definir os intervalos de confiança para cada esperança de x_i com nível de confiança $(1-\alpha)100\%$ como:

$$x'_i \hat{\beta} \pm t_{\frac{\alpha}{2}, n-p-1} \sqrt{\hat{\sigma}^2 x'_i (X'X)^{-1} x_i}$$

Também, utilizando a notação de x'_i para cada linha de X , temos a própria matriz X escrita como um vetor

$$X = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{bmatrix}$$

O que torna o vetor de resíduos $\epsilon = y - X\hat{\beta}$ igual a

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} - \begin{bmatrix} x'_1 \hat{\beta} \\ x'_2 \hat{\beta} \\ \vdots \\ x'_n \hat{\beta} \end{bmatrix}$$

Para estimar σ^2 , de forma similar à regressão linear, pode-se dividir a soma quadrática dos resíduos por $(n-p-1)$, ou seja

$$\hat{\sigma}^2 = \frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{n - p - 1}$$

6.2 Intervalos de confiança

Para cada coeficiente $\hat{\beta}_j$, com $j = 0, 1, \dots, p$, existirá um intervalo de confiança em que a estimativa reflete satisfatoriamente a relação entre x_j e β_j utilizando a distribuição t -Student e Qui-quadrado

$$\frac{\hat{\beta}_j - \beta_j}{(\hat{\sigma}^2 C_{j+1,j+1})^{\frac{1}{2}}} \sim t_{n-p-1}$$

$$\frac{(n - p - 1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{(n-p-1)}$$

Assim, o intervalo de confiança $(1 - \alpha)100\%$ para cada parâmetro β_j é dado por

$$\hat{\beta}_j \pm t_{(\frac{\alpha}{2}, n-p-1)} \sqrt{\hat{\sigma}^2 C_{j+1,j+1}}, \text{ com } j = 1, 2, \dots, p$$

enquanto o intervalo para σ^2 é dado por

$$\left[\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\chi^2_{(\frac{\alpha}{2}, n-p)}}, \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\chi^2_{(1-\frac{\alpha}{2}, n-p)}} \right]$$

Para se julgar a validade de um modelo sobre o outro, é necessária a avaliação das opções de modelo com variáveis auxiliares, ou seja o Modelo de Regressão Linear Múltipla (MRLM), ou sem variáveis auxiliares, o Modelo de Simples. A forma mais comum de se comparar os modelos também é usada em outros cenários estatísticos por meio da tabela de análise de variância, ou ANOVA (ANalysis Of VAriance) e por um teste de hipótese com base em uma estatística com distribuição F de Fisher. Para isso, na hipótese nula supomos que os modelos simples e MRLM se sobrepõem, ou seja, que $\beta_1, \beta_2, \dots, \beta_p$ sejam todos iguais a zero. A hipótese alternativa é de que pelo menos um dos coeficientes seja não-nulo, indicando que pelo menos uma das variáveis regressoras contém significância para a modelagem do problema.

Então observamos as somas dos quadrados na forma matricial, semelhante à forma linear simples.

$$SQT = SQR_{eg} + SQE$$

$$\sum_{i=1}^n (y_i - \hat{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Então, usando a notação matricial, para cada soma de quadrados, temos

$$SQT = \sum_{i=1}^n (y_i - \bar{y})^2$$

$$= \sum_{i=1}^n (y_i - \bar{y})(y_i - \bar{y})$$

como $\sum_{i=1}^n y_i(x_i - \bar{x}) = \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})$

então $SQT = \sum_{i=1}^n y_i(y_i - \bar{y})$

$$= \sum_{i=1}^n y_i^2 - \bar{y} \sum_{i=1}^n y_i$$

$$= \sum_{i=1}^n y_i^2 - n\bar{y}^2$$

$$SQT = \mathbf{y}'\mathbf{y} - n\bar{y}^2$$

para $SQE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$

$$= (\mathbf{y} - \mathbf{X}\hat{\beta})'(\mathbf{y} - \mathbf{X}\hat{\beta})$$

$$= \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}\hat{\beta} - \hat{\beta}'\mathbf{X}'\mathbf{y} + \hat{\beta}'\mathbf{X}'\mathbf{X}\hat{\beta}$$

$$= \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}\hat{\beta}$$

$$SQE = \mathbf{y}'\mathbf{y} - \hat{\beta}'\mathbf{X}'\mathbf{y}$$

e por último, partindo da definição de SQT , $SQR_{eg} = SQT - SQE$

$$= \mathbf{y}'\mathbf{y} - n\bar{y}^2 - \mathbf{y}'\mathbf{y} + \hat{\beta}'\mathbf{X}'\mathbf{y}$$

$$SQR_{eg} = \hat{\beta}'\mathbf{X}'\mathbf{y} - n\bar{y}^2$$

Dessa forma, a tabela ANOVA para o modelo de regressão linear múltipla fica:

Tabela 1 – Análise de Variância (ANOVA)

Fonte de variação	GL (Graus de liberdade)	SQ (Soma de quadrados)	QM (quadrado médio)	F_0
Regressão	p	SQR_{eg}	MS_{reg}	$\frac{MS_{reg}}{MS_e}$
Erro	$n - p - 1$	SQE	MS_e	
Total	$n - 1$	SQT		

com

$$MS_{reg} = SQR_{eg}/p$$

e

$$MS_e = SQE/(n - p - 1)$$

Então, se F_0 for maior do que ${}_pF_{(n-p-1)(\alpha)}$, rejeita-se a hipótese H_0 .

Então, podemos analisar a contribuição que cada variável regressora possui no modelo, tornando-o mais ou menos complexo conforme a colocação e retirada de variáveis da regressora. Assim, se quisermos analisar a relevância de uma variável x_k , podemos formular as seguintes hipóteses:

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0$$

Essas hipóteses podem ser testadas com base na distribuição t de Student em que

$$T_{\beta_k} = \frac{\hat{\beta}_k}{\sqrt{\hat{\sigma}^2 C_{k+1,k+1}}}$$

Então, para um dado nível de significância α , a hipótese H_0 , ou seja, de que β_k é irrelevante para o modelo, é rejeitada se $|T_{\beta_k}| \geq t_{(\frac{\alpha}{2}, n-p-1)}$.

Também podemos fazer o teste para o quadrado de T . Sabendo que o quadrado de uma variável aleatória que segue uma distribuição t-student com n graus de liberdade segue uma distribuição F de Fisher com 1 e n graus de liberdade, podemos rejeitar H_0 se $T_{\beta_k}^2 \geq {}_1F_{(n-p-1)(\alpha)}$.

Podemos testar conjuntamente mais de um parâmetro ao mesmo tempo, utilizando $\hat{\beta}^*$, um subvetor de r componentes de $\hat{\beta}$. Por resultados de Estatística e Álgebra Linear, temos que

$$\frac{1}{\hat{\sigma}^2}(\hat{\beta}^* - \beta^*)' \Sigma^{*-1} (\hat{\beta}^* - \beta^*) \sim \chi_r^2$$

De forma análoga e por resultados de probabilidade

$$\frac{1}{\hat{\sigma}^2}(\hat{\beta}^* - \beta^*)' \Sigma^{*-1} (\hat{\beta}^* - \beta^*) \sim_r F_{(n-p-1)}$$

Assim, para verificar a contribuição conjunta de β^* para o modelo, devemos testar as hipóteses de forma similar ao caso de uma variável regressora:

$$H_0 : \beta^* = 0$$

$$H_1 : \beta^* \neq 0$$

Rejeitando-se H_0 se $\frac{1}{\sigma^2}(\hat{\beta}^* - \beta^*)' \Sigma^{*-1} (\hat{\beta}^* - \beta^*) >_r F_{(n-p-1)}$.

Agora, além de comprarmos o modelo múltiplo com o modelo simples, podemos comparar diversos modelos múltiplos entre si. Para efeitos de nomenclatura, chamaremos o modelo com todas as variáveis regressoras de **modelo completo** e o modelo com menos variáveis, conforme a hipótese H_0 , de **modelo reduzido**. Por exemplo, em um caso com diversas observações, supondo um modelo com variáveis x_1 e x_2 observáveis, podemos observar se o modelo linear é o mais adequado para descrever a tendência dos dados

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

ou podemos, de acordo com alguma hipótese advinda da física, ou de observação da dispersão dos dados, ou simplesmente como uma hipótese aleatória para tentar incrementar o ajuste, a introdução de um elemento quadrático de x_2 . Então o modelo anterior seria a forma reduzida e o novo modelo

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2$$

Seria o modelo completo. Logo a hipótese que seria testada é a da relevância da contribuição de β_3 para o modelo, ou seja

$$H_0 : \beta_3 = 0$$

$$H_1 : \beta_3 \neq 0$$

Então, para formalizar a generalização dos testes de hipótese, sejam β a matriz de parâmetros, com dimensão $p + 1$, e $\mathbf{X}$ a matriz de modelo, os dois são particionados na forma

$$\beta = \begin{bmatrix} \beta_{(1)} \\ \beta_{(2)} \end{bmatrix} \quad e \quad X = [X_{(1)} | X_{(2)}]$$

onde $\beta_{(1)}$ e $X_{(1)}$ são, respectivamente, os vetores de parâmetros e matriz do modelo **reduzido**; enquanto o índice (2) indica as variáveis extras que completam o modelo. Ou seja, $\beta_{(2)}$ possui dimensão r . Então agora, além de $SQR_{eg}(\beta_{(1)})$ e $SQR_{eg}(\beta)$, temos uma nova grandeza, denominada **soma de quadrados de regressão extra** com a adição de $\beta_{(2)}$ no modelo com $\beta_{(1)}$, com notação $SQR_{eg}(\beta_2 | \beta_{(1)}) = SQR_{eg}(\beta - SQR_{eg}(\beta_{(1)}))$.

Então, as hipóteses a serem testadas para comparar os modelos são

$$H_0 : \beta_{(2)} = \mathbf{0}$$

$$H_1 : \beta_{(2)} \neq \mathbf{0}$$

utilizando a estatística $\frac{SQReg(\beta_2|\beta_{(1)})/r}{\hat{\sigma}^2}$ com distribuição F, com r e $(n-p-1)$ graus de liberdade. Para calcular as devidas estatísticas para testar as hipóteses, podemos então dispôr das tabelas ANOVA para o modelo completo e modelo reduzido.

Tabela 2 – ANOVA para o modelo completo

FV	GL	SQ	QM	F_0
Regressão	p	$SQReg(\beta)$	$SQReg(\beta)/p$	$\frac{SQReg(\beta)/p}{SQE/(n-p-1)}$
Erro	$n-p-1$	SQE	$SQE/(n-p-1)$	
Total	$n-1$	SQT		

Tabela 3 – ANOVA para o modelo reduzido

FV	GL	SQ	QM	F_0
Regressão	$p-r$	$SQReg(\beta_1)$	$SQReg(\beta_{(1)})/(p-r)$	$\frac{SQReg(\beta_{(1)})/(p-r)}{SQE_{(1)}/(n-p-1+r)}$
Erro	$n-(p+1-r)$	$SQE_{(1)}$	$SQE/(n-p-1+r)$	
Total	$n-1$	SQT		

Utilizando a tabela do modelo completo e adicionando-se duas linhas contendo as variações de $(\beta_{(1)})$ e $(\beta_{(2)}|\beta_{(1)})$, obtemos uma tabela de ANOVA para comparar os dois modelos.

Tabela 4 – ANOVA para o modelo completo - SQ da regressão particionada

FV	GL	SQ	QM	F_0
Regressão	p	$SQReg(\beta)$		
$(\beta_{(1)})$	$(p-r)$	$(SQReg(\beta_{(1)}))$		
$(\beta_{(2)} \beta_{(1)})$	(r)	$(SQReg(\beta) - SQReg(\beta_{(1)}))$	$\frac{(SQReg(\beta) - SQReg(\beta_{(1)}))}{r}$	$\frac{(SQReg(\beta) - SQReg(\beta_{(1)}))/r}{\hat{\sigma}^2}$
Erro	$n-p-1$	SQE	$SQE/(n-p-1)$	
Total	$n-1$	SQT		

Para definirmos algumas grandezas futuramente, é importante definir a distribuição normal multivariada, a versão multidimensional da distribuição normal. Para isso, consideremos um vetor $\mathbf{W}$ de variáveis aleatórias, particionado em dois subvetores da forma

$$\mathbf{W} = \begin{bmatrix} Y \\ \mathbf{X} \end{bmatrix}$$

em que $\mathbf{X}$ é

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}$$

A função de densidade conjunta de $\mathbf{W}$ é dada por

$$f(\mathbf{W}) = \frac{1}{(2\pi)^{\frac{p+1}{2}} \sqrt{|\Sigma|}} \exp \left[-\frac{1}{2} (\mathbf{W} - \mu)' \Sigma^{-1} (\mathbf{W} - \mu) \right]$$

Onde: $\mu = E[\mathbf{W}] = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}_{(p+1)}$ e $\Sigma = \begin{bmatrix} \sigma_{11} & \sigma'_{12} \\ \sigma_{12} & \Sigma_{22} \end{bmatrix}_{(p+1) \times (p+1)}$ a matriz de variâncias e covariâncias de $\mathbf{W}$ particionada também em

$$\sigma_{11} = V[Y]$$

σ_{12} , o vetor das covariâncias entre Y e X_i com $i = 1, 2, \dots, p$ e Σ_{22} a matriz de variâncias e covariâncias do vetor $\mathbf{X}$.

Dessa forma, $\mathbf{W}$ segue uma distribuição conjunta normal multivariada, com vetor de médias μ e matriz de covariâncias Σ . De forma análoga à distribuição normal, utiliza-se a notação $\mathbf{W} \sim NM_{(p+1)}(\mu : \Sigma)$. Para qualquer vetor aleatório $\mathbf{A}$ com distribuição normal multivariada, valem as seguintes propriedades:

- Qualquer subvetor $\mathbf{B}$ de $\mathbf{A}$ possui distribuição normal multivariada, com dimensão igual à do subvetor e com os respectivos vetor de médias e matriz de variâncias e covariâncias.
- A distribuição condicional de um subvetor $\mathbf{B}$ dado um outro subvetor $\mathbf{C}$ também é normal multivariada de dimensão igual à de $\mathbf{B}$.
- Qualquer combinação linear dos elementos de $\mathbf{A}$ terá distribuição normal multivariada.

Dessa forma, podemos concluir que $\mathbf{X}$ segue distribuição normal multivariada com vetor de médias μ_2 e matriz de variâncias e covariâncias Σ_{22} . Isso implica que sua função de densidade conjunta é dada por

$$f_2(x) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma_{22}|^{\frac{1}{2}}} \exp \left[-\frac{1}{2} (x - \mu_2)' \Sigma_{22}^{-1} (x - \mu_2) \right]$$

Além disso, a distribuição condicional de Y dado x é normal univariada com função de densidade dada por

$$h(y|x) = \frac{1}{\sqrt{2\pi\sigma_{11.2}}} \exp \left[-\frac{1}{2} \frac{(y - \mu_1 - (\Sigma_{22}^{-1} \sigma_{12})'(x - \mu_2))}{\sigma_{11.2}} \right]$$

onde $\sigma_{11.2} = \sigma_{11} - \sigma'_{12} \Sigma_{22}^{-1} \sigma_{12}$.

Assim, temos que a esperança condicional de $\mathbf{Y}$ dado $\mathbf{x}$ é expressa por

$$E[Y|\mathbf{x}] = \mu_1 + (\Sigma_{22}^{-1} \sigma_{12})'(\mathbf{x} - \mu_2)$$

enquanto a variância condicional é expressa por

$$V[Y|\mathbf{x}] = \sigma_{11.2}$$

constante em relação ao subvetor $\mathbf{x}$.

Assim, podemos analisar $\sigma_{11.2}$ como σ_{11} menos a redução da variância de Y quando condicionamos Y ao subvetor $\mathbf{x}$. Dessa forma, observando essa redução de variância, temos

$$\sigma'_{12}\Sigma_{22}^{-1}\sigma_{12} = \sigma_{11} - V[Y|\mathbf{x}] = V[Y] - V[Y|\mathbf{x}]$$

Dividindo os lados por $V[Y]$, temos:

$$\frac{\sigma'_{12}\Sigma_{22}^{-1}\sigma_{12}}{\sigma_{11}} = \frac{\sigma'_{12}\Sigma_{22}^{-1}\sigma_{12}}{V[Y]} = \frac{V[Y] - V[Y|\mathbf{x}]}{V[Y]}$$

Essa expressão pode ser interpretada como a redução proporcional da variância, podendo-se analisá-la independente da escala observada. Portanto, podemos defini-la como Coeficiente de determinação populacional, denotado por

$$\rho_{12\dots p+1}^2 = \frac{\sigma'_{12}\Sigma_{22}^{-1}\sigma_{12}}{\sigma_{11}}$$

6.2.1 Análise dos resíduos

O objetivo da análise dos resíduos é de identificar a adequação do modelo às suposições iniciais e conferir a contribuição de cada variável regressora para a capacidade preditiva do modelo.

Ainda por questões de notação, denotaremos por x'_i a i -ésima linha da matriz de modelo $\mathbf{X}$:

$$\mathbf{X} = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{bmatrix}$$

O vetor de valores estimados de esperanças de Y é dado pela multiplicação da matriz de modelo por $\hat{\beta}$

$$\begin{bmatrix} x'_1 \hat{\beta} \\ x'_2 \hat{\beta} \\ \vdots \\ x'_n \hat{\beta} \end{bmatrix} = \mathbf{X} \hat{\beta}$$

e o vetor de resíduos é dado por

$$\begin{bmatrix} y_1 - x'_1 \hat{\beta} \\ y_2 - x'_2 \hat{\beta} \\ \vdots \\ y_n - x'_n \hat{\beta} \end{bmatrix} = y - \mathbf{X} \hat{\beta} = \mathbf{e}$$

Cuja distribuição é dada por

$$\mathbf{X} \hat{\beta} \sim N(\mathbf{X}\beta; \sigma^2 \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')$$

A esperança e variâncias dos resíduos podem ser desenvolvidas, respectivamente, por

$$\begin{aligned} E[\mathbf{e}] &= E[y - \mathbf{X} \hat{\beta}] \\ &= E[y] - E[\mathbf{X} \hat{\beta}] \\ &= \mathbf{X} \hat{\beta} - \mathbf{X} \hat{\beta} = \mathbf{0} \end{aligned}$$

$$\begin{aligned} Var[\mathbf{e}] &= Var[y - \mathbf{X} \hat{\beta}] \\ &= Var[y] + Var[\mathbf{X} \hat{\beta}] - 2Cov[y, \mathbf{X} \hat{\beta}] \\ &= \sigma^2 I + \sigma^2 \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - 2Cov[y, \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y] \\ &= \sigma^2 I + \sigma^2 \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - 2\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'Cov[y, y] \\ &= \sigma^2 I + \sigma^2 \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - 2\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\sigma^2 I \\ &= \sigma^2 I - \sigma^2 \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \\ &= \sigma^2 [I - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'] \end{aligned}$$

Ou seja, podemos escrever que

$$\mathbf{e} \sim NM_n(\mathbf{0}; \sigma^2[I - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'])$$

Para efeitos de simplificação de notação, podemos definir $\mathbf{H}$ como

$$\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

ou seja: $\hat{y} = \mathbf{X}\hat{\beta} = \mathbf{H}y$

Adicionalmente, definimos também os elementos da diagonal principal de $\mathbf{H}$ como "valores $\mathbf{h}$ ", ou seja, $\mathbf{h} = h_{ii}, i = 1, \dots, n$, onde $0 < h_{ii} < 1, i = 1, \dots, n$. Além disso, alguns resultados úteis não demonstrados podem ser conferidos a seguir. Primeiramente

$$h_{ii} = \sum_{j=1}^n h_{ij}^2$$

O que adiciona a interpretação de que o valor $\mathbf{h}$, h_{ii} é uma forma de avaliar a influência da i -ésima observação na composição de cada $\hat{y}_j, j = 1, 2, \dots, n$. Além disso:

$$\sum_{i=1}^n h_{ii} = k$$

Onde k é o número de colunas de $\mathbf{X}$. Isso implica que k/n é a média dos valores $\mathbf{h}$.

Podemos, então, para auxiliar cálculos mais à frente, definir duas novas formas de resíduos. Os resíduos padronizados, denotados z_i definidos como

$$z_i = \frac{e_i}{\sqrt{\hat{\sigma}(1 - h_{ii})}}, i = 1, \dots, n$$

E resíduos estudentizados, definidos como

$$z_i^* = \frac{e_i}{\sqrt{\hat{\sigma}_{(i)}(1 - h_{ii})}}, i = 1, \dots, n$$

Onde $\hat{\sigma}_{(i)}$ é a soma média dos quadrados dos resíduos de MRLM, ajustado **sem** a i -ésima observação. Dessa forma, z_i^* tem distribuição t de Studente com $(n - k - 1)$ graus de liberdade, cuja hipótese supões que o MRLM seja adequado.

Para auxiliar-nos a avaliar a influência de cada observação na estimativa dos coeficientes, podemos definir a estatística de Cook, ou distância de Cook (Cook, 1982), que pode ser

expressa de duas formas, ambas denotadas por D_i . A primeira sendo

$$D_i = \frac{(\hat{\beta} - \hat{\beta}_{(i)})' \mathbf{X}' \mathbf{X} (\hat{\beta} - \hat{\beta}_{(i)})}{k \hat{\sigma}^2}, i = 1, \dots, n$$

Onde $\hat{\beta}_{(i)}$ é o estimador de quadrados mínimos do vetor β omitindo-se a i -ésima observação.

A segunda forma é dada por

$$D_i = (z_i)^2 \frac{h_{ii}}{1 - h_{ii}} \frac{1}{k}$$

Das duas formas, quanto maior a Estatística de Cook, maior a influência daquela observação na estimativa dos coeficientes da regressora.

Para auxiliar também na investigação do papel de cada variável regressora no modelo, definiremos também os resíduos parciais, denotados por e_{ij}^* dados por

$$e_{ij}^* = y_i - \mathbf{x}_i' \hat{\beta} + x_{ij} \hat{\beta}_j$$

Onde x_{ij} é o j -ésimo elemento da linha $\mathbf{x}_i'$ da matriz $\mathbf{X}$ e $\hat{\beta}_j$ é o j -ésimo elemento do vetor $\hat{\beta}$.

6.2.2 Validação dos modelos múltiplos

O pressuposto para avaliação dos resíduos $\epsilon = Y - \hat{Y}$ é a forma de sua distribuição, normal, bem como a homocedasticidade e independência entre seus resíduos:

- Normalidade:

$$\epsilon \sim N(0, \sigma_\epsilon^2)$$

- Independência

$$f(\epsilon) = \prod f(\epsilon_i)$$

Tais exigências são "muito fortes", mas necessárias para o ajuste adequado. Quando a independência é verificada, a correlação será nula $cor(\epsilon_i, \epsilon_j) = 0; \quad \forall i \neq j$.

- Homocedasticidade

$$H_0 : \sigma_i^2 = \sigma_j^2; \quad \text{para } i \neq j$$

6.2.2.1 Teste de hipótese para a inclinação β_i :

Supondo a normalidade para a i -ésima inclinação β_i , a hipótese a ser testada é:

- Hipótese:

$$\begin{cases} H_0 : \beta_i = 0 \\ H_1 : \beta_i \neq 0 \end{cases}$$

- Quantidade Pivotal: Seja a distribuição t-Student

$$t = \frac{\hat{\beta}_i - \beta_i}{s/\sqrt{s_{xx}}} \quad \forall i = 1, 2, 3 \dots$$

com $n - 2$ graus de liberdade. Quando a hipótese nula não for rejeitada, o coeficiente angular é dito significativo e deverá estar no modelo.

O intervalo de confiança é dado por:

$$\hat{\beta}_i - t_{\alpha/2} \cdot s/\sqrt{s_{xx}} < \beta_i < \hat{\beta}_i + t_{\alpha/2} \cdot s/\sqrt{s_{xx}}$$

O intervalo de confiança é um indicativo da variabilidade do coeficiente estimado.

6.2.2.2 Teste de hipótese para o intercepto β_0 :

Supondo a normalidade para o intercepto β_0 , a hipótese a ser testada é:

- Hipótese:

$$\begin{cases} H_0 : \beta_0 = 0 \\ H_1 : \beta_0 \neq 0 \end{cases}$$

- Quantidade Pivotal: Seja a distribuição t-Student

$$t = \frac{\hat{\beta}_0 - \beta_0}{s/\sqrt{\sum_{i=1}^n x_i^2/(ns_{xx})}}$$

com $n - 2$ graus de liberdade.

Se a hipótese nula não for rejeitada, o intercepto é significativo e diferente de zero. Pode-se também avaliar através do intervalo de confiança:

$$\hat{\beta}_0 - t_{\alpha/2} \cdot s/\sqrt{\sum_{i=1}^n x_i^2/(ns_{xx})} < \beta_0 < \hat{\beta}_0 + t_{\alpha/2} \cdot s/\sqrt{\sum_{i=1}^n x_i^2/(ns_{xx})}$$

6.3 Método dos Mínimos Quadrados Generalizado

Existe uma diferença entre o métodos dos mínimos quadrados generalizados (MMQG) e dos ordinários (MMQO). Quando as premissas para os resíduos, homocedasticidade, por exemplo, desse último não são atendidas (violadas), a técnica generalizada MMQG minimiza os quadrados dos resíduos levando em consideração a estrutura da correlação e variância dos erros, o que o MMO não o faz. Sejam as novas premissas:

- **Esperança do resíduo:** Assumimos que o erro médio depende da observação $\mathbf{X}$,

$$E(\epsilon|\mathbf{X}) = 0$$

para quaisquer observação.

- **Variância do resíduo:** Dado que a definição da variância é o segundo momento centrado:

$$\sigma^2 = E(X^2) - E^2(X)$$

podemos assumir que a média é o tópico anterior $E(\epsilon|\mathbf{X}) = 0$, então o segundo momento não centrado é a própria variância

$$E(\epsilon^T \epsilon|\mathbf{X}) = \sigma^2 \Omega$$

sendo Ω a matriz positiva definida.

$$\sigma^2 \Omega = \begin{bmatrix} \sigma_{11}^2 & a_{12} & a_{13} & . & . & . & a_{1n} \\ a_{21} & \sigma_{22}^2 & x_{23} & . & . & . & a_{2n} \\ a_{31} & x_{32} & \sigma_{33}^2 & . & . & . & a_{3n} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ a_{n1} & a_{n2} & a_{n3} & . & . & . & a_{nn} \end{bmatrix}.$$

De tal forma que a distribuição dos resíduos para *MMQG* será dada por

$$\mathbf{e} \sim NM_n(\mathbf{0}; [(\mathbf{X}'\mathbf{X})^{-1}\sigma^2\Omega(\mathbf{X}\mathbf{X}')])$$

Para o caso especial, em que $\Omega = \mathbf{1}$, assume a matriz identidade, os resíduos recaem na forma ordinária (MMQO). Nesse caso, a exigência da **homocedasticidade** para o caso ordinário leva a matriz seguinte:

$$\sigma^2 \mathbf{\Omega} = \sigma^2 \mathbf{1} = \begin{bmatrix} \sigma_{11}^2 & 0 & 0 & . & . & . & 0 \\ 0 & \sigma_{22}^2 & 0 & . & . & . & 0 \\ 0 & 0 & \sigma_{33}^2 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & \sigma_{nn}^2 \end{bmatrix}.$$

Observe que a homocedasticidade é um caso especial para $\sigma^2 \mathbf{\Omega}$.

Para o caso geral minimizamos o erro por

$$\epsilon^T \epsilon = (y - \mathbf{X}\beta)^T \mathbf{\Omega}^{-1} (y - \mathbf{X}\beta)$$

sendo resolvido pelas mesmas técnicas MMQO, resultado em

$$\hat{\beta} = (\mathbf{X}^T \mathbf{\Omega}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{\Omega}^{-1} \mathbf{Y}$$

A matriz $\mathbf{\Omega}$ pode assumir qualquer forma, inclusive uma matriz de variâncias e covariâncias, com $E(\epsilon_t \epsilon_{t-k}) = \gamma(k)$:

$$\sigma^2 \mathbf{\Omega} \rightarrow \mathbf{\Sigma}$$

$$\sigma^2 \mathbf{\Omega} = \mathbf{\Sigma} = \begin{bmatrix} \sigma_{11}^2 = \gamma(0) & \gamma(1) & \gamma(2) & . & . & . & \gamma(t-1) \\ \gamma(1) & \sigma_{22}^2 = \gamma(0) & \gamma(3) & . & . & . & \gamma(t-1) \\ \gamma(2) & \gamma(3) & \sigma_{33}^2 = \gamma(0) & . & . & . & \gamma(t-3) \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ \gamma(t-1) & \gamma(t-2) & 0 & . & . & . & \sigma_{nn}^2 = \gamma(0) \end{bmatrix}.$$

6.4 Planejamento de Experimentos: fatorial 2^k

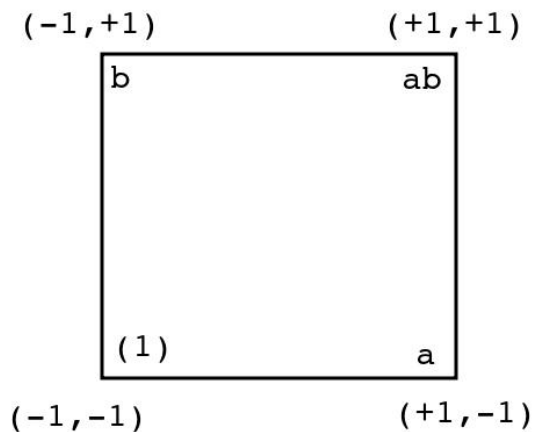
O planejamento experimental é muito importante no setor farmacêutico, produção de alimentos, agronomia e indústria de transformação. A técnica experimental fatorial 2^k utiliza-se das técnicas de regressão múltipla para experimentos laboratoriais ou em campo. Aqui apresentamos para o caso $k = 2$, sendo o mesmo mecanismo para k superiores.

6.4.1 Experimento 2^k

Suponha dois fatores A, B, com dois níveis cada um + e -. Cada fator representa uma medida distinta da outra. Pode ser volume vs. pressão, carga vs. tensão, nível de glicose vs. pressão sistólica, ou quaisquer outros pares de medida. Cada fator apresenta dois níveis, um limite mínimo (-) e um máximo (+). Cada observação (1), a, b e ab, contemplam com um pool de n observações, vide figura 3.

$$2^k = p$$

Figura 3 – Planejamento 2^k



Fonte: Autoria própria.

Tabela 5 – A forma experimental.

	A		Média
B	b	ab	$\frac{b+ab}{2n}$
	(1)	a	$\frac{(1)+a}{2n}$
Média	$\frac{(1)+b}{2n}$	$\frac{a+ab}{2n}$	

Fonte: Autoria própria.

Suponha a resposta da diferença entre as médias do nível alto (H) com relação ao baixo (L);

$$\text{efeito} = \bar{y}_H - \bar{y}_L$$

Então, podemos definir os efeitos principais:

$$A = \bar{y}_{H_A} - \bar{y}_{L_A} = \frac{a + ab}{2n} - \frac{(1) + b}{2n} = \frac{ab + a - b - (1)}{2n}$$

$$A = \frac{ab + a - b - (1)}{2n} = \frac{\text{Contraste A}}{2n}$$

e para o efeito B;

$$B = \bar{y}_{H_B} - \bar{y}_{L_B} = \frac{b + ab}{2n} - \frac{(1) + a}{2n} = \frac{ab - a + b - (1)}{2n}$$

$$B = \frac{ab - a + b - (1)}{2n} = \frac{\text{Contraste B}}{2n}$$

Além dos efeitos principais A e B, temos a interação entre ambos os efeitos, que chamamos de efeito de interação AB:

$$AB = \frac{(1) - a - b + ab}{2n} = \frac{\text{Contraste AB}}{2n}$$

6.4.2 Relação com modelos de regressão

O efeito é definido por:

$$\text{efeito} = \frac{\text{contraste}}{2^{k-1}(n)}$$

resultando nos coeficientes de regressão linear

$$\text{Coeficiente de regressão} = \frac{\text{contraste}}{2^k(n)} = \frac{\text{efeito}}{2}$$

Para um modelo

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \beta_{p_1 p_2} x_{p_1 p_2} \equiv y = \beta_0 + \sum_{i=1}^{m < p} \beta_i x_i + \sum_{i < j}^p \beta_{ij} x_i x_j$$

Para a seguinte forma

$$\mathbf{Y} = \mathbf{X}\beta$$

temos

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_{12} \end{bmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

com $\mathbf{y}$

$$\mathbf{y} = \begin{bmatrix} (1) \\ a \\ b \\ ab \end{bmatrix}$$

sendo $\mathbf{X}$ a matriz composta pelos termos

$$\begin{aligned} (-1) &= [1, 1, 1, 1] \\ a &= [-1, 1, -1, 1] \\ b &= [1, -1, 1, -1] \\ a * b = ab &= [1, -1, -1, 1] \end{aligned}$$

Que compõe a matriz:

$$\mathbf{X} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} = 2^2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Observe que aparece o termo 2^k para $k = 2$ fatores: A e B.

Ao encontrar a inversa temos:

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 0.25 & 0 & 0 & 0 \\ 0 & 0.25 & 0 & 0 \\ 0 & 0 & 0.25 & 0 \\ 0 & 0 & 0 & 0.25 \end{bmatrix} = \frac{1}{2^2} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \frac{1}{2^2} \mathbf{I}$$

Portanto, para 2^k fator temos o resultado para apenas uma única réplica

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{2^k} \mathbf{I}$$

Em

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \dots \\ \beta_{p_1 p_2} \end{bmatrix} = \frac{1}{2^k} \mathbf{I}(\mathbf{X}'\mathbf{y})$$

se $k=2$ termos $(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{4}$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_{12} \end{bmatrix} = \frac{1}{2^2} \mathbf{X}'\mathbf{y} = \frac{1}{2^2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} (1) \\ a \\ b \\ ab \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_{12} \end{bmatrix} = \begin{bmatrix} \frac{(1)+a+b+ab}{2^2} \\ \frac{-(1)+a-b+ab}{2^2} \\ \frac{-(1)-a+b+ab}{2^2} \\ \frac{(1)-a-b+ab}{2^2} \end{bmatrix}$$

Por fim, teremos o seguinte modelo de regressão experimental

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_{12}$$

$$\begin{aligned} y = & \left(\frac{(1) + a + b + ab}{2^2} \right) + \left(\frac{-(1) + a - b + ab}{2^2} \right) x_1 \\ & + \left(\frac{-(1) - a + b + ab}{2^2} \right) x_2 + \left(\frac{(1) - a - b + ab}{2^2} \right) x_{12} \end{aligned}$$

Para o caso de n-réplica teremos

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{2^k \cdot (n)} \mathbf{I}$$

Em

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \dots \\ \beta_{p_1 p_2} \end{bmatrix} = \frac{1}{2^k \cdot (n)} \mathbf{I}(\mathbf{X}'\mathbf{y})$$

6.5 Aplicações

As aplicações a seguir, demonstram alguns exemplos teóricos e práticos que utilizam MLM nos usos de algumas ferramentas matemáticas.

6.5.1 Exemplos de aplicações dos MLMs

Apresentamos aqui exemplos aplicáveis de conceitos teóricos e práticos com o uso da regressão múltipla. Tal regressão auxilia o professor na apresentação e explicação dos conceitos de nível superior aplicado em áreas técnicas.

- **Estimando coeficientes** Em alguns casos nos cursos técnicos, o profissional deverá determinar os coeficientes por linearização. Um caso clássico na literatura é a equação dos gases ideais. Essa equação depende de duas constantes γ e C , os quais devem ser estimados por regressão.

– a) Exemplo: Estimação de parâmetros físicos

A pressão P de um gás que corresponde a vários volumes de V é registrada.

$V \text{ (cm}^3\text{)}$	50	60	70	90	100
$P \text{ (kg/cm}^2\text{)}$	64,7	51,3	40,5	25,9	7,8

A equação de estado dos gases ideais é dada pela forma funcional

$$PV^\gamma = C$$

Encontre as constantes γ e C .

R: Para encontrarmos uma função linear, vamos aplicar o logarítmo em ambos os lados da equação:

$$PV^\gamma = C \cdot \epsilon$$

$$\ln PV^\gamma = \ln C \cdot \epsilon$$

$$\ln P + \gamma \ln V = \ln C + \ln \epsilon$$

$$\ln P = \ln C - \gamma \ln V + \ln \epsilon$$

Observe que a equação fica linearizada e podemos associar os elementos para $y = \beta_0 + \beta_1 x + \epsilon$ por:

$$y = \ln P$$

$$x = \ln V$$

$$\beta_0 = \ln C$$

$$\beta_1 = -\gamma$$

Encontrando os coeficientes $\hat{\beta}_0$ e $\hat{\beta}_1$ temos:

$$\hat{\beta}_1 = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Para resolver, podemos usar uma tabela prática e calcular os coeficientes:

Tabela 6 – Tabela auxiliar para cálculo.

V	P	$x = \ln V$	$y = \ln P$	x^2	xy
64,7	50	4,170	3,912	15,304	16,312
51,3	60	3,938	4,094	16,764	16,122
40,5	70	3,701	4,248	18,050	15,725
25,9	90	3,254	4,500	20,248	14,643
7,8	100	2,054	4,605	21,208	9,460
	soma	17,117	21,360	91,573	72,262

Fonte: Autoria própria.

Encontrando os coeficientes:

$$\hat{\beta}_1 = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2} = -2,65347$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 14,7589$$

Ou seja,

$$\hat{\beta}_1 = -\gamma = -2,65347 \rightarrow \gamma = 2,65347$$

$$\hat{\beta}_0 = \ln \hat{C} = 14,7589 \rightarrow \hat{C} = e^{14,7589} \rightarrow \hat{C} = 2.568.862,88$$

A equação estimada foi

$$PV^{2,65} = 2.568.862$$

No entanto, para saber se a equação é válida, precisamos avaliar os resíduos de ϵ e aqui cabe uma observação. Ao avaliar o modelo proposto, a resposta da homocedasticidade

não é atendida para o modelo logarítmico, necessitando de um modelo de segunda ordem.

- **Conceitos de demografia e geografia**

Outra aplicação seria a interação com as outras disciplinas do ensino médio, como por exemplo a demografia dos países dentro da geografia. Um dos conceitos que mais geram dúvidas são os conceitos de fertilidade e natalidade. Enquanto a primária é a capacidade da população gerar novos indivíduos a segunda refere-se à quantidade de nascidos vivos. Para uma população se manter estável é necessário que a fertilidade encontra-se na quantidade de 2,1 filhos por mulher. O exemplo a seguir mostra-nos o ajuste linear com intervalos de confiança para a estimação.

- **b) Fertilidade**

Suponha os países da América do Sul. Verifique a taxa de fertilidade por idade. Quais países estão com menor taxa de reposição populacional?

Tabela 7 – Tabela taxa de fertilidade por países da América do Sul.

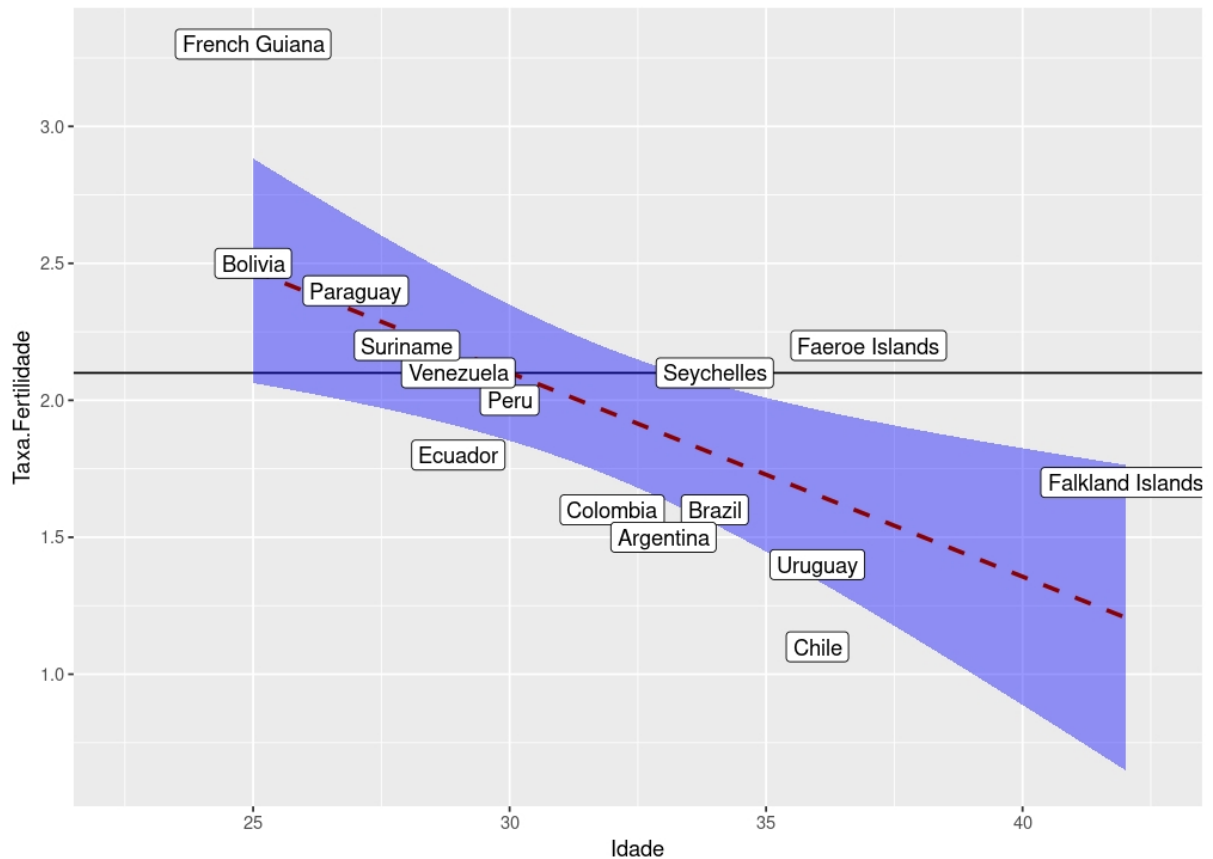
País	Idade	Taxa.Fertilidade	pop
Brazil	34	1.6	211998573
Colombia	32	1.6	52886363
Argentina	33	1.5	45696159
Peru	30	2.0	34217848
Venezuela	29	2.1	28405543
Chile	36	1.1	19764771
Ecuador	29	1.8	18135478
Bolivia	25	2.5	12413315
Paraguay	27	2.4	6929153
Uruguay	36	1.4	3386588
Suriname	28	2.2	634431
French Guiana	25	3.3	308522
Seychelles	34	2.1	130418
Faeroe Islands	37	2.2	554
Falkland Islands	42	1.7	347

Fonte: Autoria própria.

R: Podemos usar a regressão linear simples, mas com intervalo de confiança. Esse intervalo mede a variabilidade dos coeficientes linear e angular da regressão. A figura seguinte mostra que a fertilidade decai com a idade. Países com maior idade média tendem a ter menor taxa de fertilidade, impactando na reposição da população. A figura mostra-nos um ajuste de $R^2 = 44\%$ para

$$\text{Taxa de fertilidade} = 4.33413 - 0.07445 \text{ Idade}$$

Figura 4 – Taxa de fertilidade por idade nos países



Fonte: Autoria própria.

Isso mostra que há queda na América do Sul com relação à fertilidade entre os países. A figura mostra-nos uma linha horizontal em 2.1 filhos por mulher. Os países Peru, Equador, Colômbia, Brasil, Argentina, Uruguai, Chile e as ilhas inglesas Falkland estão abaixo da linha de reposição. Somando essas populações representam 88,7% da população do subcontinente que futuramente podem ter eventuais problemas de reposição de pessoal no mercado de trabalho e previdência.

• Utilizando equações EDO

Obviamente as equações EDO são difíceis de explicar para cursos que são de nível técnico. Nesse caso, a regressão múltipla pode ser um aliada na compreensão do mecanismo das equações diferenciais. O exemplo a seguir, refere-se à área da Engenharia Biomédica: Estimar a forma da abertura de lesões. Uma possibilidade é usar regressões de banco de dados reais e mostrar os conceitos de campo (gradiente, divergente e rotacional) usando funções geradas por regressões múltiplas. A seguir, mostramos um exemplo de câmeras térmicas para avaliar lesões cutâneas de decúbito em UTI's. A proposta é apresentar os conceitos físicos sobre aumento/ diminuição térmica usando conceitos matemáticos do campo vetorial. Dessa forma, o professor pode demonstrar na prática e com dados reais o funcionamento do campo vetorial na

área de saúde, por exemplo, para cursos de técnico em enfermagem, técnico em radiologia ou afins. Outra forma de se utilizar a regressão é para avaliar casos com alto volume de dados cuja interação ocorre de maneira muito complexa, aproximando uma função suave e contínua de uma situação real.

– **b) Exemplo na área biomédica**

Como no caso visto em (Voziniak *et al.*, 2023), lesões por pressão são um tópico complexo nas áreas da medicina, enfermagem e engenharia biomédica. A equação que modela a perfusão sanguínea, ou seja, a quantidade de sangue que irriga o tecido, é dada pela equação de Pennes (Hh., 1948):

$$\begin{aligned} \rho_{skin} c_{skin} \frac{\partial T_{skin}}{\partial t} &= k_{skin} \nabla^2 T_{skin} \\ &- \omega_b \rho_b c_b (T_{skin} - T_a) \\ &+ q_{met} + q_s \end{aligned}$$

R: Segundo Voziniak *et al.*(2023), apesar de a câmera térmica capturar apenas a temperatura, a equação de Pennes utiliza o Laplaciano da temperatura, o qual pode ser obtido utilizando pacotes de métodos numéricos. Dessa forma, a imagem gerada pela câmera, uma matriz cujas células correspondem a pixels da imagem e a cada uma é atribuída um valor de temperatura, pode ser convertida numa matriz em que a cada pixel é atribuído um valor do Laplaciano da temperatura naquele ponto. Utilizando do ajuste de curvas na câmera térmica, encontramos o resultado:

$$\begin{aligned} T &= \beta_0 + \beta_1 x + \beta_2 y + \beta_3 x^2 + \beta_4 y^2 + \beta_5 xy + \beta_6 x^3 + \beta_7 x^2 y + \beta_8 xy^2 + \beta_9 x^4 \\ &+ \beta_{10} y^4 + \beta_{11} x^2 y^2 + \beta_{12} x^5 + \beta_{13} y^5 + \beta_{14} x^3 y^2 + \beta_{15} x^2 y^3 + \beta_{16} xy^4 + \beta_{17} x^4 y^1 \end{aligned}$$

via

$$\hat{\beta} = (X'X)^{-1}X'y$$

O gradiente encontrado foi:

$$\begin{aligned}
\vec{\nabla}T &= [\beta_1 + 2\beta_3x + \beta_5y + 3\beta_6x^2 + 2\beta_7xy + \beta_8y^2 \\
&+ 4\beta_9x^3 + 2\beta_{11}xy^2 + 5\beta_{12}x^4 + 3\beta_{14}x^2y^2 \\
&+ 2\beta_{15}xy^3 + \beta_{16}y^4 + 4\beta_{17}x^3y^1]\hat{i} \\
&+ [\beta_2 + 2\beta_4y + \beta_5x + \beta_7x^2 + 2\beta_8xy + 4\beta_{10}y^3 \\
&+ 2\beta_{11}x^2y + 5\beta_{13}y^4 + 2\beta_{14}x^3y + 3\beta_{15}x^2y^2 \\
&+ 4\beta_{16}x^1y^3 + \beta_{17}x^4]\hat{j} \\
&+ 0\hat{k}
\end{aligned}$$

Figura 5 – O problema real por termografia

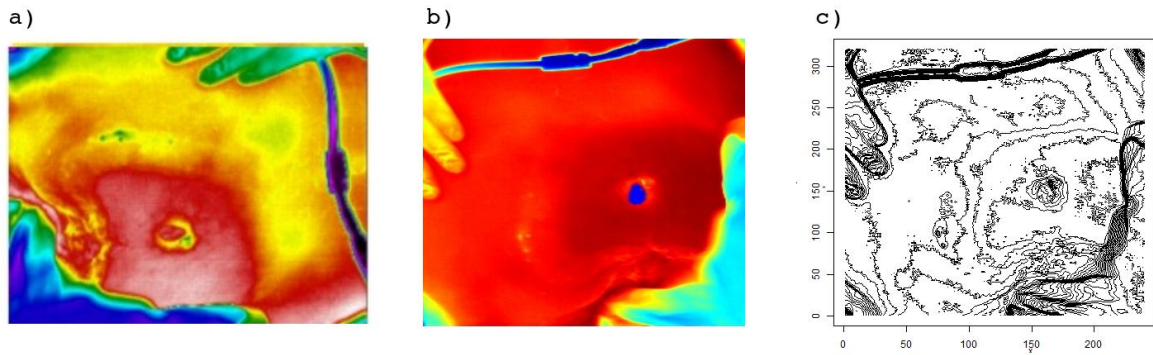
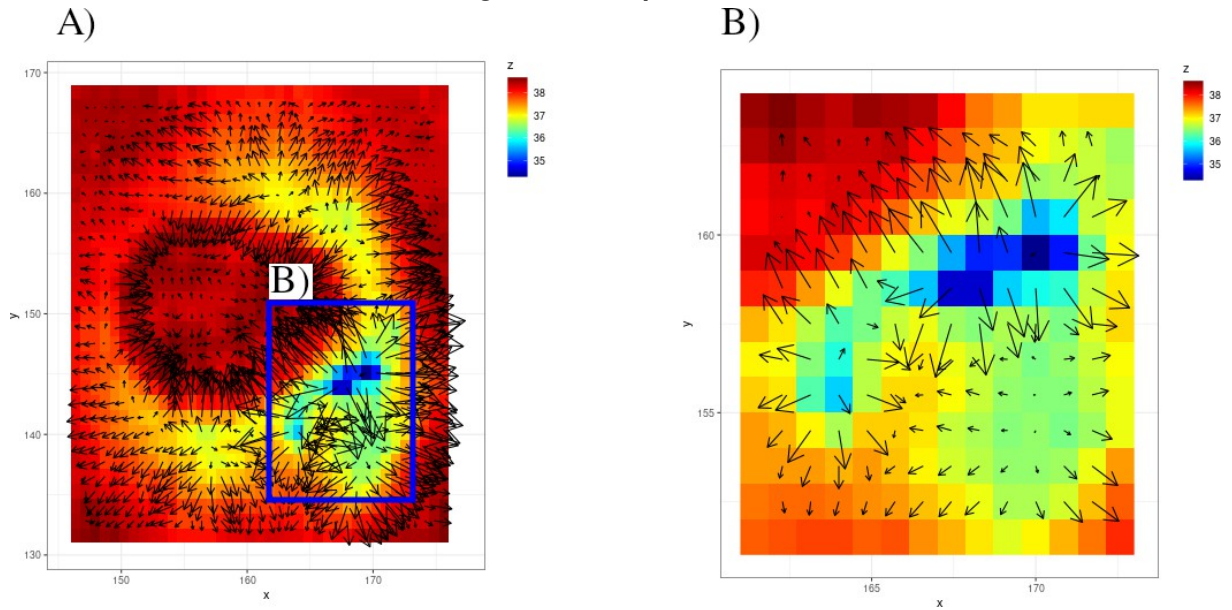


Figura 6 – Campo Vetorial



O laplaciando encontrado foi

$$\begin{aligned}\nabla^2 T &= 2\beta_3 + 6\beta_6 x + 2\beta_7 y + 12\beta_9 x^2 + 2\beta_{11} y^2 + 20\beta_{12} x^3 + 6\beta_{14} xy^2 \\ &+ 2\beta_{15} y^3 + 12\beta_{17} x^2 y + 2\beta_4 + 2\beta_8 x + 12\beta_{10} y^2 + 2\beta_{11} x^2 \\ &+ 20\beta_1 x^3 + 2\beta_{14} x^3 + 6\beta_{15} x^2 y + 12\beta_{16} xy^2\end{aligned}$$

Dependendo da região, podemos obter $\nabla^2 T > 0$ mostrando perda de calor e $\nabla^2 T < 0$ aumento de temperatura local. Uma vez que o modelo é ajustável para a temperatura T , podemos encontrar o gradiente da função quanto o laplaciano. Se o laplaciano for positivo, significa que a região está perdendo calor e, deve-se ao fechamento da ferida ou, a depender do valor temperatura pode ser indício de necrose local. Quando contrário, o laplaciano for negativo pode ser reperusão sanguínea ou sangramento local. Essa técnica foi estudada na iniciação científica.

• Experimentação

O professor pode usar de experimentos em sala de aula ou de resultados já observados e publicados. Assim, técnicos em laboratório poderão compreender como são feitos experimentos controlados e como os resultados são encontrados e fornecidos. O interessante aqui é a utilização frequente das operações matriciais para ajuste do modelo. Podemos citar como exemplos, curvas de operação de tiristores, diodos, curvas de cargas de motores e, inclusive, efeitos de corrosão, como no caso do exemplo a seguir.

– c) Exemplo experimental

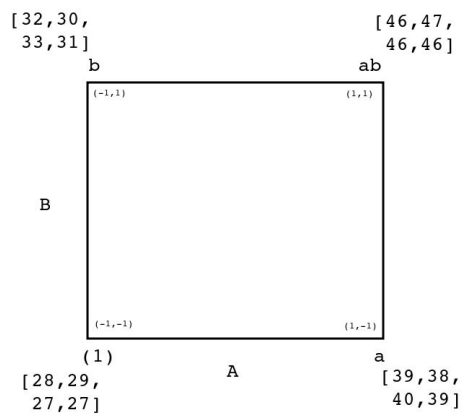
O dióxido de carbono na presença da água, pode causar a corrosão em metais aço carbono devido a formação de ácido carbônico H_2CO_3 que é provocado por reações eletroquímicas que abrangem a dissolução anódica do metal e a evolução catódica do hidrogênio provocando a corrosão. Suponha um experimento fatorial 2^k para avaliar a corrosão do aço carbono sob condições da presença de CO_2 (421 ppm a 580 ppm) e temperatura da água (50 $^{\circ}C$ a 100 $^{\circ}C$). Mensura-se a quantidade de substrato oxidado no metal por $\mu g/mm^3$. Avalie se existe interação ente a temperatura da água e a presença de CO_2 no efeito da corrosão.

Para isso define-se dois fatores: A) presença de CO_2 2 níveis: valor mínimo 421 ppm (-) e máximo de 580 ppm (+) e B) temperatura da água ambiente T com dois níveis: valor mínimo 50 $^{\circ}C$ e máximo de 100 $^{\circ}C$ (+).

Cada medida foi feita quatro réplicas conforme mostrado na figura seguinte:

Tabela 8 – Tabela experimentação 2^k

Fator		interação	Combinação	Réplicas			
A	B	AB		I	II	III	IV
-	-	+	baixo, baixo	28	29	27	27
+	-	-	alto, baixo	39	38	40	39
-	+	-	baixo, alto	46	47	46	46
+	+	+	alto, alto	32	30	33	31

Fonte: Autoria própria.**Figura 7 – Planejamento 2^k para o experimento de corrosão****Fonte: Autoria própria.**

$$X = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Fazendo $\mathbf{X}'\mathbf{X}$, obtemos

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 16 & 0 & 0 & 0 \\ 0 & 16 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 16 \end{bmatrix}$$

Encontrando a inversa:

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 0.0625 & 0 & 0 & 0 \\ 0 & 0.0625 & 0 & 0 \\ 0 & 0 & 0.0625 & 0 \\ 0 & 0 & 0 & 0.0625 \end{bmatrix} = \frac{1}{16} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \frac{1}{2^2 \cdot (4)} \cdot \mathbf{I}$$

Observe que temos 2^2 pontos com 4 réplicas cada uma. O próximo passo é encontrar os coeficientes do modelo do experimento:

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_{12} \end{bmatrix} = \frac{1}{2^k \cdot (n)} \mathbf{I}(\mathbf{X}'\mathbf{y})$$

Para $\mathbf{Y}$ temos

$$\mathbf{Y} = \begin{bmatrix} 28 \\ 39 \\ 32 \\ 46 \\ 29 \\ 38 \\ 30 \\ 47 \\ 27 \\ 40 \\ 33 \\ 46 \\ 27 \\ 39 \\ 31 \\ 46 \end{bmatrix}$$

Calculando, temos o seguinte resultado

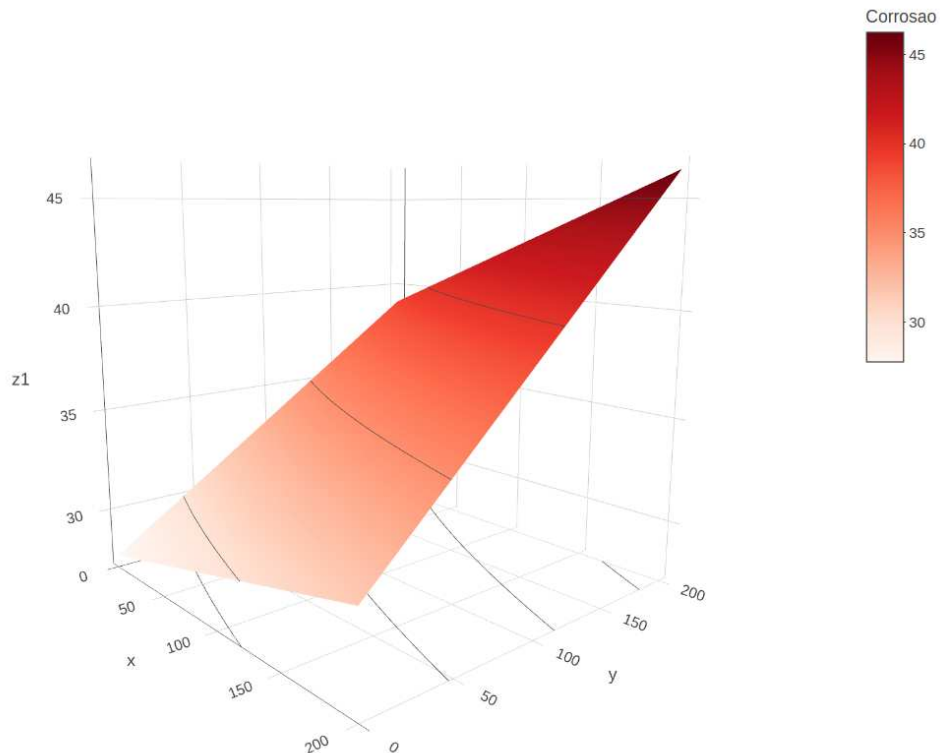
$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_{12} \end{bmatrix} = \frac{1}{2^2 \cdot (4)} \mathbf{I}(\mathbf{X}'\mathbf{y}) = \begin{bmatrix} 36,125 \\ 6,500 \\ 2,750 \\ 0,875 \end{bmatrix}$$

Resultado no modelo

$$y = 36,125 + 6,500x_1 + 2,750x_2 + 0,875x_1x_2$$

Sendo que o modelo pode ser visualizado pela superfície.

Figura 8 – Superfície para a corrosão



Fonte: Autoria própria.

E, assim podemos encontrar a curva de nível:

A análise de variância verifica se a interação entre os efeitos A:B é significativo. Nesse caso, a interação entre a concentração de CO_2 e a temperatura T é significativa com $p - valor = 0,002813$. Outra observação importante é que o quadrado médio do resíduo é baixo o suficiente quando comparado com os efeitos isoladamente e a interação.

Tabela 9 – Tabela ANOVA para 2^k

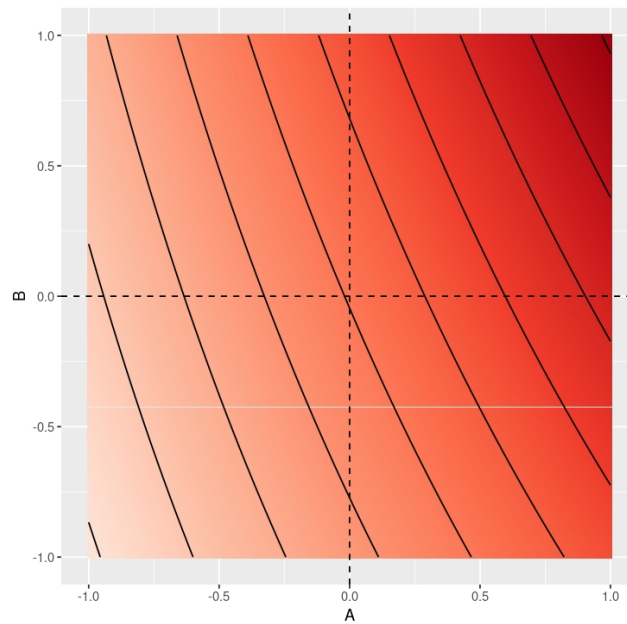
Factor	Df	Sum Sq	Mean Sq	F value	Pr(>F)
A	1	676.00	676.00	772.57	2.907e-12 ***
B	1	121.00	121.00	138.29	6.058e-08 ***
A:B	1	12.25	12.25	14.00	0.002813 **
Residuals	12	10.50	0.87		
—					
Signif. codes:	0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

Fonte: Autoria própria.

A tabela a seguir verifica se os coeficientes da regressora são significativos. Ou seja, avalia se os coeficientes β_0 , β_1 , β_2 e β_{12} são importantes para o modelo. Em todos os casos, a resposta é positiva, indicando que o modelo encontrado foi:

$$y = 36,125 + 6,5x_1 + 2,75x_2 + 0,875x_{12}$$

Figura 9 – Curva de nível para a corrosão



Fonte: Autoria própria.

com ajuste de $R^2 = 93,54\%$.

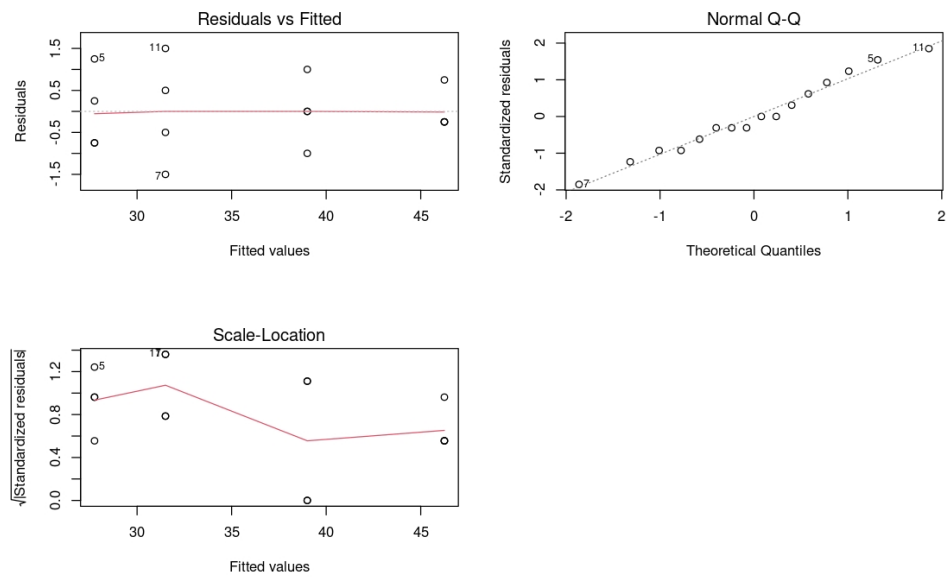
Tabela 10 – Tabela avaliação dos coeficientes β_i

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	36.1250	0.2339	154.477	<2e-16 ***
A	6.5000	0.2339	27.795	2.91e-12 ***
B	2.7500	0.2339	11.759	6.06e-08 ***
A:B	0.8750	0.2339	3.742	0.00281 **
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
Residual standard error: 0.9354 on 12 degrees of freedom				
Multiple R-squared: 0.9872, Adjusted R-squared: 0.984				
F-statistic: 308.3 on 3 and 12 DF,				

Fonte: Autoria própria.

No entanto, o ajuste do modelo será válido se, e somente se, os resíduos forem independentes e apresentarem a homocedasticidade e normalidade dos dados para $\epsilon \sim N(0, \sigma_\epsilon)$. Para testar a normalidade usamos o teste de Shapiro-Wilk para os resíduos. A técnica consiste em verificar se a amostra converge para uma distribuição aproximadamente normal, sendo que a hipótese nula do teste é que a população é normalmente distribuída. O teste avaliou $W = 0,97756$, com $p - \text{valor} = 0,9413$ indicando a normalidade. Para a homocedasticidade podemos avaliar o gráfico seguinte composto por valores ajustados (fitted values) e os resíduos (residuals). Nesse caso, os dados devem ser dispersos o suficiente.

Esses exemplos mostram que o uso das técnicas de regressão múltipla pode ser útil na construção de conceitos avançados da matemática superior como funções em várias dimen-

Figura 10 – Validação dos resíduos**Fonte: Autoria própria.**

sões, equações diferenciais, gradientes, divergentes, algumas teorias que envolvem campos através do uso de regressão múltipla e, dessa forma, apresentar os conceitos aos alunos do ensino técnico que podem melhor interpretar os conceitos em suas próprias áreas.

7 CONCLUSÃO

É importante ressaltar que a Matemática envolvida, dentro das diversas áreas interdisciplinares abordadas e no conteúdo apresentado, é extensa, mas não necessita de conhecimento a mais do que o fornecido durante a formação do professor, compilando conceitos elementares de Cálculo de Várias Variáveis, Álgebra Linear e Geometria.

A Regressão Múltipla se mostrou uma ferramenta abrangente de modelagem matemática, a qual converte dados observáveis em equações com capacidade limitada, mas funcional, de predição. Essa ferramenta se mostra não só útil para o desenvolvimento da ciência como na engenharia e na análise de dados em geral, visto que, desde que os passos de validação sejam respeitados, diversos contextos podem se beneficiar da análise de regressão.

As pontuações feitas neste estudo ressaltam que o avanço da tecnologia tende a demandar uma Matemática mais robusta dos estudantes em formação, bem como dos profissionais já formados que buscam se atualizar. Para que os estudantes tenham pleno aprendizado das técnicas requeridas pelo mercado, o professor também deve ter conhecimento profundo sobre o que está sendo ensinado. Assim, não restritamente à Regressão Múltipla, a preparação do docente deve contemplar não só o algoritmo ensinado em si, mas também a Matemática que cria as técnicas a serem utilizadas. Dessa forma, o conhecimento solidifica o aprendizado da técnica, também removendo a camada de misticismo dos métodos ensinados, de forma que suas limitações fiquem claras, incentivando os estudantes a buscarem alternativas ou métodos complementares quando determinada técnica não é suficiente para resolver os problemas encontrados.

O ferramental apresentado aqui altera a perspectiva da ordem pedagógica dos conteúdos. Convencionalmente se espera que um estudante aprenda matrizes, polinômios, CDI, Estatística, vetores, etc. e seja capaz de aplicar esses conceitos na sua vida profissional e acadêmica. O proposto aqui é que a Regressão Múltipla pode ser a situação desencadeadora de aprendizagem. Dada a técnica e demonstradas as possibilidades de sua aplicação, os conteúdos a serem aprendidos para seu entendimento descolam-se da obrigatoriedade e se aproximam do desafio a ser superado para dominar uma tecnologia de resolução de problemas.

REFERÊNCIAS

- BORGIO, L. A. O cálculo diferencial e integral no ensino profissionalizante, . Trabalho de Conclusão de Curso (Licenciatura em Matemática) - Universidade Tecnológica Federal do Paraná,, 2019.
- BRASIL. **Base Nacional Comum Curricular**. [S.l.]: Ministério da Educação, 2018.
- CANDIDO, A. A literatura e a formação do homem. *In: Textos de intervenção*. São Paulo: Duas Cidades, 2002. p. 67. (Coleção Espírito Crítico).
- CASAGRANDE, R. . **Inteligência Artificial e a educação além da curva**. [S.l.]: Edição do Autor, 2019.
- CAZORLA, C. C. G. I. M. **O papel do letramento estatístico na implementação dos Temas Contemporâneos Transversais da BNCC**. [S.l.: s.n.], 2021. 88–111 p.
- COOK, S. W. R. D. Residuals and influence in regression. **Monographs on Statistics and Applied Probability**, , 1982.
- DIAS, A. A. d. S. CÁlculo diferencial e integral (cdi) no ensino da física na educaÇÃo básica: um estudo. **Revista Inova Ciência amp; Tecnologia / Innovative Science amp; Technology Journal**, v. 5, n. 2, p. 64–73, jun. 2019. Disponível em: <https://periodicos.iftm.edu.br/index.php/inova/article/view/459>.
- DREYFUS ANATOLI KOUROPATOV, G. R. T. Research as a resource in a high-school calculus curriculum. **ZDM Mathematics Education** **53**, , 2021.
- FEDERAL, S. de Estado de E. D. **ENSINO MÉDIO CATÁLOGO DE PERCURSOS EDUCACIONAIS ESTRUTURADOS 2025**. [S.l.]: Seed DF, 2025.
- FILHO, N. de O. B.; LEÃO, W. da S. A matemática É difícil? como os números podem nos enganar. **RECIMA21 - Revista Científica Multidisciplinar - ISSN 2675-6218**, [S. l.], v. 4, n. 6, p. e463295 DOI: 10.47820/recima21.v4i6.3295. Disponível em: <https://recima21.com.br/index.php/recima21/article/view/3295>. Acesso em: 6 maio. 2024, , 2023.
- FORUM, W. E. . **The Future of Jobs**. [S.l.], 2020.
- HAYKIN, S. **Redes neurais: princípios e prática**. [S.l.]: Bookman Editora, 2001.
- HH., P. Analysis of tissue and arterial blood temperatures in the resting human forearm. **J. Appl. Physiol.**, v. 1,, p. 93–122, 1948.
- KUBINIEC, K.; MOFFETT, K. B.; BLOUNT, K. Street-scale urban air temperatures predicted by simple high-resolution cover- and shade-weighted surface temperature mosaics in a variety of residential neighborhoods. **Remote Sensing**, v. 17, n. 11,, 2025. ISSN 2072-4292. Disponível em: <https://www.mdpi.com/2072-4292/17/11/1932>.
- OCDE . **PISA Vocational Education and Training (VET): Assessment and Analytical Framework**, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/b0d5aaf9-en>. [S.l.], 2024.
- OZIGIS, M. S. *et al.* Characterizing crop distribution and the impact on forest conservation in central africa. **Remote Sensing**, v. 17, n. 11,, 2025. ISSN 2072-4292. Disponível em: <https://www.mdpi.com/2072-4292/17/11/1958>.

PAGANI, N. S. G. A. E. M. L. Derivadas e integrais no ensino médio integrado ao técnico: O que pensam professores e alunos. **Anais do Encontro de Produção Discente PUCSP/Cruzeiro do Sul. São Paulo**, , 2014.

ROCHA, J. S. de M. O ensino de cálculo no ensino médio. **Dissertação (mestrado) - Universidade Federal de São João del Rei , Departamento Matemática e Estatística.**, , 2018.

ROMANIELLO, V.; GANCI, G. Exploitation of oco-3 satellite data to analyse carbon dioxide emissions from the mt. etna volcano. **Remote Sensing**, v. 17, n. 11,, 2025. ISSN 2072-4292. Disponível em: <https://www.mdpi.com/2072-4292/17/11/1918>.

SANTANA, G. C. L. D. *et al.* Analfabetismo digital. **CADERNO DISCENTE**, v. 8, n. 1, p. 65–71, 2023.

SILVA, T. S. d.; PASQUALLI, R.; SPESSATTO, M. B. Desafios da implementação do novo ensino médio: o que dizem os professores. **Educação Em Foco**, **28(1)**, e28007, , Recuperado de <https://periodicos.ufjf.br/index.php/edufoco/article/view/39210>, 2023.

SPINA, C. d. O. C. Modelagem matemática no processo ensino-aprendizagem do cálculo diferencial e integral para o ensino médio, . Universidade Estadual Paulista (Unesp),, 2002.

STATISTA, . 2020. Acessado em 06 de junho de 2025. Disponível em: <https://www.statista.com/statistics/871513/worldwide-data-created/>.

VOZINIAK, P. *et al.* Avaliação do gradiente térmico na superfície de pele normal e com lesão por pressão. *In: . 2023. Anais [...]* [S.l.: s.n.], 2023.

WU, e. a. X. Big data and artificial intelligence in cancer research. **Trends in cancer**, v. 10,2,, 2024.

ANEXO A – Lei N.º 9.610, de 19 de Fevereiro de 1998



Presidência da República
Casa Civil
Subchefia para Assuntos Jurídicos

LEI Nº 9.610, DE 19 DE FEVEREIRO DE 1998¹.

Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências.

O PRESIDENTE DA REPÚBLICA Faço saber que o Congresso Nacional decreta e eu sanciono a seguinte Lei:

Título I - Disposições Preliminares

Art. 1º Esta Lei regula os direitos autorais, entendendo-se sob esta denominação os direitos de autor e os que lhes são conexos.

Art. 2º Os estrangeiros domiciliados no exterior gozarão da proteção assegurada nos acordos, convenções e tratados em vigor no Brasil.

Parágrafo único. Aplica-se o disposto nesta Lei aos nacionais ou pessoas domiciliadas em país que assegure aos brasileiros ou pessoas domiciliadas no Brasil a reciprocidade na proteção aos direitos autorais ou equivalentes.

Art. 3º Os direitos autorais reputam-se, para os efeitos legais, bens móveis.

Art. 4º Interpretam-se restritivamente os negócios jurídicos sobre os direitos autorais.

Art. 5º Para os efeitos desta Lei, considera-se:

I - publicação - o oferecimento de obra literária, artística ou científica ao conhecimento do público, com o consentimento do autor, ou de qualquer outro titular de direito de autor, por qualquer forma ou processo;

II - transmissão ou emissão - a difusão de sons ou de sons e imagens, por meio de ondas radioelétricas; sinais de satélite; fio, cabo ou outro condutor; meios óticos ou qualquer outro processo eletromagnético;

III - retransmissão - a emissão simultânea da transmissão de uma empresa por outra;

IV - distribuição - a colocação à disposição do público do original ou cópia de obras literárias, artísticas ou científicas, interpretações ou execuções fixadas e fonogramas, mediante a venda, locação ou qualquer outra forma de transferência de propriedade ou posse;

V - comunicação ao público - ato mediante o qual a obra é colocada ao alcance do público, por qualquer meio ou procedimento e que não consista na distribuição de exemplares;

VI - reprodução - a cópia de um ou vários exemplares de uma obra literária, artística ou científica ou de um fonograma, de qualquer forma tangível, incluindo qualquer armazenamento permanente ou temporário por meios eletrônicos ou qualquer outro meio de fixação que venha a ser desenvolvido;

VII - contrafação - a reprodução não autorizada;

VIII - obra:

a) em co-autoria - quando é criada em comum, por dois ou mais autores;

b) anônima - quando não se indica o nome do autor, por sua vontade ou por ser desconhecido;

c) pseudônima - quando o autor se oculta sob nome suposto;

d) inédita - a que não haja sido objeto de publicação;

e) póstuma - a que se publique após a morte do autor;

f) originária - a criação primígena;

g) derivada - a que, constituindo criação intelectual nova, resulta da transformação de obra originária;

¹ Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l9610.htm.

h) coletiva - a criada por iniciativa, organização e responsabilidade de uma pessoa física ou jurídica, que a publica sob seu nome ou marca e que é constituída pela participação de diferentes autores, cujas contribuições se fundem numa criação autônoma;

i) audiovisual - a que resulta da fixação de imagens com ou sem som, que tenha a finalidade de criar, por meio de sua reprodução, a impressão de movimento, independentemente dos processos de sua captação, do suporte usado inicial ou posteriormente para fixá-lo, bem como dos meios utilizados para sua veiculação;

IX - fonograma - toda fixação de sons de uma execução ou interpretação ou de outros sons, ou de uma representação de sons que não seja uma fixação incluída em uma obra audiovisual;

X - editor - a pessoa física ou jurídica à qual se atribui o direito exclusivo de reprodução da obra e o dever de divulgá-la, nos limites previstos no contrato de edição;

XI - produtor - a pessoa física ou jurídica que toma a iniciativa e tem a responsabilidade econômica da primeira fixação do fonograma ou da obra audiovisual, qualquer que seja a natureza do suporte utilizado;

XII - radiodifusão - a transmissão sem fio, inclusive por satélites, de sons ou imagens e sons ou das representações desses, para recepção ao público e a transmissão de sinais codificados, quando os meios de decodificação sejam oferecidos ao público pelo organismo de radiodifusão ou com seu consentimento;

XIII - artistas intérpretes ou executantes - todos os atores, cantores, músicos, bailarinos ou outras pessoas que representem um papel, cantem, recitem, declamem, interpretem ou executem em qualquer forma obras literárias ou artísticas ou expressões do folclore.

Art. 6º Não serão de domínio da União, dos Estados, do Distrito Federal ou dos Municípios as obras por eles simplesmente subvencionadas.