



30 ANOS DE STOCHASTIC DUAL DYNAMIC PROGRAMMING (SDDP)

Mario Veiga
mario@psr-inc.com

Temário

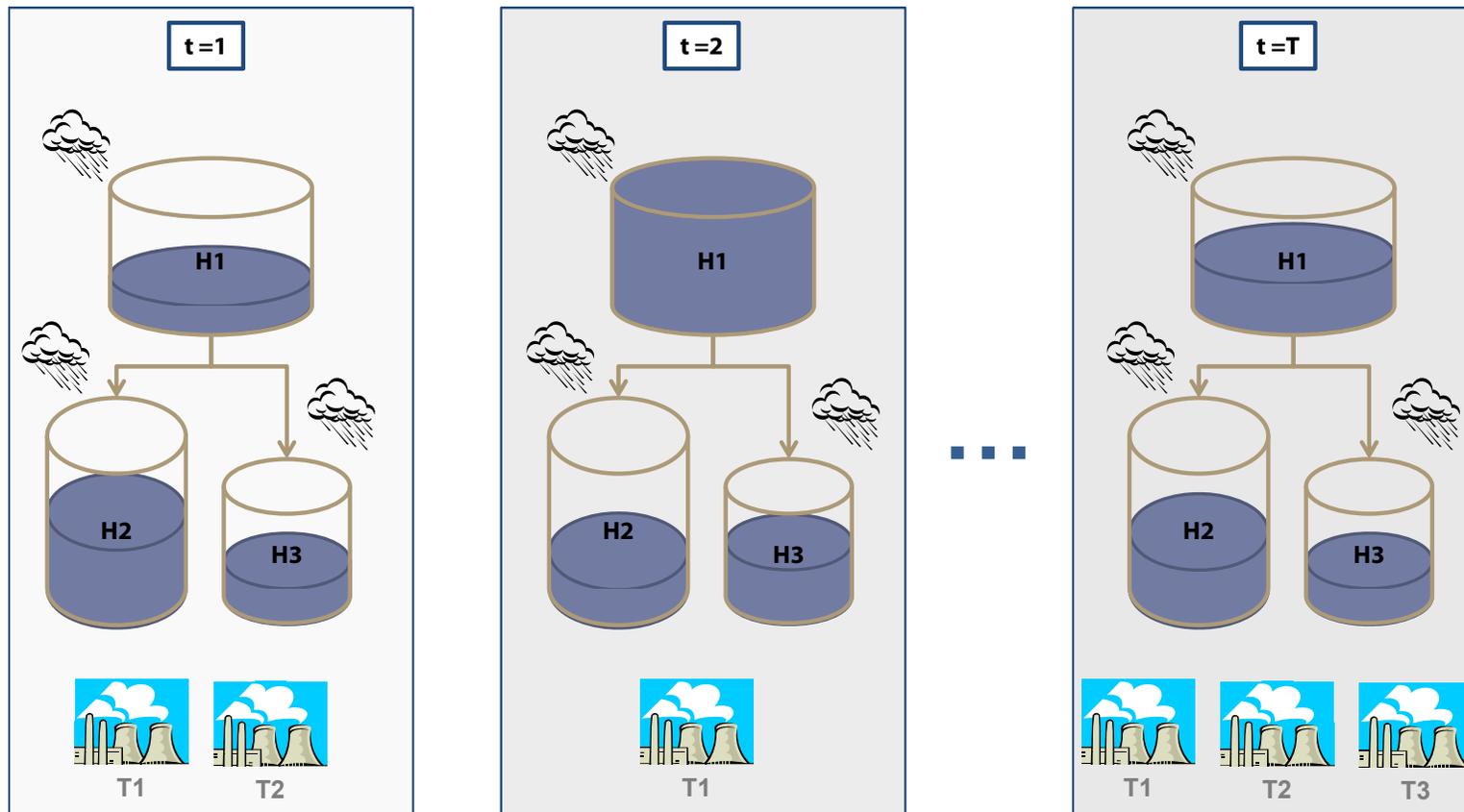
- ▶ **Motivação: operação de sistemas elétricos com reservatórios**
- ▶ **O algoritmo SDDP**
- ▶ **Aplicações do SDDP**
- ▶ **Extensão para aversão ao risco**
- ▶ **Extensão para não convexidade**
- ▶ **SDDP e Reinforcement Learning**

Temário

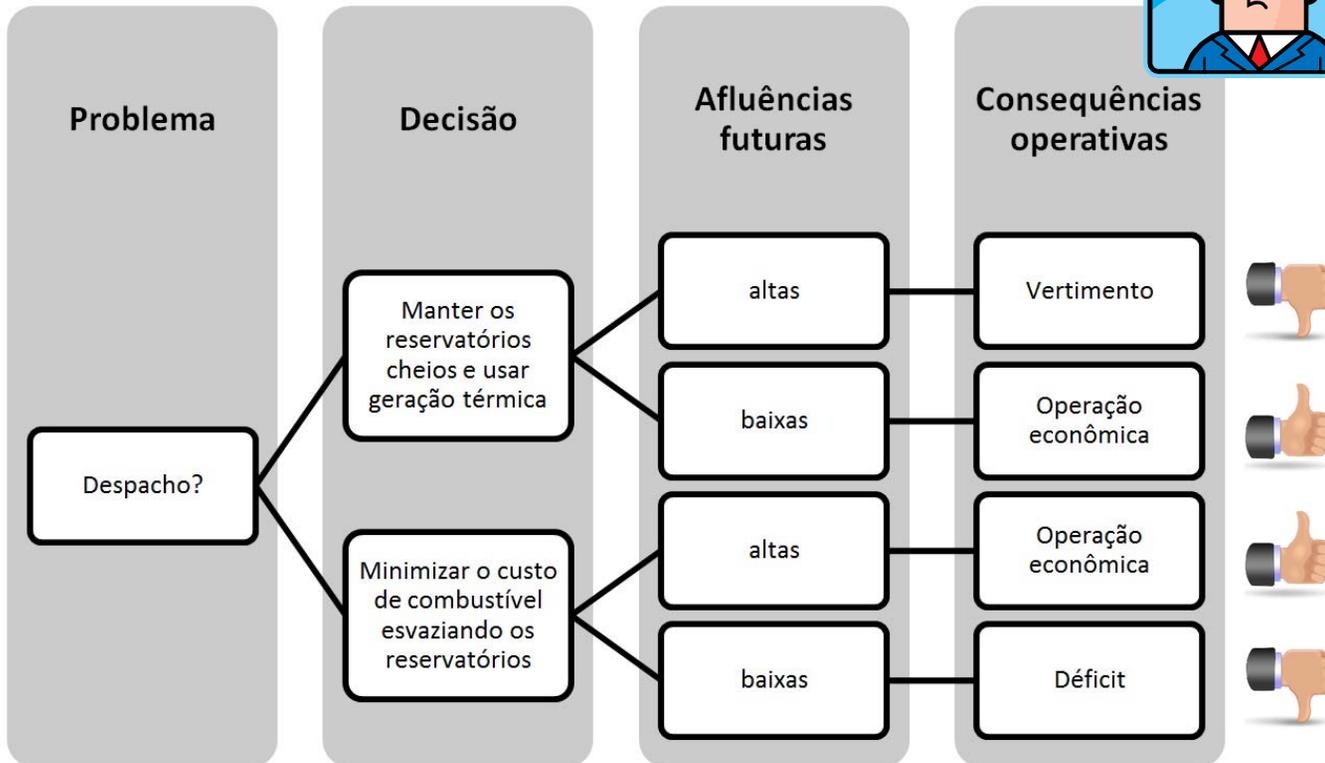
- ▶ **Motivação: operação de sistemas elétricos com reservatórios**
- ▶ O algoritmo SDDP
- ▶ Aplicações do SDDP
- ▶ Extensão para aversão ao risco
- ▶ Extensão para não convexidade
- ▶ SDDP e Reinforcement Learning

Operação de sistemas elétricos

Problema de otimização *multi-estágio e estocástico* (incertezas sobre afluências, demanda etc.)

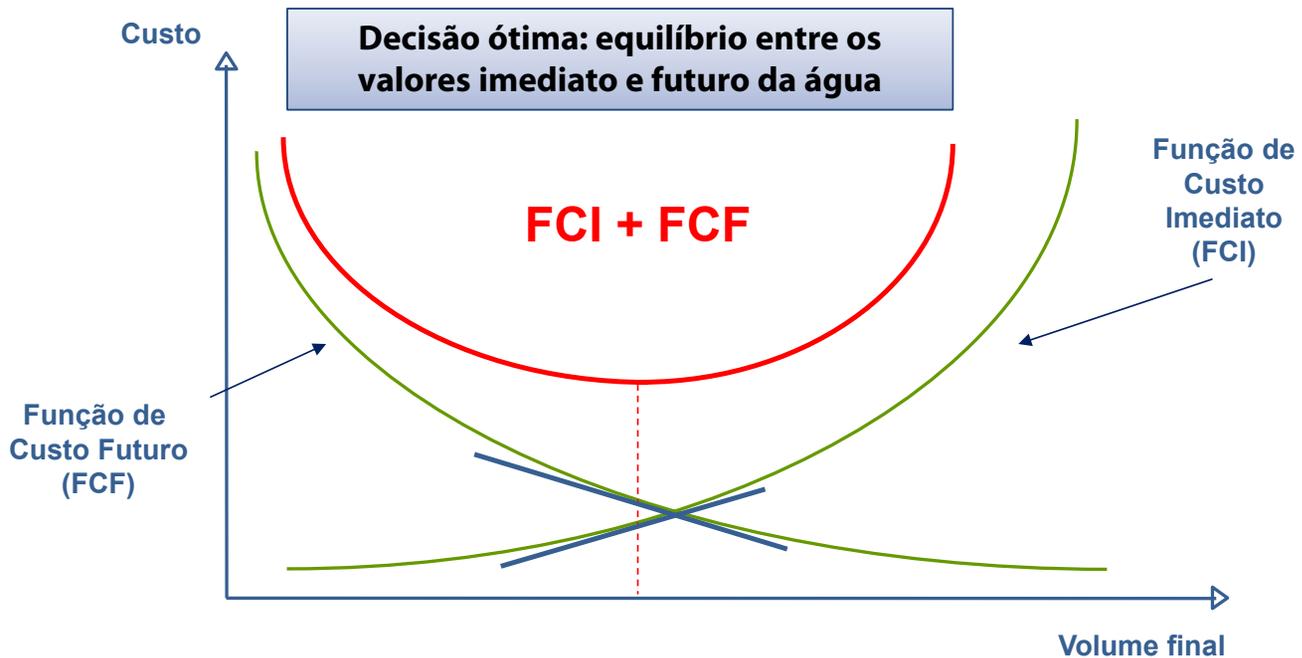


É possível resolver com árvore de decisões?



NÃO!
o número de ramos da árvore cresce exponencialmente com o número de decisões e estágios. Por exemplo, 10 decisões ao longo de 24 meses (2 anos) $\Rightarrow 10^{24} \approx (10^6)^4 = 1 \text{ milhão}^4$

É possível resolver com Programação Dinâmica Estocástica (SDP)?



MALDIÇÃO DA DIMENSIONALIDADE

A SDP requer a discretização dos estados (armazenamento):

1 reservatório $\Rightarrow 10^1$ estados

5 reservatórios $\Rightarrow 10^5$ estados

Sistemas reais podem ter 100 reservatórios (!)

Temário

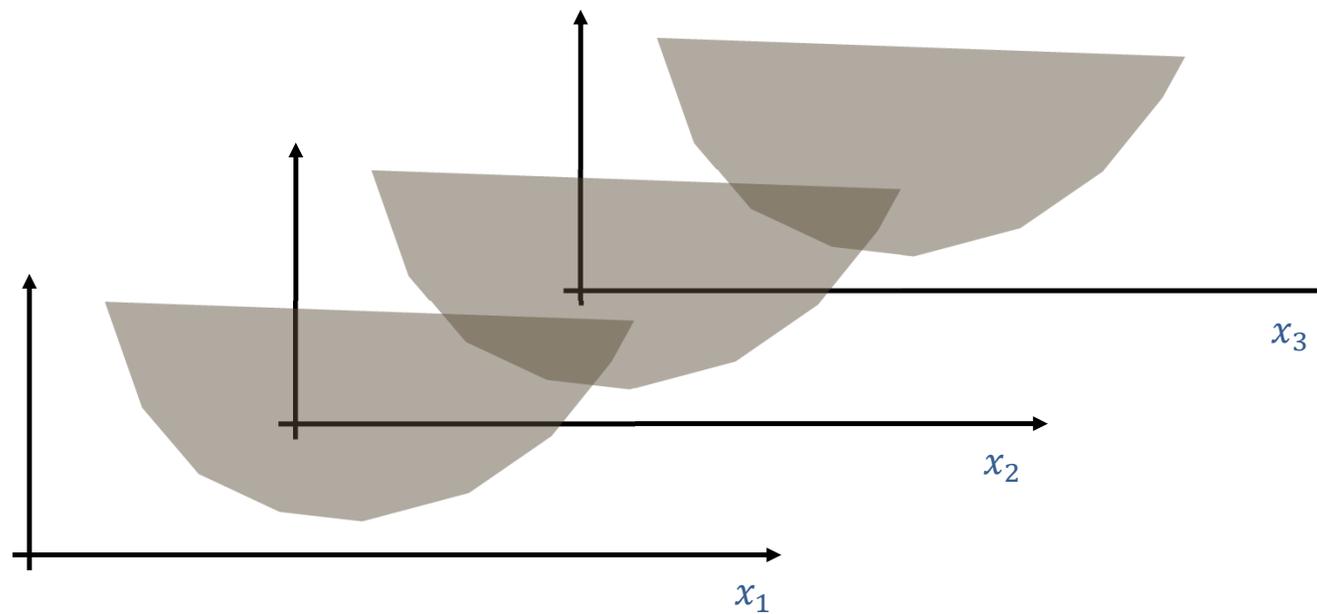
- ▶ **Motivação: operação de sistemas elétricos com reservatórios**
- ▶ **O algoritmo SDDP**
- ▶ **Aplicações do SDDP**
- ▶ **Extensão para aversão ao risco**
- ▶ **Extensão para não convexidade**
- ▶ **SDDP e Reinforcement Learning**

O algoritmo SDDP (com Leontina Pinto e Nora Campodónico)

Ambas do PESC!

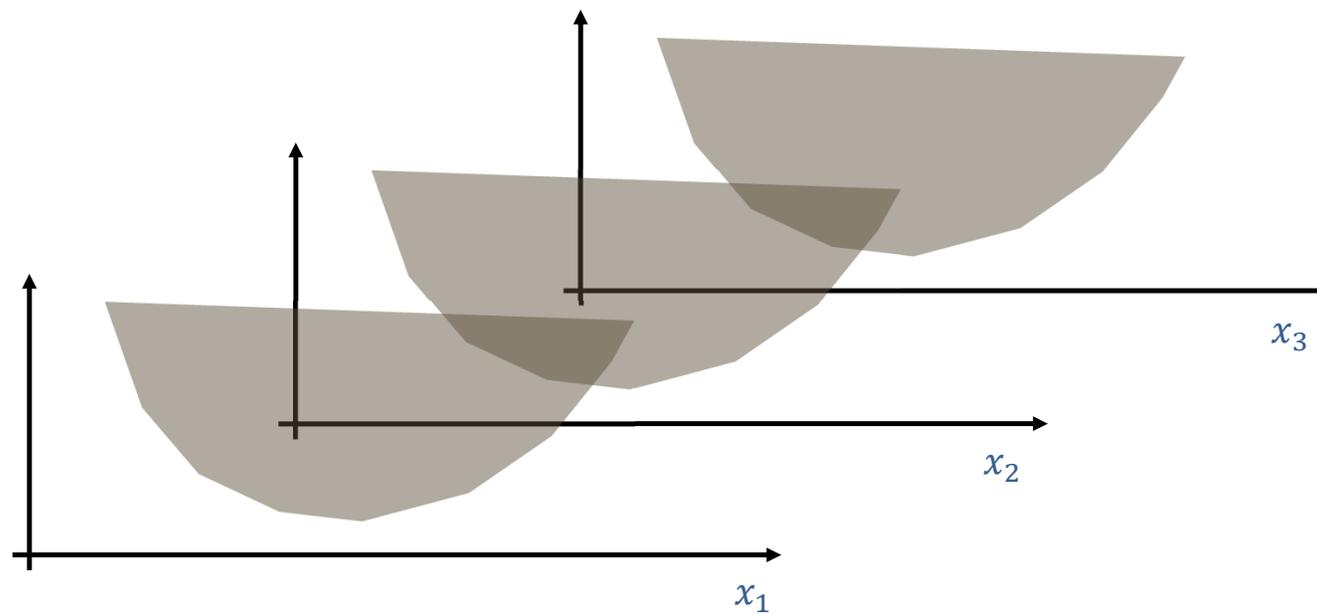
- ▶ Alivia a maldição da dimensionalidade usando hiperplanos para representar a FCF
 - Não permite precisa discretizar, permite extrapolar ao invés de interpolar
 - Estes hiperplanos são obtidos das variáveis duais do problema operativo
 - São um “lower bound” da FCF
- ▶ Concentra os hiperplanos na região “plausível” do espaço de estados
 - Por exemplo, é implausível que 50% dos reservatorios estejam cheios e 50% vazios
- ▶ Como obter estados plausíveis? Simulando a operação do sistema (**passo forward**)
 - A simulação forward também produz um upper bound
- ▶ **Passo backward**: recursão de SDP para os estados identificados na “forward”
 - Os novos hiperplanos melhoram a representação da FCF \Rightarrow novo lower bound

Função de custo futuro (FCF)

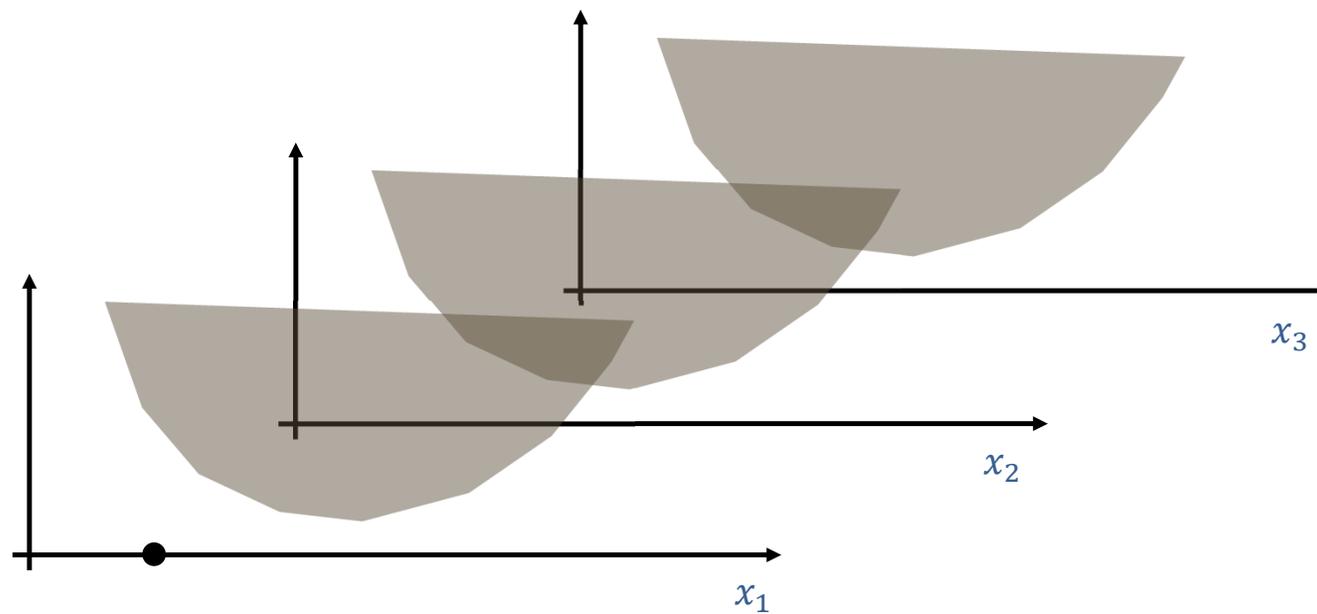


Baseado em slides do Shabbir Ahmed

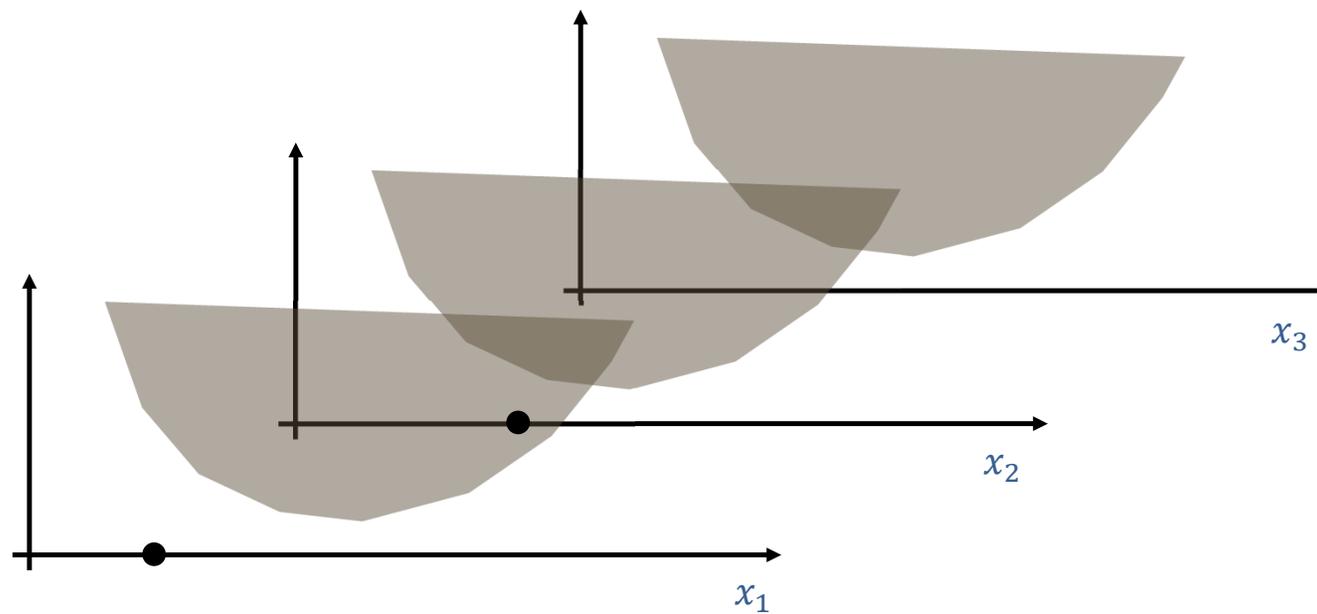
Forward iteração 1



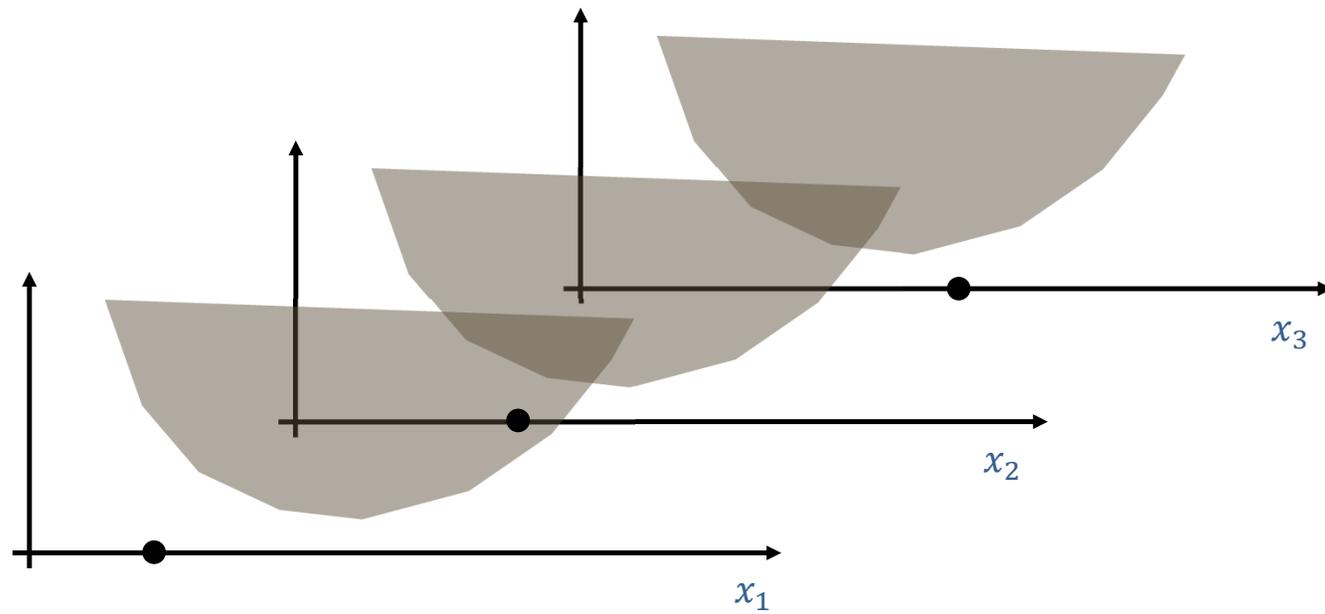
Forward iteração 1



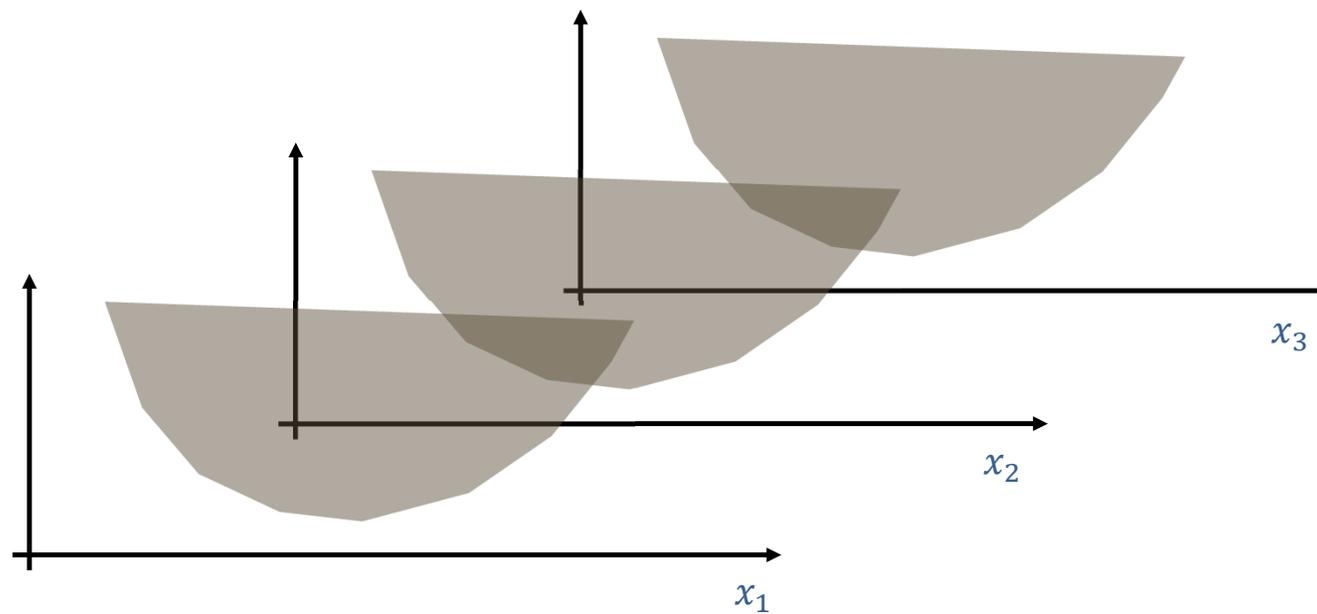
Forward iteração 1



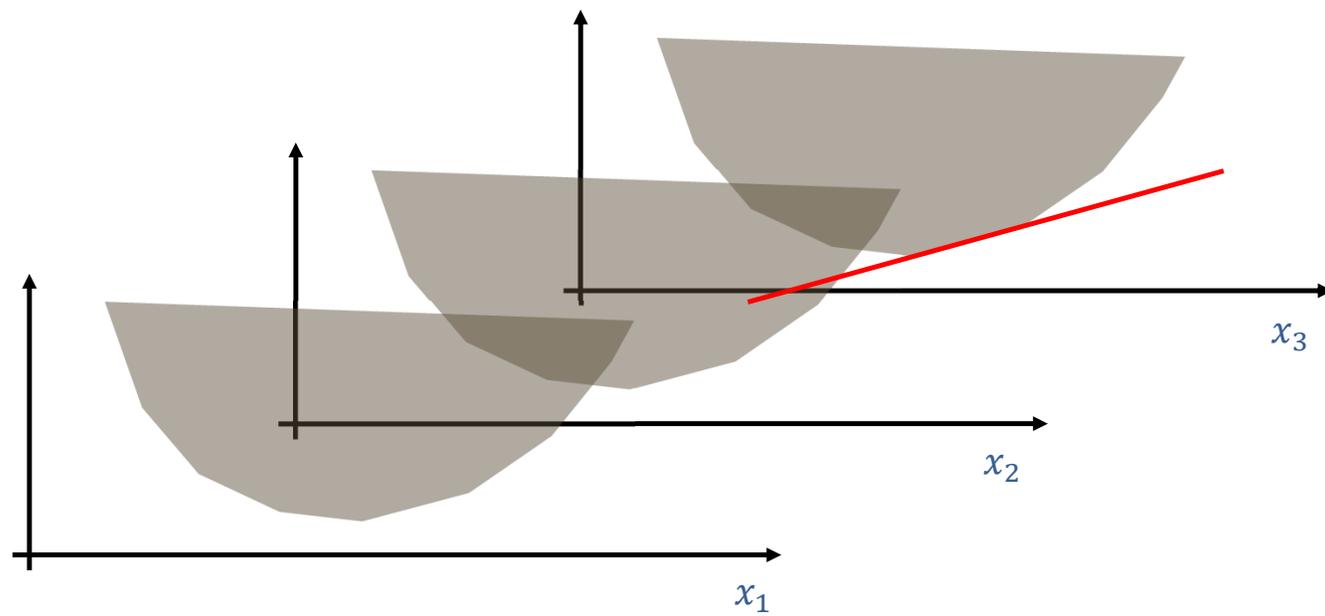
Forward iteração 1



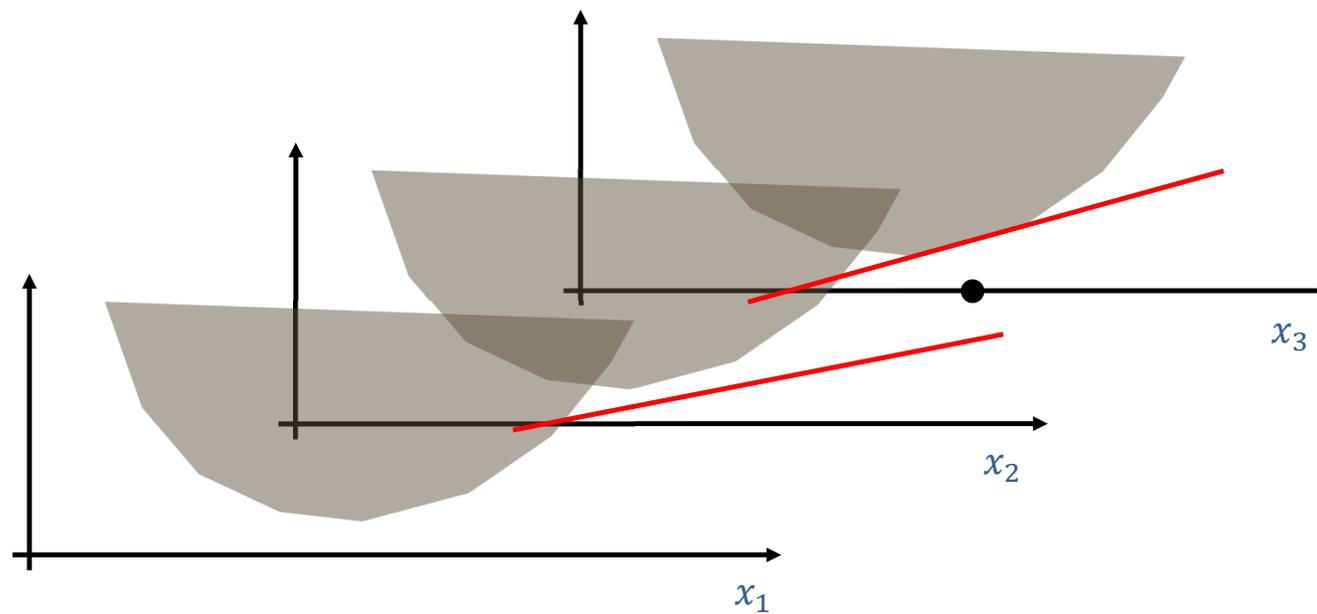
Backward iteração 1



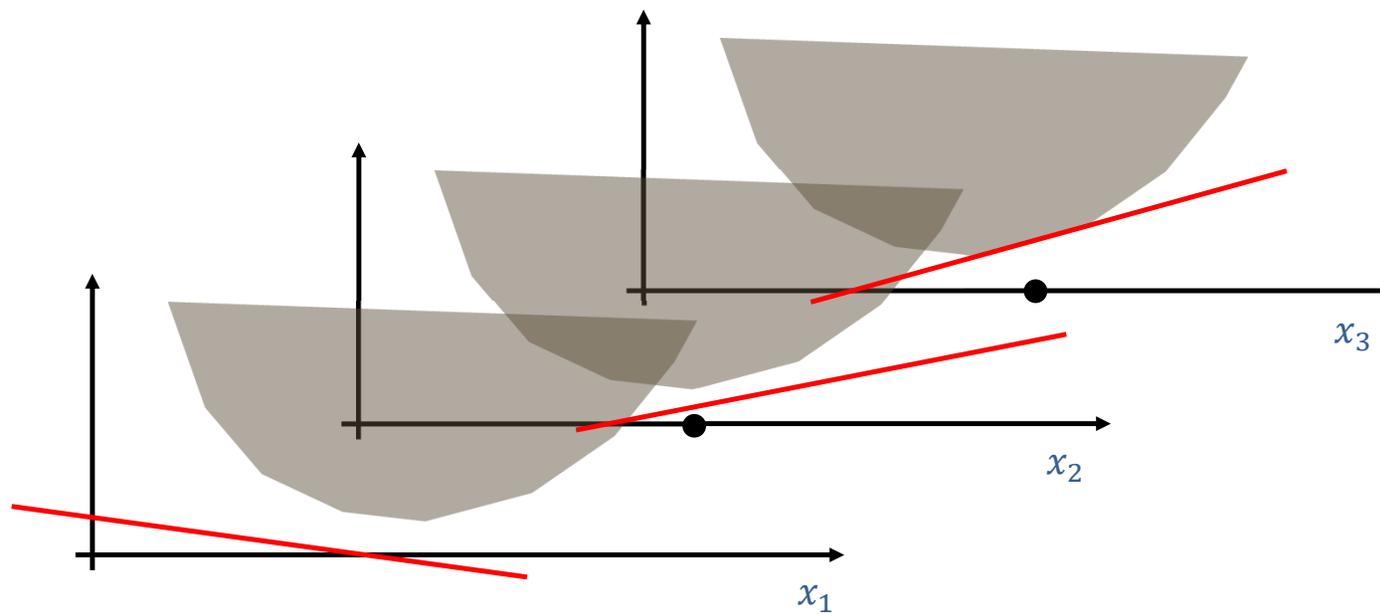
Backward iteração 1



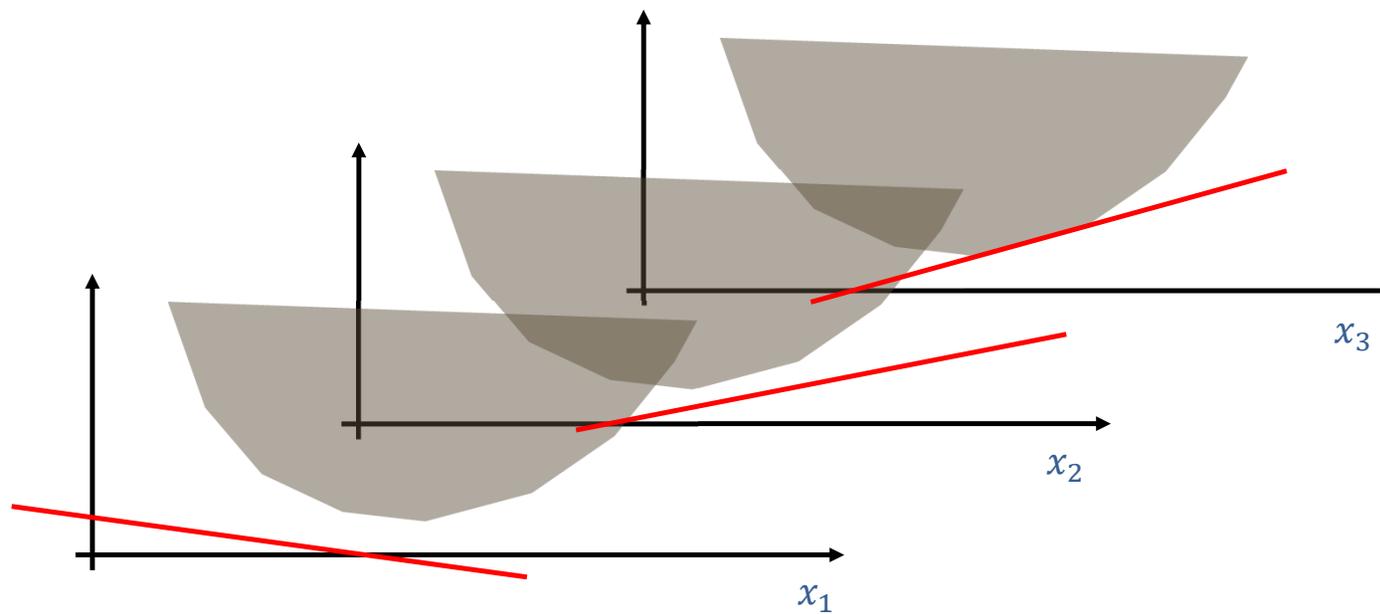
Backward iteração 1



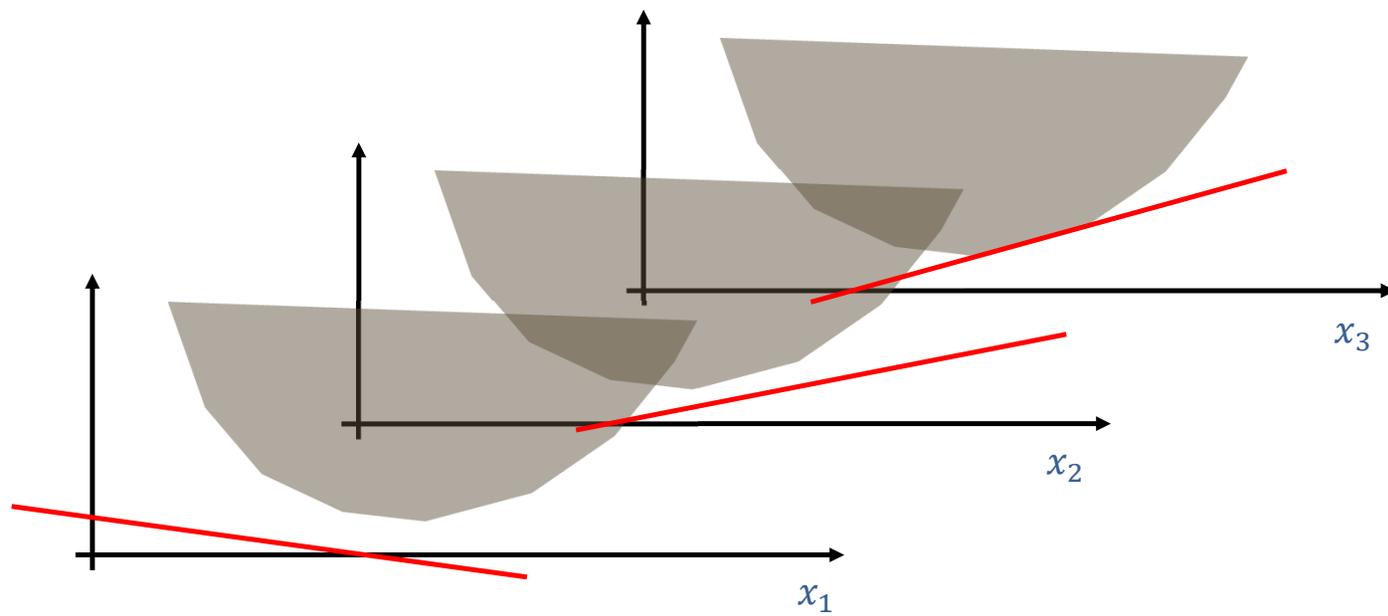
Backward iteração 1



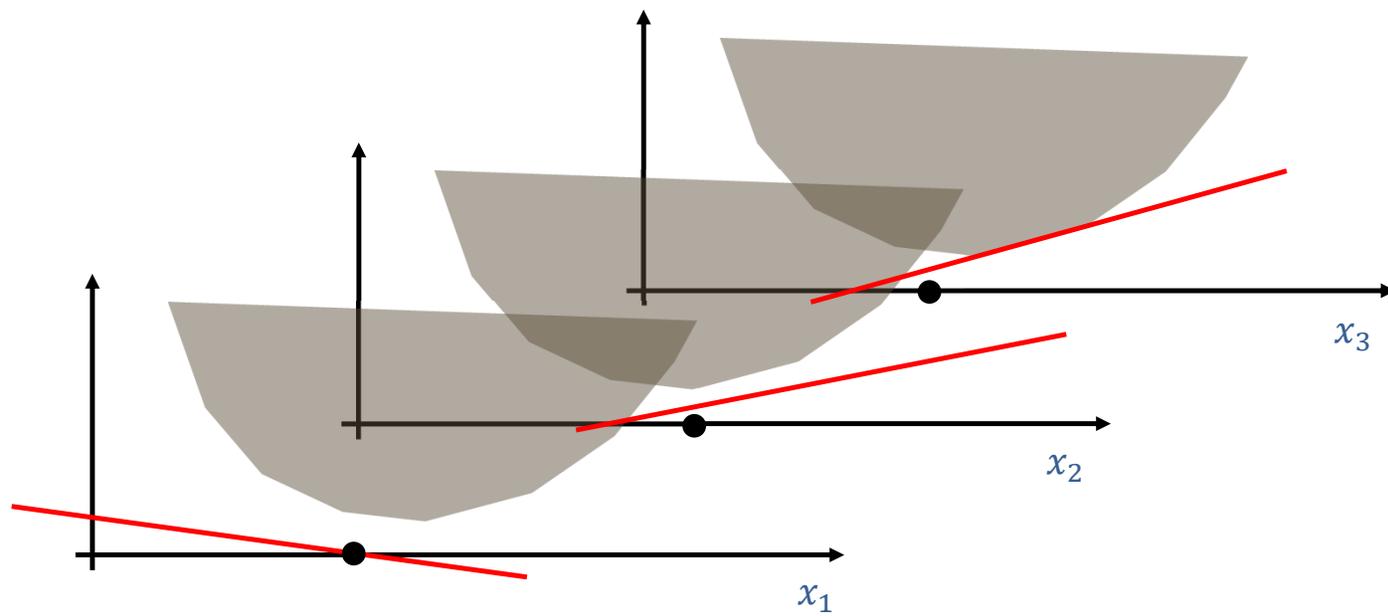
Backward iteração 1



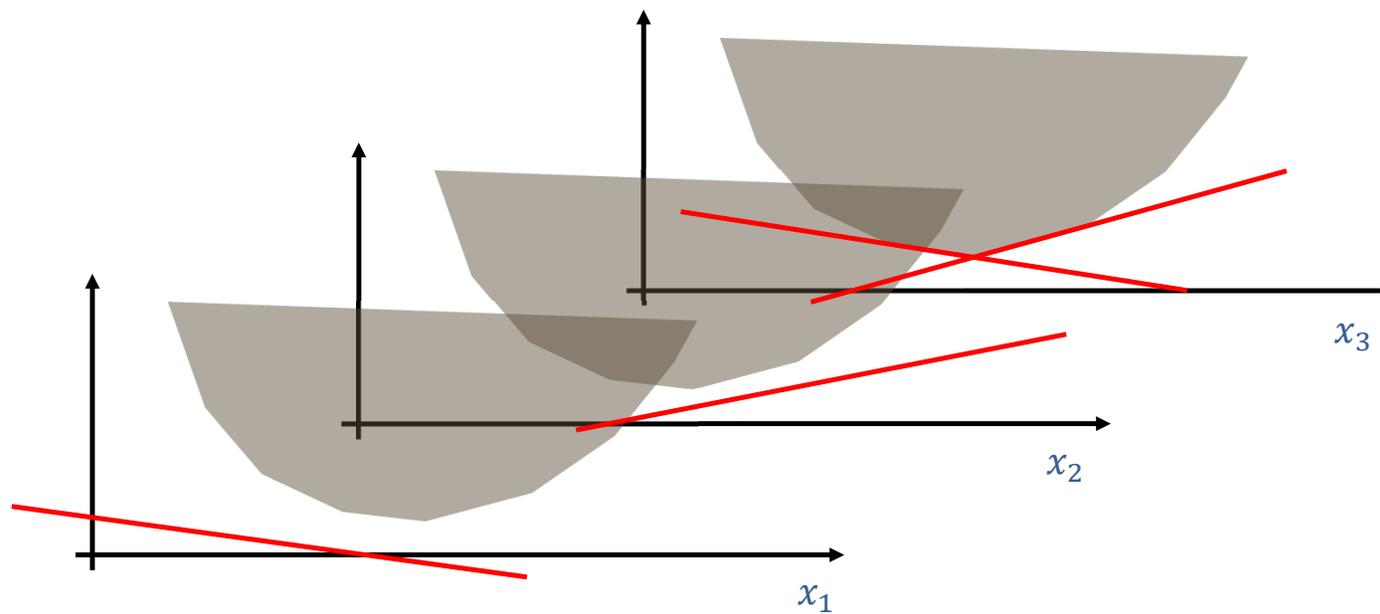
Forward iteração 2



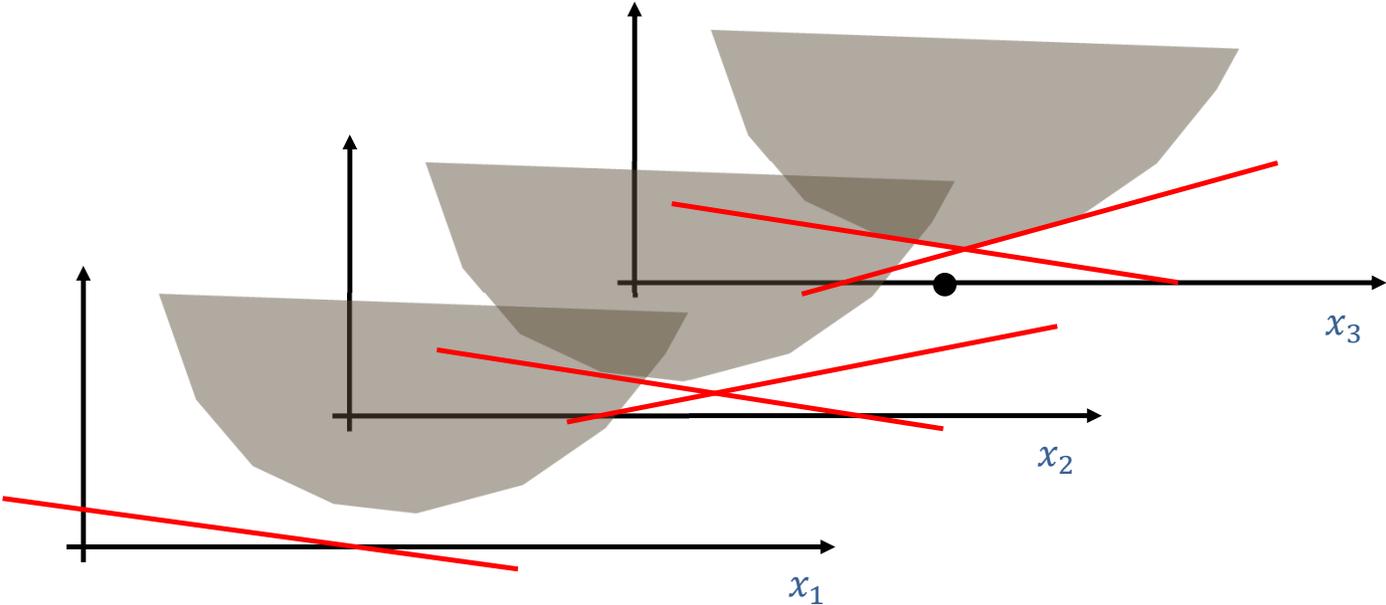
Forward iteração 2



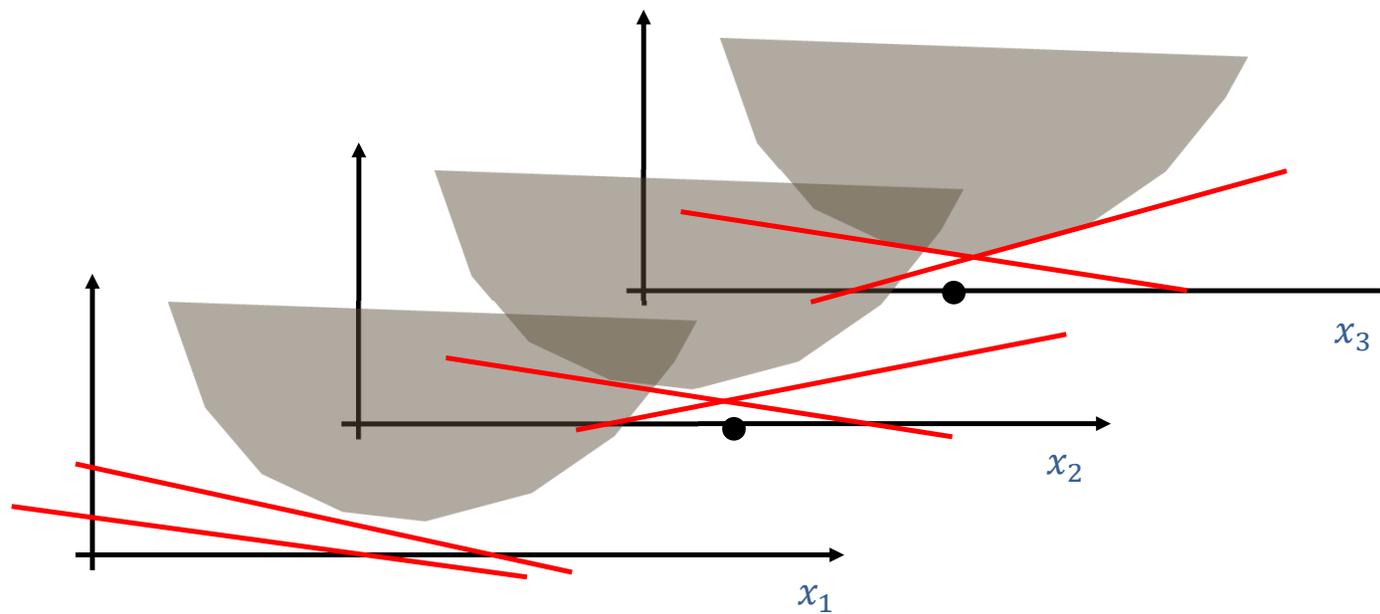
Backward iteração 2



Backward iteração 2



Backward iteração 2



Vantagens do SDDP

- ▶ **Convergência demonstrada: Theorem:** If sampling is done with replacement and basic dual optimal solutions are used to construct Benders' cuts, then, w.p. 1, **SDDP converges** in a finite number of iterations to an optimal solution. [Chen & Powell, 1999], [Philpott & Guan, 2008], [Shapiro, 2011]...
- ▶ Produz limites inferior e superior a cada iteração
- ▶ Permite obter soluções viáveis a cada iteração
 - A forward pode ser feita com modelo não convexo
- ▶ Fácil adaptar para processamento distribuído

⇒ **Interesse acadêmico e industrial no algoritmo**

The screenshot shows the AWS Case Study page for PSR. The page layout includes the AWS logo and navigation menu at the top. The main content area features a search bar for 'CUSTOMER STORES' with a 'Use Case' dropdown menu and a 'Go' button. Below this, there are links for 'Get an Amazon.com Gift Card' and 'Take Part in an AWS Usability Study'. A prominent yellow button says 'Get Started with AWS for Free' with a 'Create Free Account' sub-button. The article title is 'AWS Case Study: PSR' with the PSR logo. The 'About PSR' section describes the company's history and services. The 'The Challenge' section discusses the adoption of cloud computing. The 'Why Amazon Web Services' section explains the benefits of moving to the cloud.

SDDP & academia

- ▶ Original paper (1991): 1306 citations (Google scholar)
- ▶ ICSP 2016: 5 sessions, 22 talks and 2 plenaries (SDDP and SDDiP)
- ▶ ISMP 2018: Dedicated 3-day simposia with 20 talks
- ▶ ICSP 2019: 21 presentations out of 250! Almost 10%
- ▶ Julia and JuMP

- <https://github.com/odow/SDDP.jl>
- <https://github.com/blegat/StructDualDynProg.jl>
- <https://github.com/JuliaOpt/StochDynamicProgramming.jl>
- <https://gitlab.com/stochastic-control/StOpt>

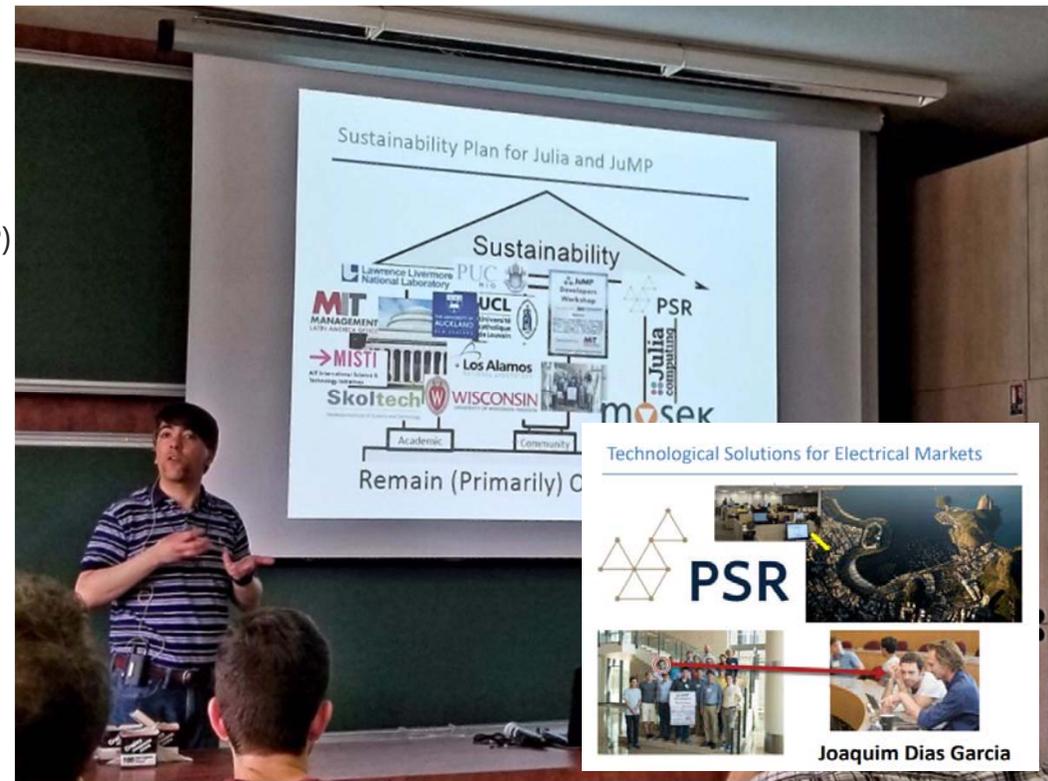


▶ MATLAB

- FAST "(Finally An SDDP Toolbox) is an easy-to-use Stochastic Dynamic Dual Programming toolbox for Matlab. It helps you model and solve your problem easily and quickly!"

▶ Python:

- <https://github.com/lingquant/msppy>
- <https://pypi.org/project/PySDDP>

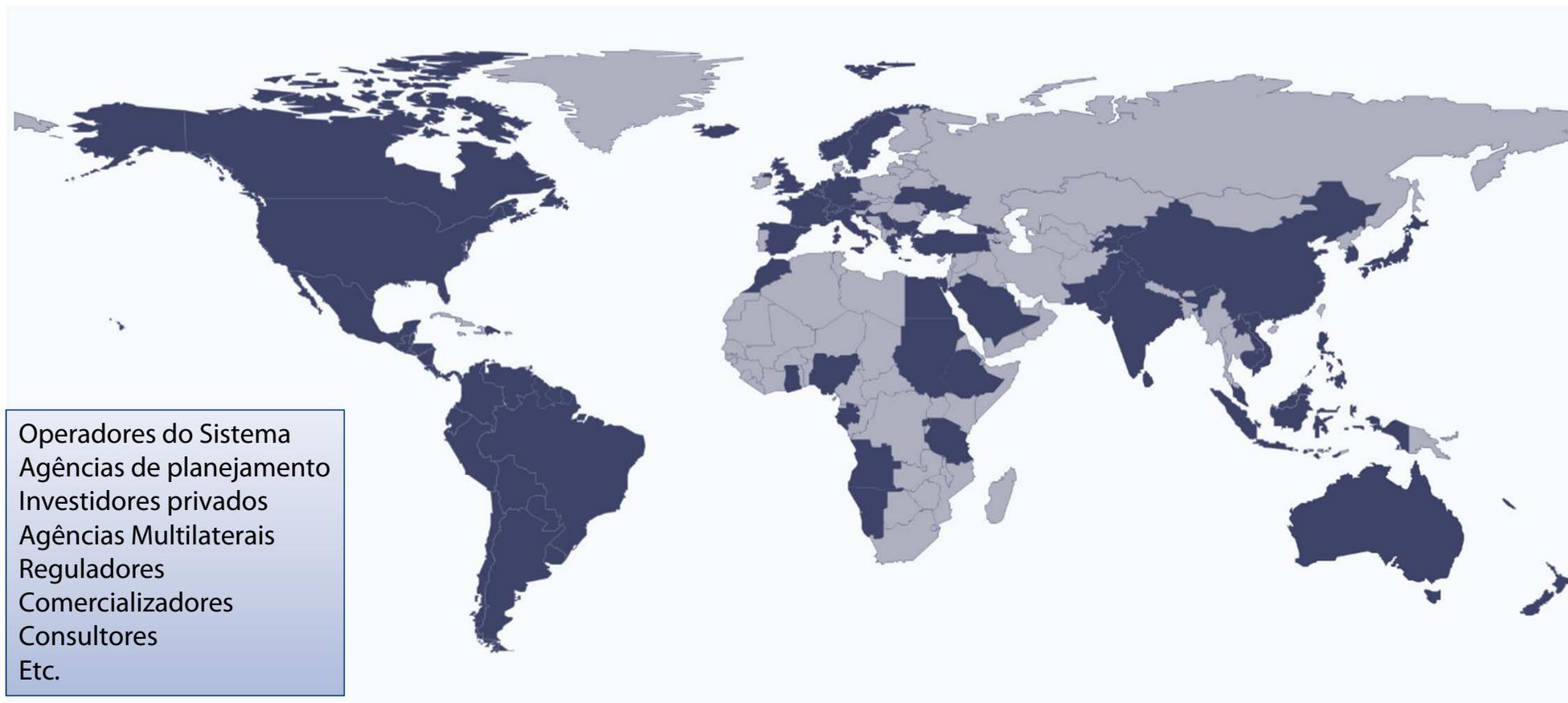


Joaquim Dias Garcia

Temário

- ▶ **Motivação: operação de sistemas elétricos com reservatórios**
- ▶ **O algoritmo SDDP**
- ▶ **Aplicações do SDDP**
- ▶ **Extensão para aversão ao risco**
- ▶ **Extensão para não convexidade**
- ▶ **SDDP e Reinforcement Learning**

O SDDP é aplicado em mais de setenta países



Brasil

- ▶ Capacidade instalada: 125 GW
- ▶ 160 hidrelétricas (85 reservatórios), 140 térmicas, 14 GW renováveis
- ▶ Rede de transmissão: 5 mil barras, 7 mil circuitos

Variáveis de estado: 85 (armazenamento) + 160 x 2 = 320 (vazões AR-2) = **405**

Estágios: 120 meses (10 anos)

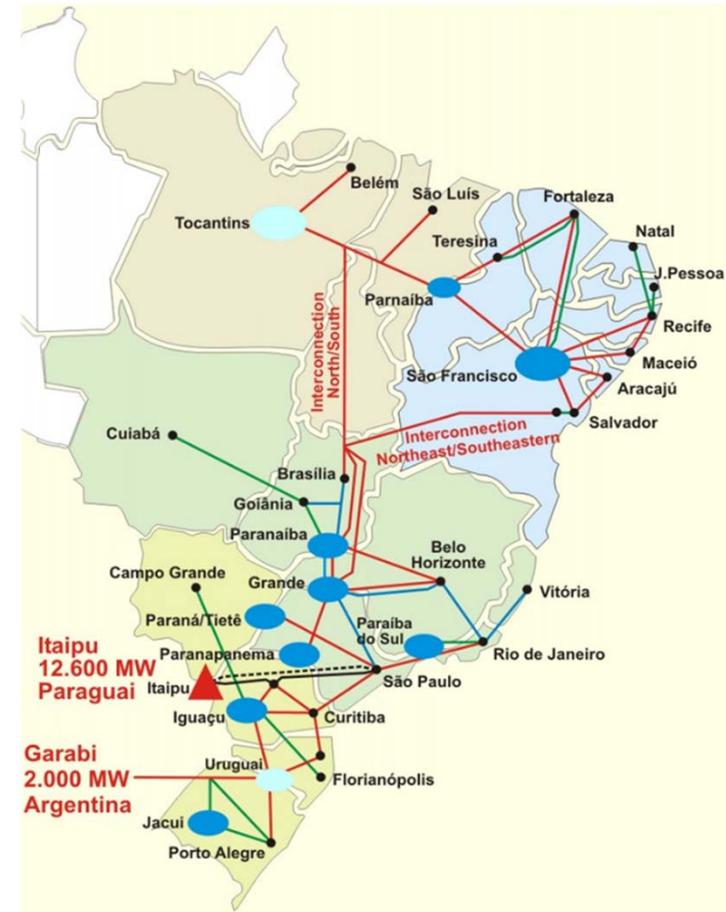
Simulação *forward*: 1.200 cenários

“Aberturas” *backward*: 30

Problemas por estágio/iteração: **36 mil**

Número de iterações: 10

Tempo total de execução: **110** minutos
25 servidores com 16 processadores cada



US Pacific Northwest

- Capacidade instalada: 63 GW
 - Conectado à Costa Oeste (300 GW)
- 80 hidrelétricas (40 reservatórios)
- 75 termelétricas
- Estágios: 104 semanas (2 anos)
Variáveis de cada problema: 30 mil
Restrições: 50 mil
Total de problemas: **2 milhões**
Tempo de execução: **130** minutos

The new SDDP Nordic

3 YEAR FORECASTS WITH IMPROVED STACK AND HYDROLOGY

In April 2008 we presented the first SDDP forecast for the Nord Pool market at the annual Montel spring conference. Our estimate was very bearish for May compared with the market, but more bullish later in the summer. It turned out that the delivered price for May was even lower than we forecasted. After that head start our medium term forecast have become an important reference for the Nordic market, most recently when the market really turned bearish this June.

Our goal is to always perform better and deliver better services to our clients, and over the years we have seen some areas of improvement. Most notably is the new price areas in both Norway and Sweden, but we also wanted a better coupling between our hydrological (HBV) models and a full revision of the stack.

Hence, over the last year we have put a lot of effort in recalibrating the SDDP model at the same time as publishing our weekly forecast. Last week we published our first forecast with the recalibrated SDDP Nordic. The SDDP methodology is developed by PSR in Brazil, a strategic partner of Thomson Reuters.



NEW FEATURES

The main new feature of the new SDDP model is a detailed modeling of all the 12 Nord Pool price areas. The historical inflow series have been updated as well, based on the years 1981 to 2007. However, from week 1 to week 40 in the SDDP forecast we use the latest HBV long term forecast based on the latest EC00 ens the first two weeks and historic temperatures and precipitation thereafter. There is good match between the price areas and the hydro regions, although there are some minor deviations between NO5 and the corresponding hydro region (NO6 in the map over hydro regions).

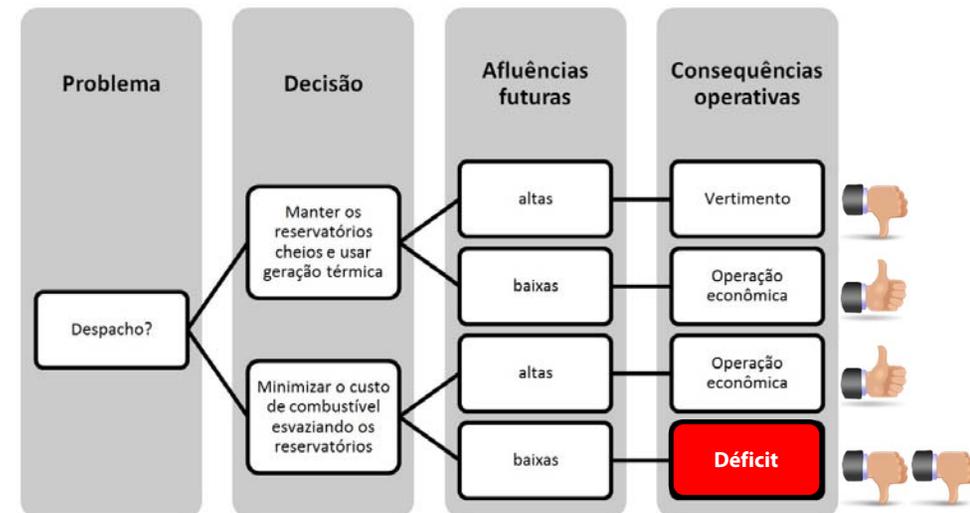
We have updated the load using weekly load levels that has been temperature corrected against

Temário

- ▶ **Motivação: operação de sistemas elétricos com reservatórios**
- ▶ **O algoritmo SDDP**
- ▶ **Aplicações do SDDP**
- ▶ **Extensão para aversão ao risco**
- ▶ **Extensão para não convexidade**
- ▶ **SDDP e Reinforcement Learning**

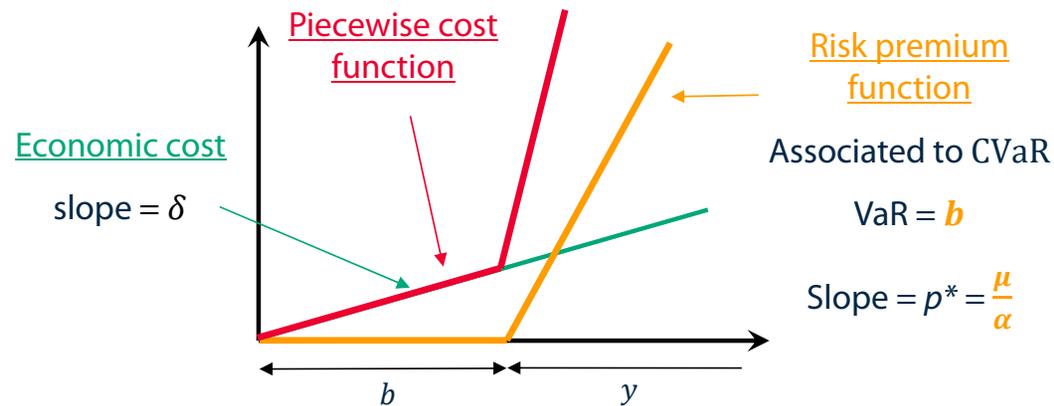
Abordagem neutra x avessa ao risco

- ▶ Minimização do valor esperado → abordagem neutra ao risco
 - Busca por políticas que minimizam o custo de operação *em média*
 - Do ponto de vista econômico, seria razoável aceitar déficits de energia ocasionais em troca da redução dos custos operativos no longo prazo
- ▶ Mesmo com penalizações pelos déficits, o valor esperado pode não representar o real impacto para a sociedade
- ▶ Por esta razão, operadores do sistema têm interesse em adotar políticas operativas **avessas ao risco**



Abordagem #1: Restrição CVaR associada à confiabilidade de suprimento

- ▶ Objetivo: determinar uma política operativa que minimize o valor esperado do custo operativo, sujeito à uma restrição que limita o déficit de energia baseada em CVaR
 - Exemplo: média dos $\alpha = 1\%$ piores déficits ($\text{CVaR}_{1\%}$) menor que 5% da demanda
 - Abordagem geral de solução: relaxação Lagrangeana (permite separabilidade)

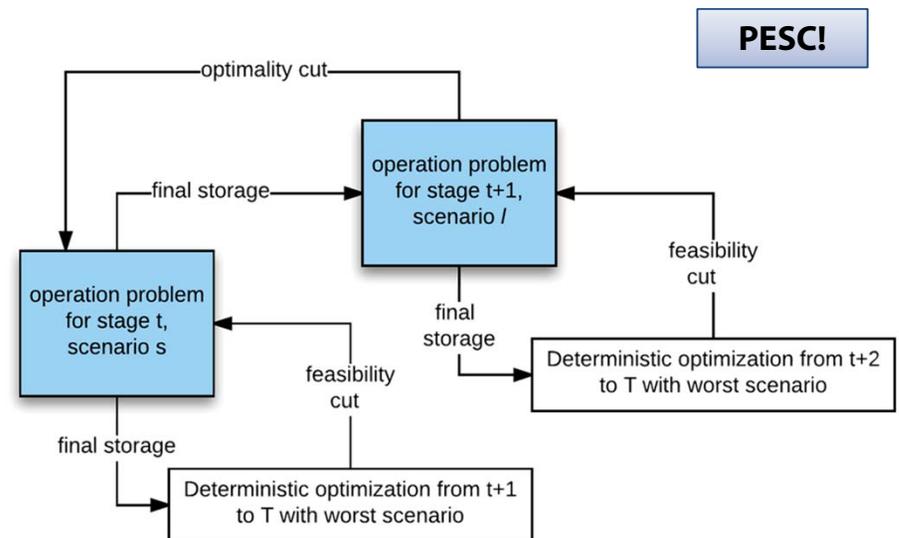
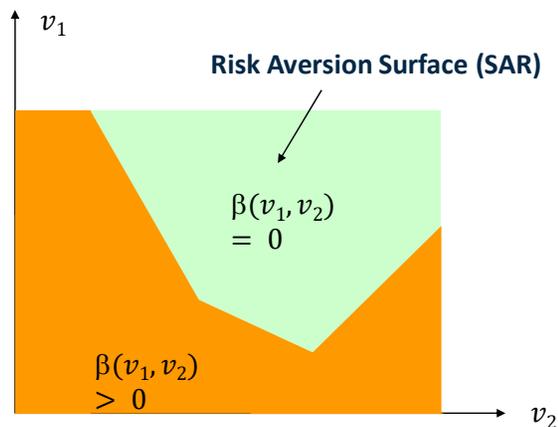


"Representação de restrições de aversão a risco de CVaR em Programação Dinâmica Dual Estocástica com aplicação ao planejamento da operação de sistemas hidrotérmicos". Tese de Doutorado – L. C. Costa Junior - COPPE/2013

PESCI!

Abordagem #2: Superfície de Aversão ao Risco (SAR)

- ▶ Antecedentes: Curva de Aversão ao Risco (CAR), proposta pela PSR e implantada após racionamento de 2002
- ▶ Proposta pela PSR em 2008 como generalização da CAR para múltiplos reservatórios
- ▶ Baseada em cortes de viabilidade associados a um conjunto de cenários críticos



PESC!

Abordagem #3: CVaR associado aos custos operativos

- ▶ Atribuir um peso relativo maior aos cenários de afluências desfavoráveis na minimização dos custos imediato e futuro
 - Minimização de uma função de risco
 - Exemplo: ponderação entre $\mathbb{E}[CO]$ e $CVaR[CO]$
 - [Shapiro, 2011], [Philpott, De Matos, 2010], [Shapiro et al., 2013]

$$\text{Min } (1 - \lambda)\mathbb{E}(\cdot) + \lambda CVaR_{\alpha\%}(\cdot)$$

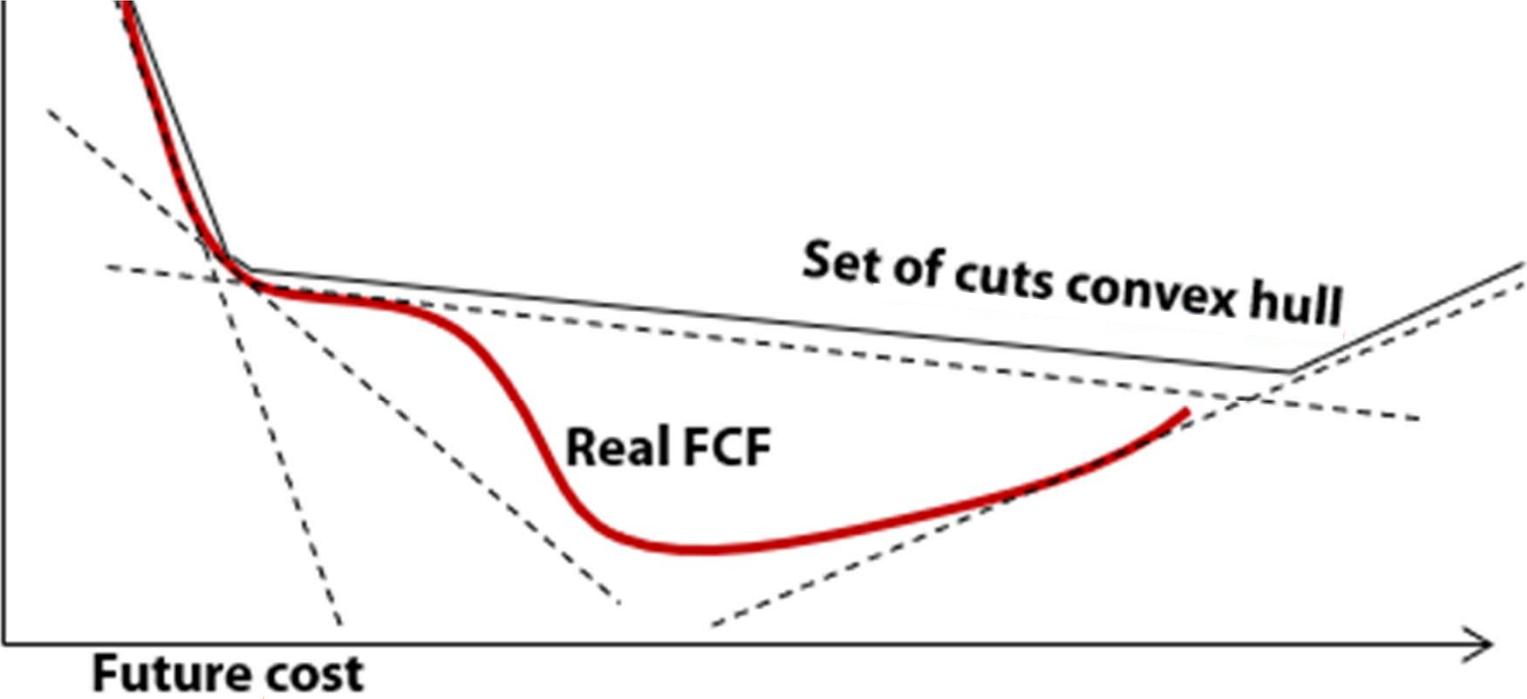
- ▶ A abordagem de CVaR-custo representa *indiretamente* o critério de segurança de suprimento
 - maior $\lambda \rightarrow$ maior peso para os cenários desfavoráveis \rightarrow maior despacho térmico preventivo \rightarrow maior segurança de suprimento

Colaboração Shapiro com
Joari Costa (PESCI)

Temário

- ▶ **Motivação: operação de sistemas elétricos com reservatórios**
- ▶ **O algoritmo SDDP**
- ▶ **Aplicações do SDDP**
- ▶ **Extensão para aversão ao risco**
- ▶ **Extensão para não convexidade**
- ▶ **SDDP e Reinforcement Learning**

Limitações no caso de não convexidade



Abordagens de solução para o caso não convexo

- ▶ Lagrangian Cuts - 2014
 - Best convex outer approximation – **Fernanda Thomé**, M. Pereira, S. Granville, M. Fampa
- ▶ SDDiP (stochastic dual dynamic integer programming) - 2019
 - All states must be binary variables (uses Lagrangian cuts) - J. Zou, S. Ahmed, A. Sun
- ▶ MIDAS (mixed integer dynamic approximation scheme) - 2019
 - Future cost functions must be monotone - A. Philpot, F. Wahid, J. Bonnans
- ▶ SLDP (stochastic lipschitz dynamic programming) – unpublished
 - Tent shaped function as lower approximations - S. Ahmed, F. Cabral, B. da Costa
- ▶ **Locally Valid Cuts**
- ▶ H. Abgottspon (2014)
- ▶ SDDP-SVM (**Julio Alberto Dias** et al. 2016) – A seguir

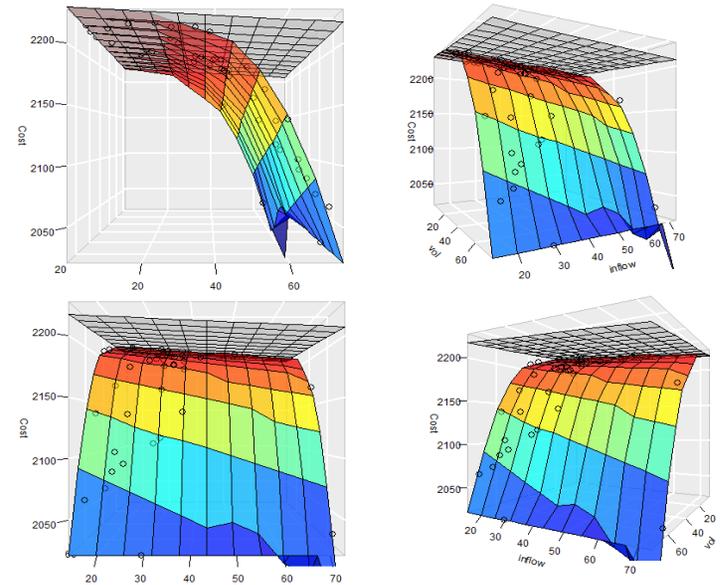
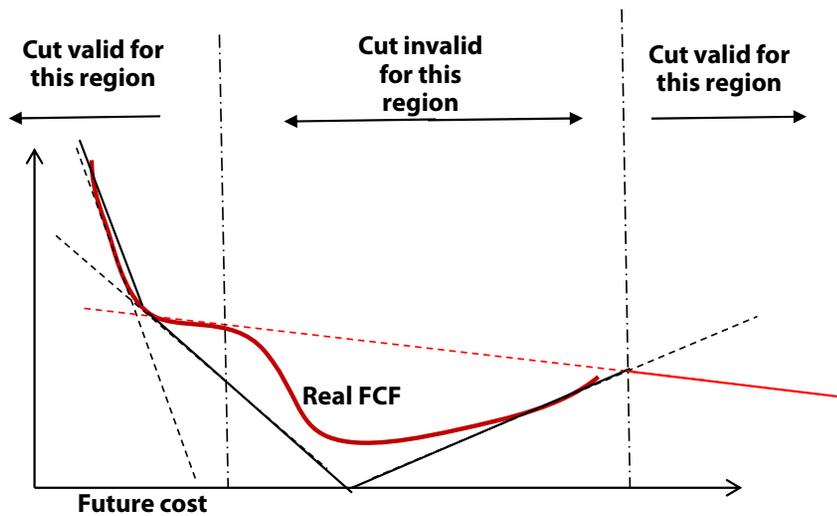
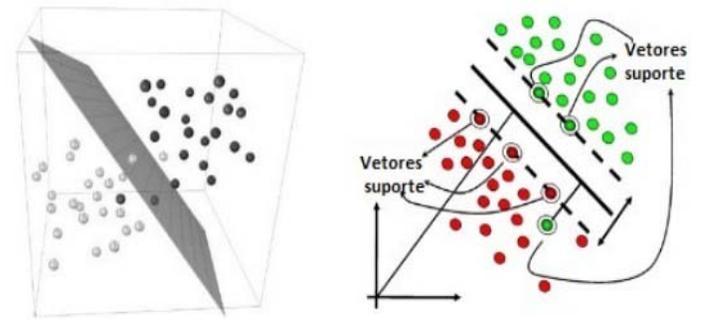
COPPE/Elétrica



PESC!

SDDP-SVM

- ▶ A FCF é representada como a união de regiões localmente convexas. Um classificador baseado em Support Vector Machine (SVM) determina quais hiperplanos são válidos em cada região

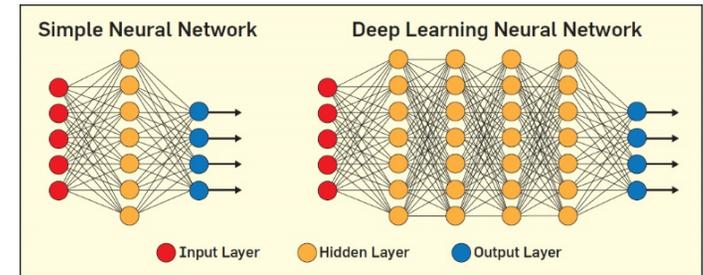


Temário

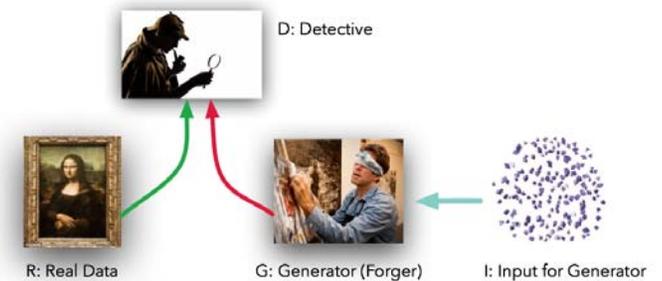
- ▶ **Motivação: operação de sistemas elétricos com reservatórios**
- ▶ **O algoritmo SDDP**
- ▶ **Aplicações do SDDP**
- ▶ **Extensão para aversão ao risco**
- ▶ **Extensão para não convexidade**
- ▶ **SDDP e Reinforcement Learning**

A revolução da IA/ML

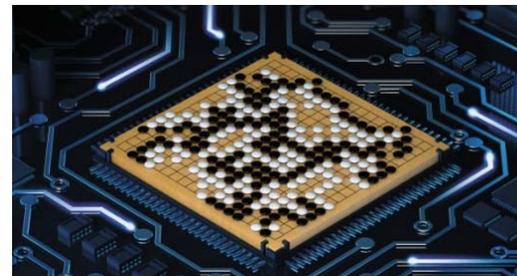
- ▶ Melhor capacidade de previsão \Rightarrow Deep Learning



- ▶ Melhor capacidade de criação de cenários \Rightarrow GAN



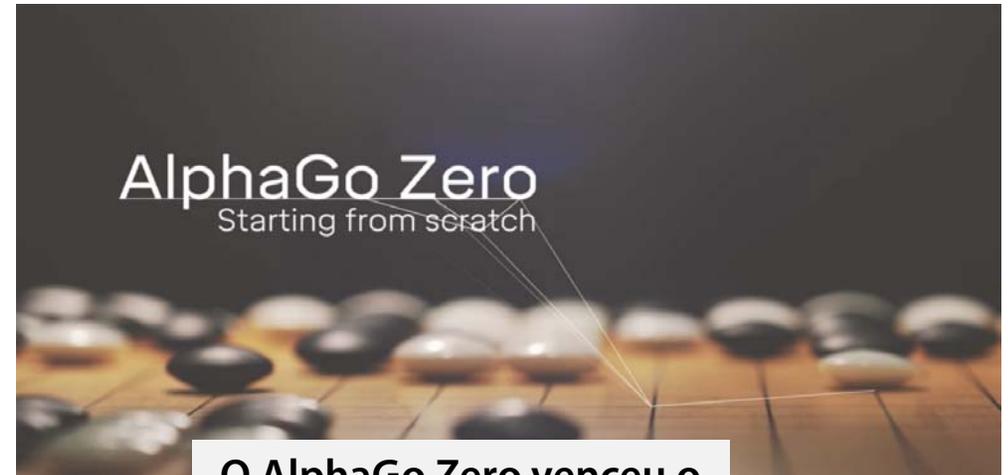
- ▶ Decisões multi-estágio sob incerteza
 \Rightarrow Reinforcement Learning



Reinforcement Learning: AlphaGo e AlphaGo Zero



O AlphaGo
venceu o
campeão mundial
4 x 1



O AlphaGo Zero venceu o
AlphaGo por 100 x zero!

Como visto, o SDDP é um algoritmo para
tomada de decisão multi-estágio sob incerteza
Existe alguma relação entre RL e SDDP?

REINFORCEMENT LEARNING E SDDP

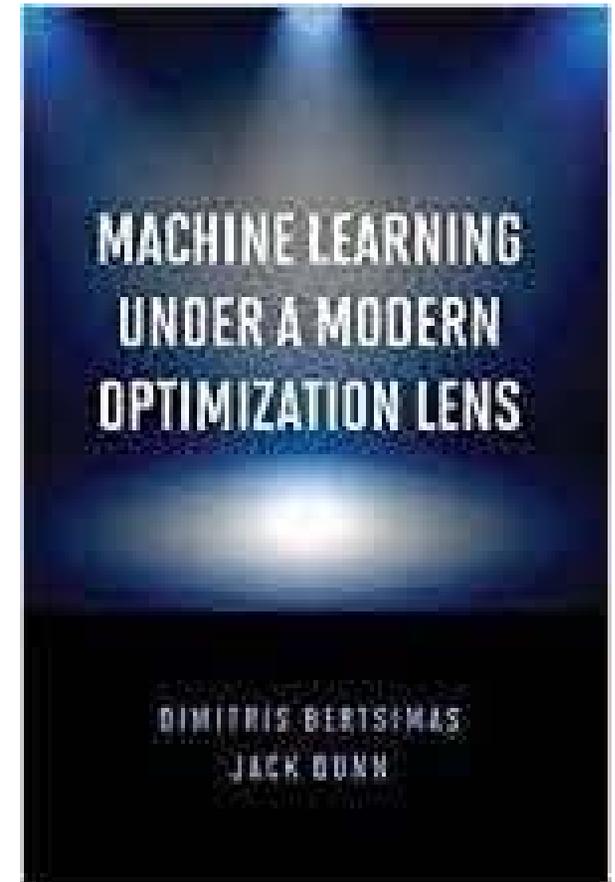
Learning continuous Q -functions using generalized Benders cuts

Joseph Warrington, *Member, IEEE*

Abstract— Q -functions are widely used in discrete-time learning and control to model future costs arising from a given control policy, when the initial state and input are given. Although some of their properties are understood, Q -functions generating optimal policies for continuous problems are usually hard to compute. Even when a system model is available, optimal control is generally difficult to achieve except in rare cases where an analytical solution happens to exist, or an explicit exact solution can be computed. It is typically necessary to discretize the state and action spaces, or parameterize the Q -function with a basis that can be hard to select *a priori*. This paper describes a model-based algorithm based on generalized Benders theory that yields ever-tighter outer-approximations of the optimal Q -function. Under a strong duality assumption, we prove that the algorithm yields an arbitrarily small Bellman optimality error at any finite number of arbitrary points in the state and input space, in finite iterations. Under additional assumptions, the same guarantee holds when the inputs are determined online by the algorithm's updating Q -function. We demonstrate these properties numerically on scalar and 8-dimensional systems

[5]. However, performing even one iteration of the canonical algorithms, such as value iteration, then has an undesirable exponential cost. ADP methods have arisen to find more tractable parameterizations of the continuous V -function. Several are based on continuous extensions of the “linear programming approach” to ADP [6], in which a valid lower bound on the optimal value function is maximized. Examples include the quadratic lower bound in [14], and the polynomial derived using sum-of-squares techniques in [13]. Approximate V -functions represented as the pointwise maximum of multiple lower-bounding functions have been used in [1], [9], [10], [15]. Recent work utilizing a piecewise-linear maximum representation [16] has extended the Benders decomposition argument used for linear multi-stage decision problems in Dual DP (DDP, [12]), to a general nonlinear, infinite-horizon setting. In this paper we adapt the Benders approach from [16] to learn Q -functions. We define an algorithm that successively

- [12] M. V. F. Pereira and L. M. V. G. Pinto. Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming*, 52(1-3):359–375, 1991.

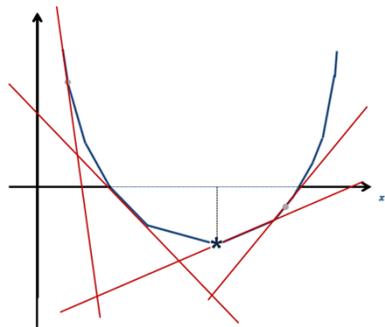


SDDP x RL (parceria com Kunumi)

- ▶ Bellman value function representation $V_t(x)$
- ▶ Approximation method: multi-dimensional **piecewise linear** functions

Benders Cuts

Globally convergent
Convex functions

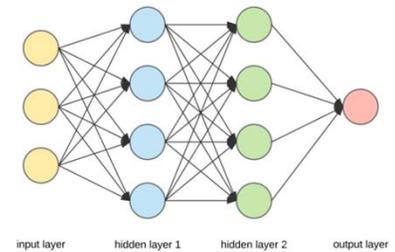
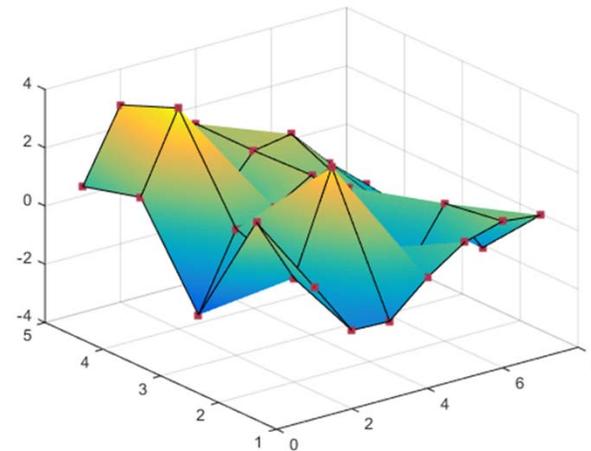


Recent unifying works:

Using RL for convex problems, Non-linear SDDP (SDDiP, SDDP-SVM, SLDP...)

Neural Networks*

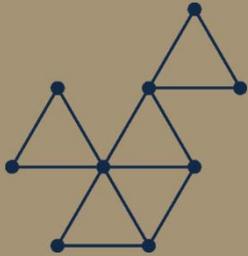
Locally Convergent
Non-linear functions



Em resumo...

- ▶ O caso do SDDP ilustra os benefícios de uma parceria entre universidade e indústria; somos muito gratos a tod@s os professores e colegas do PESC
- ▶ Estes benefícios só devem aumentar, pois o ativo mais valioso para as empresas será a capacidade analítica e de inovação das pessoas
- ▶ A otimização, computação de alto desempenho e data science, que estão no núcleo do PESC, estão entre os instrumentos mais importantes para estas inovações

**50 IS THE
NEW 30**



 www.psr-inc.com

 psr@psr-inc.com

 +55 21 3906-2100

 /psrenergy

 @psrenergy

 @psrenergy

MUITO OBRIGADO

