

MÉTODOS DE SEGUNDA ORDEM PARA PROBLEMAS COMPÓSITOS DESCONTÍNUOS

Relatório Científico Primeiro Relatório Científico Anual do projeto na
modalidade Bolsa no País - Doutorado, fomentado pela Fundação de Amparo à
Pesquisa do Estado de São Paulo.

Projeto FAPESP #2024/20168-8

Pesquisador Responsável: Paulo José da Silva e Silva
Doutorando: Gabriel Belém Barbosa

Campinas, 7 de fevereiro de 2026

Informações Gerais do Projeto

- Título do projeto:

Métodos de segunda ordem para problemas compósitos descontínuos

- Nome do pesquisador responsável:

Paulo José da Silva e Silva

Doutorando: Gabriel Belém Barbosa

- Instituição sede do projeto:

**Departamento de Matemática Aplicada, Instituto de Matemática, Estatística e
Computação Científica da Universidade Estadual de Campinas**

- Equipe de pesquisa:

Paulo José da Silva e Silva

Gabriel Belém Barbosa

- Número do projeto de pesquisa:

2024/20168-8

- Período de vigência:

01/03/2025 a 28/02/2029

- Período coberto por este relatório científico:

01/03/2025 a 28/02/2026

Resumo

Este projeto visa investigar e avançar no campo da otimização compósita não convexa, focando na recente incorporação de informação de segunda ordem nos modelos iterativos. Duas referências recentes se destacam nessa área: *Forward-backward envelope for the sum of two nonconvex functions: further properties and nonmonotone line-search algorithms* [Themeles 2018] e *The Indefinite Proximal Gradient Method* [Leconte 2024]. Vários problemas do mundo real, incluindo recuperação de sinal esparso, processamento de imagem e otimização de portfólio, podem ser modelados na forma do problema de interesse. Esse trabalho também envolverá extensa experimentação numérica em várias classes de problemas para avaliar o desempenho dos algoritmos propostos. Os experimentos se concentrarão em comparar o desempenho de diferentes estratégias de busca linear sob diferentes hipóteses, especialmente estratégias não monótonas, o impacto da incorporação de informações quasi-Newton nos modelos do subproblema e a eficácia de cada atualização. Possivelmente novas estratégias híbridas também serão analisadas. O projeto pretende avançar no entendimento de métodos como o ZeroFPR [Themeles 2018] e de gradiente proximal indefinido [Leconte 2024] no tratamento de problemas de otimização compósita não convexa, fornecer orientação prática na seleção de estratégias de busca linear adequadas e atualizações quasi-Newton para diferentes estruturas de problema e desenvolver algoritmos eficientes e robustos para resolver uma classe mais ampla de problemas de otimização compósita não convexa, levando potencialmente a soluções aprimoradas em importantes áreas de aplicação.

Abstract

This project aims to investigate and advance the field of nonconvex composite optimization, focusing on the recent incorporation of second-order information into iterative models. Two recent sources stand out in this area: *Forward-backward envelope for the sum of two nonconvex functions: further properties and nonmonotone line-search algorithms* [Themeles 2018] and *The Indefinite Proximal Gradient Method* [Leconte 2024]. Several real-world problems, including sparse signal recovery, image processing, and portfolio optimization, can be modeled in the form of the problem of interest. This work will also involve extensive numerical experimentation on several types of problems to evaluate the performance of the proposed algorithms. The experiments will focus on comparing the performance of different line search strategies with or without relaxed assumptions, especially nonmonotone variants, the impact of incorporating quasi-Newton information into the subproblem and the effectiveness of each update. New hybrid strategies also possibly will be analysed. The project aims to contribute to advance the understanding of methods such as ZeroFPR [Themeles 2018] and indefinite proximal gradient [Leconte 2024] in the treatment of non-convex composite optimization problems, to provide practical guidance in the selection of appropriate line search strategies and quasi-Newton updates for different problem structures, and to develop efficient and robust algorithms to solve a broader class of non-convex composite optimization problems, potentially leading to improved solutions in important application areas.

Sumário

Informações Gerais do Projeto	i
Resumo	ii
Abstract	iii
1 Realizações no período	1
1.1 Atividades Acadêmicas	1
1.2 Atividades de Pesquisa	2
1.2.1 Definição do Problema e Notação	2
1.2.2 Revisão Bibliográfica	2
1.3 Desenvolvimento Algorítmico e Experimental	2
1.3.1 Estratégias Avançadas de Validação Cruzada	3
1.3.2 Integração Híbrida e Escala Dinâmica	3
1.3.3 Resultados Preliminares	4
2 Descrição e avaliação do apoio institucional recebido no período	6
3 Participação em evento científico	7
4 Plano de atividades para o próximo período	8
Referências Bibliográficas	8

1 Realizações no período

Durante a vigência deste primeiro período do projeto de doutorado (03/2025 a 02/2026), as atividades planejadas foram cumpridas, com destaque para a conclusão dos créditos obrigatórios, aprovação em exames de qualificação e participação em evento científico, além do andamento da pesquisa bibliográfica e experimental.

1.1 ATIVIDADES ACADÊMICAS

No âmbito das exigências do programa de Doutorado em Matemática Aplicada do IMECC/Unicamp, foram realizadas as seguintes atividades:

- Coeficiente de Rendimento (CR): O aluno mantém um CR perfeito de 4,0000;
- Disciplinas Cursadas no Doutorado (1º Semestre de 2025):
 - MT504 — Fluxos em Redes: Aprovado com conceito A;
 - MT853 — Tópicos em Otimização: Aprovado com conceito A;
- Aproveitamento de Créditos: As disciplinas obrigatórias MT401 — Análise Aplicada e MT402 — Matrizes foram aproveitadas do Mestrado (registradas como MT801 e MT802 no 2º semestre de 2025), cumprindo os requisitos de créditos obrigatórios;
- Total de Créditos: Foram totalizados 18 créditos no Doutorado (incluindo aproveitamentos), avançando significativamente rumo à integralização dos créditos exigidos pelo programa;
- Estágio Docente: Realização do Estágio de Capacitação Docente (PED C) na disciplina MS211 - Cálculo Numérico (código CD003), sob supervisão, durante o 2º semestre de 2025;
- Exame de Proficiência em inglês: Aprovação no Exame de Proficiência em Inglês II (Escrito), realizado em 30/05/2025;
- Exames de Qualificação:
 - Aprovação no Exame de Qualificação na área de Análise Aplicada (MT401), realizado em 25/08/2025;
 - Aprovação no Exame de Qualificação na área de Matrizes (MT402), realizado em 27/08/2025.

1.2 ATIVIDADES DE PESQUISA

As atividades de pesquisa concentraram-se na investigação e aprimoramento de métodos de segunda ordem para otimização compósita não convexa.

1.2.1 Definição do Problema e Notação

Considere o problema de regressão linear com ruído:

$$y = Ax + \varepsilon, \quad (\text{P})$$

onde $y \in \mathbb{R}^n$ é o vetor resposta e $A \in \mathbb{R}^{n \times p}$ é a matriz de desenho. O objetivo é recuperar o vetor de coeficientes $x \in \mathbb{R}^p$, assumindo que a solução verdadeira é esparsa (poucos coeficientes não nulos), uma hipótese comum e desejável em cenários onde $p \gg n$.

Para impor essa esparsidade, abordamos o problema de regularização ℓ_0 :

$$\min_{x \in \mathbb{R}^p} \frac{1}{2} \|Ax - y\|_2^2 + \lambda \|x\|_0, \quad (\text{R})$$

onde $\|x\|_0$ denota a pseudo-norma ℓ_0 (cardinalidade do suporte de x) e $\lambda > 0$ é o parâmetro de regularização que controla o balanço entre o erro de ajuste e a esparsidade da solução.

Durante o texto, utilizamos a seguinte notação:

- $[p] := \{1, 2, \dots, p\}$;
- O suporte do vetor x é denotado por $\text{supp}(x) = \{i \in [p] : x_i \neq 0\}$.

O problema (R) é NP-difícil [3], o que motiva o desenvolvimento de algoritmos que combinem a eficiência de métodos contínuos com a qualidade de soluções de métodos combinatoriais.

Os principais avanços são descritos nas seções a seguir.

1.2.2 Revisão Bibliográfica

Estudo aprofundado de referências centrais e do estado da arte:

- *Sparse regression at scale: branch-and-bound rooted in first-order optimization* [4];
- *First-Order Methods in Optimization* [2];
- *Fast Best Subset Selection* [3], para comparação com métodos de seleção de subconjuntos.

1.3 DESENVOLVIMENTO ALGORÍTMICO E EXPERIMENTAL

Foram realizados vários experimentos computacionais no ambiente Julia, bem como desenvolvimentos teóricos, focados na robustez e eficiência dos algoritmos para o problema de seleção de variáveis. As principais contribuições são detalhadas a seguir.

1.3.1 Estratégias Avançadas de Validação Cruzada

Implementamos e comparamos novas estratégias de Validação Cruzada (Cross-Validation - CV) para a seleção do hiperparâmetro λ ao longo do caminho de regularização. Essas estratégias foram testadas no ambiente computacional Julia. A escolha adequada de λ é fundamental para o desempenho do modelo, pois controla o equilíbrio entre o ajuste aos dados e a esparsidade da solução. Em cenários reais, onde o suporte verdadeiro é desconhecido, o erro de validação (erro de predição em um conjunto de dados não utilizado no treinamento) serve como critério para identificar o modelo que melhor generaliza.

Nosso objetivo com essas estratégias é mitigar a convergência para mínimos locais ruins. Essa estratégia é particularmente essencial em contextos nos quais a esparsidade da solução original não é conhecida, como em problemas de seleção de variáveis em aprendizado de máquina.

A abordagem padrão na literatura (e em pacotes como L0Learn e GLMNet) é percorrer o caminho de regularização de λ_{high} (solução esparsa) para λ_{low} (solução densa), utilizando a solução anterior como ponto de partida (*warmup*). No entanto, em problemas não convexos como a regularização ℓ_0 , essa estratégia gananciosa pode ficar presa precocemente em mínimos locais subótimos.

A primeira abordagem desenvolvida, denominada *CV Inverso*, inverte a lógica padrão. Esta estratégia percorre o caminho de λ_{low} a λ_{high} , iniciando de soluções mais densas. Diferente da "partida fria" (iniciar do zero), implementamos um *warmup* específico: partindo de um λ elevado, reduzimos seu valor sucessivamente (por um fator de 0,9) até que o algoritmo encontre uma solução com suporte não-nulo ($\|x\|_0 > 0$). Esse valor é então utilizado como ponto de partida para a varredura ascendente. A intuição é que, ao permitir que mais variáveis entrem no modelo inicialmente, o algoritmo pode explorar melhor o espaço de busca combinatorial antes de ser forçado a esparsificar a solução, evitando mínimos locais associados a suportes incorretos.

Além disso, desenvolvemos o *CV Adaptativo Inteligente*, uma abordagem híbrida que sonda os erros de validação nos dois extremos do intervalo (λ_{low} e λ_{high}) e seleciona a direção de varredura que apresenta o menor erro de validação inicial. O algoritmo permite ainda uma única reversão de direção caso a varredura na direção escolhida não produza uma redução significativa no erro de validação (pelo menos 1%) nas primeiras etapas, permitindo explorar vales de mínimos locais que poderiam ser acessíveis apenas pelo outro extremo. Embora exista o risco teórico de não visitar o ótimo global se a superfície do erro de validação for altamente multimodal, essa heurística demonstrou, empiricamente, um balanço favorável: obteve resultados de qualidade competitiva com o *Grid Search* exaustivo, porém com custo computacional reduzido pela metade, ao evitar o cômputo de caminhos de regularização pouco promissores.

1.3.2 Integração Híbrida e Escala Dinâmica

Recentemente, desenvolvemos variantes do método do Gradiente Proximal Espectral (como o NSPG [1] e VMNSPG [5]) projetadas para lidar com a norma ℓ_0 . Visto que esses métodos apresentam convergência acelerada em comparação a abordagens de primeira ordem — sem o custo computacional de informações de segunda ordem — decidimos investigar sua eficácia em problemas práticos de seleção de variáveis. Para isso, integramos esses métodos com a estratégia

de busca proposta. No entanto, sua aplicação direta impõe desafios: para algoritmos baseados em passos espectrais, corre-se o risco de instabilidade na seleção do suporte devido à variação da escala do passo.

No contexto espectral, o passo γ_k (uma aproximação escalar para a inversa da Hessiana) varia dinamicamente a cada iteração para capturar a curvatura local da função objetivo. O operador proximal para a regularização ℓ_0 impõe esparsidade zerando coeficientes menores que um limiar $\tau = \sqrt{2\gamma_k\lambda}$. Portanto, a agressividade do corte depende não somente do hiperparâmetro λ , mas também do passo espectral γ_k . Como γ_k oscila, um valor fixo de λ resultaria em um limiar de corte instável, fazendo com que variáveis entrassem e saíssem do suporte erradicamente. Para mitigar isso, propomos compensar essa variação: ajustamos o λ efetivo utilizando informação do γ_k , visando manter o limiar de decisão estável ao longo das iterações.

Essa técnica é fundamental para a estratégia híbrida desenvolvida, que combina as forças de diferentes classes de algoritmos em dois estágios. Na fase global, utilizamos o NSPG para escapar de mínimos locais rasos e identificar rapidamente um suporte razoável. Na fase seguinte, estratégias de *Coordinate Partial Swap* (CPSI) [3] garantem otimalidade combinatorial de ordem superior. Além disso, testamos o uso das soluções do NSPG como ponto de partida para o *Partially Greedy Cyclic Coordinate Descent* (PGCCD) [3], gerando um método híbrido. O PGCCD converge rapidamente para um mínimo de alta precisão, refinando os coeficientes no suporte identificado, antes da busca por otimalidade combinatorial.

1.3.3 Resultados Preliminares

Para validar a eficácia dos métodos, utilizamos dados sintéticos gerados segundo o modelo linear (P) conforme descrito em [3]. A matriz de dados A possui colunas normalizadas e estrutura de correlação controlada pelo parâmetro ρ (correlação entre variáveis adjacentes com decaimento exponencial ou correlação constante). O nível de ruído é determinado pelo *Signal to Noise Ratio* (SNR), que varia conforme a dificuldade do cenário analisado.

A qualidade da solução recuperada é medida pela métrica de Similaridade do Suporte (Jaccard Modificado), definida como $J(S, S^\dagger) = \frac{|S \cap S^\dagger|}{\max\{|S|, |S^\dagger|\}}$, onde S é o suporte estimado pelo algoritmo e S^\dagger é o suporte verdadeiro (com $|S^\dagger| = k^\dagger$). Esta métrica varia de 0 (disjunção total) a 1 (recuperação perfeita).

Além do cenário exponencial, avaliamos o desempenho em configurações de correlação constante, que impõem desafios diferentes à estrutura de correlação das variáveis. A Figura 1.6 apresenta os resultados para $\rho = 0.9$ e $p = 1000$.

Os resultados ilustrados nas Figuras 1.3 e 1.6 indicam a robustez da combinação proposta. Em cenários desafiadores de ruído considerável e alta correlação entre os parâmetros — seja ela exponencial ou constante — a abordagem híbrida supera ou iguala o estado da arte (LOLearn) em termos de qualidade da solução recuperada.

Particularmente no cenário de correlação exponencial (Fig. 1.3), cabe ressaltar que, embora o híbrido NSPG+PGCCD possa ser mais lento para valores superiores de n , ele atinge recuperação perfeita em regimes de n pequeno, onde seu tempo de execução permanece competitivo, oferecendo uma alternativa valiosa para cenários de dados limitados.

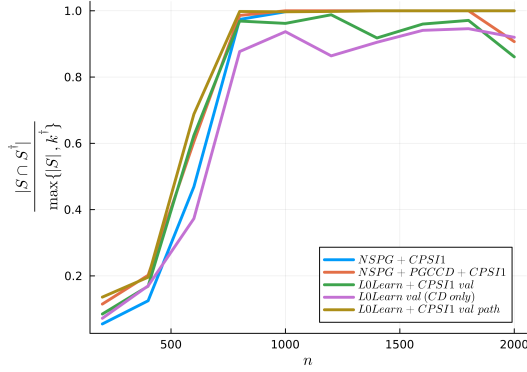


Figura 1.1: Similaridade do Suporte (Jaccard Modificado).

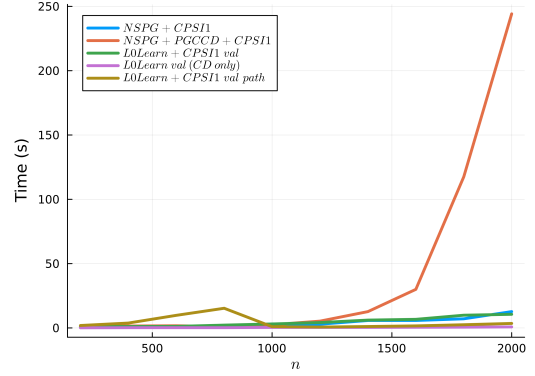


Figura 1.2: Tempo de Execução (s).

Figura 1.3: Comparação entre o método proposto (PGCCD/NSPG) e o estado da arte (variantes do L0Learn) em cenário de correlação exponencial ($\rho = 0.5$, $p = 2000$, $n = 500$, $k^\dagger = 100$, SNR = 10). Os métodos propostos atingem recuperação de suporte superior ou equivalente com tempos competitivos.

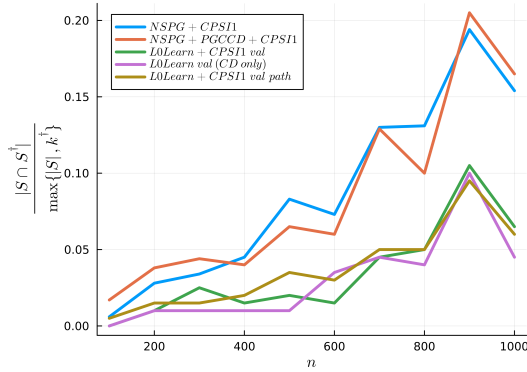


Figura 1.4: Similaridade do Suporte (Jaccard Modificado).

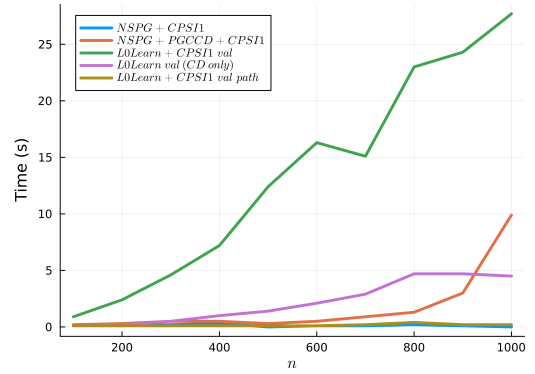


Figura 1.5: Tempo de Execução (s).

Figura 1.6: Desempenho em cenário de correlação constante ($\rho = 0.9$, $p = 1000$, $n = 250$, $k^\dagger = 20$, SNR = 5). Neste regime de alta correlação e ruído mais acentuado, os métodos propostos demonstram superioridade marcante na qualidade da solução recuperada em comparação com as alternativas, mantendo a eficiência computacional.

2 Descrição e avaliação do apoio institucional recebido no período

O Departamento de Matemática Aplicada do Instituto de Matemática, Estatística e Computação Científica (IMECC) contribuiu com o projeto através do acesso ao LOC (Laboratório de Otimização Contínua) e a seus computadores de alto desempenho, essenciais para os experimentos numéricos realizados. Além disso, o SBU (Sistema de Bibliotecas da Unicamp) forneceu acesso ao acervo bibliográfico e aos periódicos assinados pela UNICAMP e CAPES, fundamentais para a revisão bibliográfica.

3 Participação em evento científico

O aluno participou do 1st Carioca Workshop on Optimization and Applications (CariOPT 2025), realizado na Escola de Matemática Aplicada da Fundação Getúlio Vargas (FGV EMAp), no Rio de Janeiro, de 05 a 07 de maio de 2025.

Na ocasião, foi apresentado o trabalho intitulado “*Optimization of problems involving group sparsity*”, divulgando os resultados preliminares e a temática central da pesquisa.

4 Plano de atividades para o próximo período

Para o próximo período de vigência (03/2026 a 02/2027), o projeto prevê as seguintes etapas principais:

1. Atividades Acadêmicas: Cumprimento dos créditos restantes e preparação para o *Exame de Qualificação Específico* na disciplina MT601 - Métodos Computacionais de Otimização, conforme exigência do programa de pós-graduação para a linha de pesquisa;
2. Redação e Submissão de Artigo Científico: Organização dos resultados obtidos, em particular a comparação entre o método proposto (NSPG) e o estado da arte (L0Learn), bem como a análise das estratégias híbridas, visando a submissão de um artigo para um periódico de circulação internacional na área de otimização;
3. Aprofundamento dos Experimentos Numéricos: Expandir a bateria de testes para incluir:
 - Problemas de maior escala para avaliar a escalabilidade das implementações;
 - Aplicações com dados reais (e.g., genômica ou processamento de imagens) para validar a eficácia prática da regularização ℓ_0 nos modelos propostos;
 - Testes de robustez em cenários mal-condicionados e com ruído elevado;
4. Participação em Evento Internacional: Participação confirmada em um Minisimpósio na *SIAM Conference on Optimization (OP26)*, que ocorrerá de 02 a 05 de junho de 2026, na University of Edinburgh, Edimburgo, Reino Unido;
5. Estágio de Pesquisa no Exterior (BEPE): Preparação e submissão de proposta para realização de Estágio de Pesquisa no Exterior (BEPE) na *Polytechnique Montréal*, Canadá, com início previsto para setembro de 2026. O projeto será supervisionado pelo Prof. Dr. Dominique Orban, Professor Titular do Departamento de Matemática e Engenharia Industrial e pesquisador afiliado ao GERAD (Group for Research in Decision Analysis) e IVADO (Institute for Data Valorization), referência internacional em otimização numérica e computacional. As tratativas preliminares com o pesquisador já estão em andamento;
6. Investigação Teórica: Estudar a convergência global das estratégias híbridas desenvolvidas, buscando estabelecer garantias teóricas para a combinação de métodos de descida coordenada com inicializações espectrais não-monótonas;

Referências Bibliográficas

- [1] G. B. Barbosa and P. J. d. S. Silva. Relaxing smoothness conditions for non-monotone optimization methods applied to discontinuous composite problems. Manuscript in progress, 2024.
- [2] A. Beck. *First-Order Methods in Optimization*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2017.
- [3] H. Hazimeh and R. Mazumder. Fast best subset selection: Coordinate descent and local combinatorial optimization algorithms. *Operations Research*, 68(5):1517–1537, 2020.
- [4] H. Hazimeh, R. Mazumder, and A. Saab. Sparse regression at scale: branch-and-bound rooted in first-order optimization. *Mathematical Programming*, 196:1–42, 10 2021.
- [5] Y. Park, S. Dhar, S. Boyd, and M. Shah. Variable metric proximal gradient method with diagonal barzilai-borwein stepsize. In *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3597–3601, 05 2020.