



Formulação de um problema inverso não-linear

Prof. André L. A. dos Reis

Rio de Janeiro 2023

Objetivos da aula

- * A formulação de um problema inverso não linear
- * Métodos por gradiente:
 - Steepest descent
 - Newton
 - Gauss-Newton
 - Levenberg-Marquardt
- * Diferença entre os métodos

Formulação de um problema inverso não-linear

Formulação matemática

Dados preditos: dados gerados por um conjunto de parâmetros que descrevem o nosso modelo, que nos traz a relação física e matemática do problema

Modelagem direta: *sintoniza manualmente* o conjunto de parâmetros e compara com os dados observados

Dados observados: medições realizadas na superfície da Terra

Dados preditos

Dados observados

Vetor de parâmetros: conjunto de variáveis que descrevem o modelo

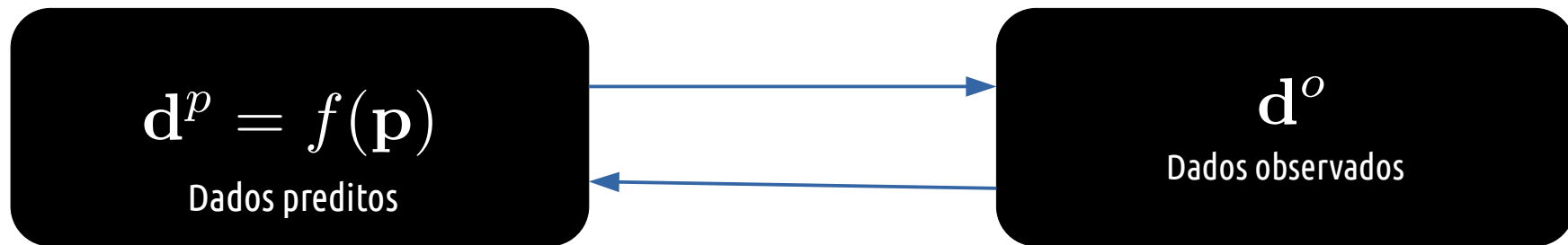
Problema inverso: *estima automaticamente* o conjunto de parâmetros a partir dos dados observados

Formulação matemática

Dados preditos: dados gerados por um conjunto de parâmetros que descrevem o nosso modelo, que nos traz a relação física e matemática do problema

Modelagem direta: *sintoniza manualmente* o conjunto de parâmetros e compara com os dados observados

Dados observados: medições realizadas na superfície da Terra



Vetor de parâmetros: conjunto de variáveis que descrevem o modelo

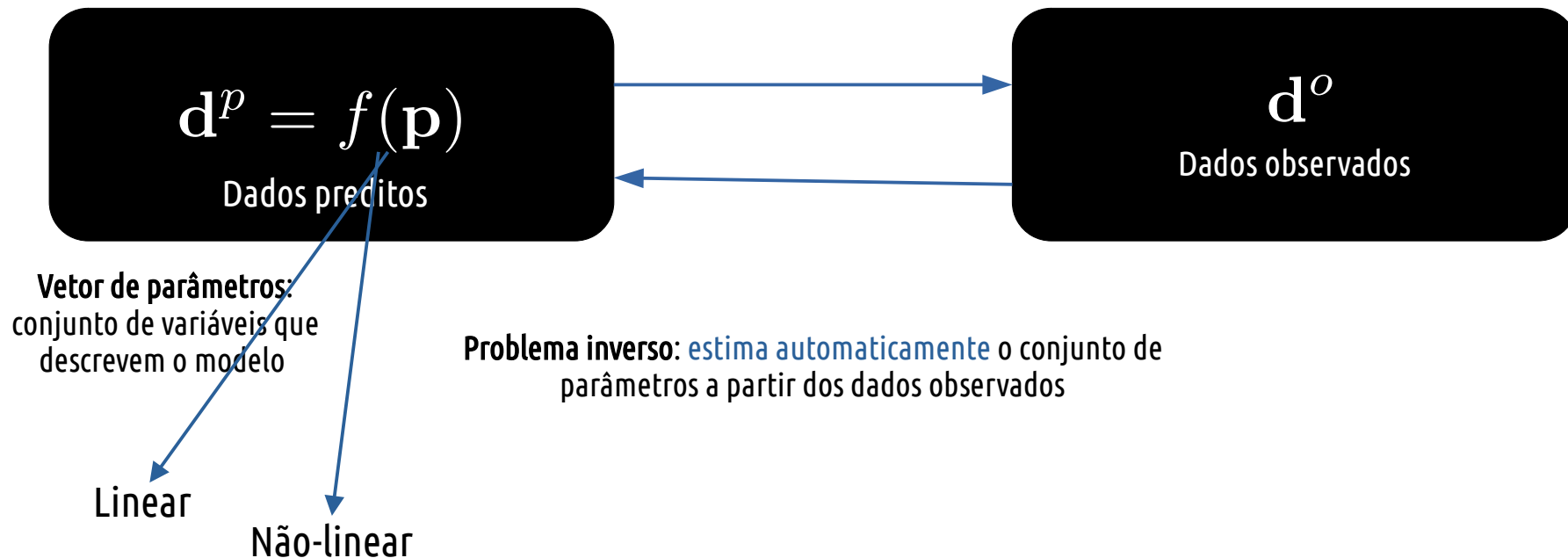
Problema inverso: *estima automaticamente* o conjunto de parâmetros a partir dos dados observados

Formulação matemática

Dados preditos: dados gerados por um conjunto de parâmetros que descrevem o nosso modelo, que nos traz a relação física e matemática do problema

Modelagem direta: *sintoniza manualmente* o conjunto de parâmetros e compara com os dados observados

Dados observados: medições realizadas na superfície da Terra

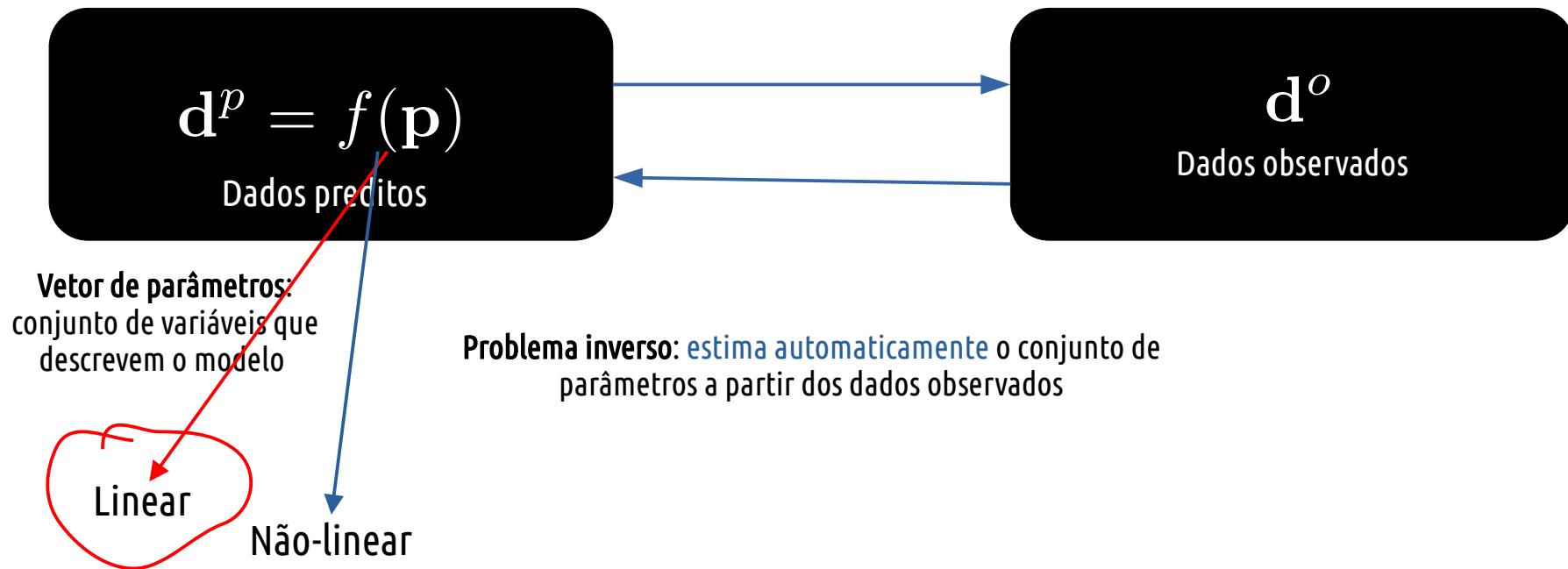


Formulação matemática

Dados preditos: dados gerados por um conjunto de parâmetros que descrevem o nosso modelo, que nos traz a relação física e matemática do problema

Modelagem direta: *sintoniza manualmente* o conjunto de parâmetros e compara com os dados observados

Dados observados: medições realizadas na superfície da Terra

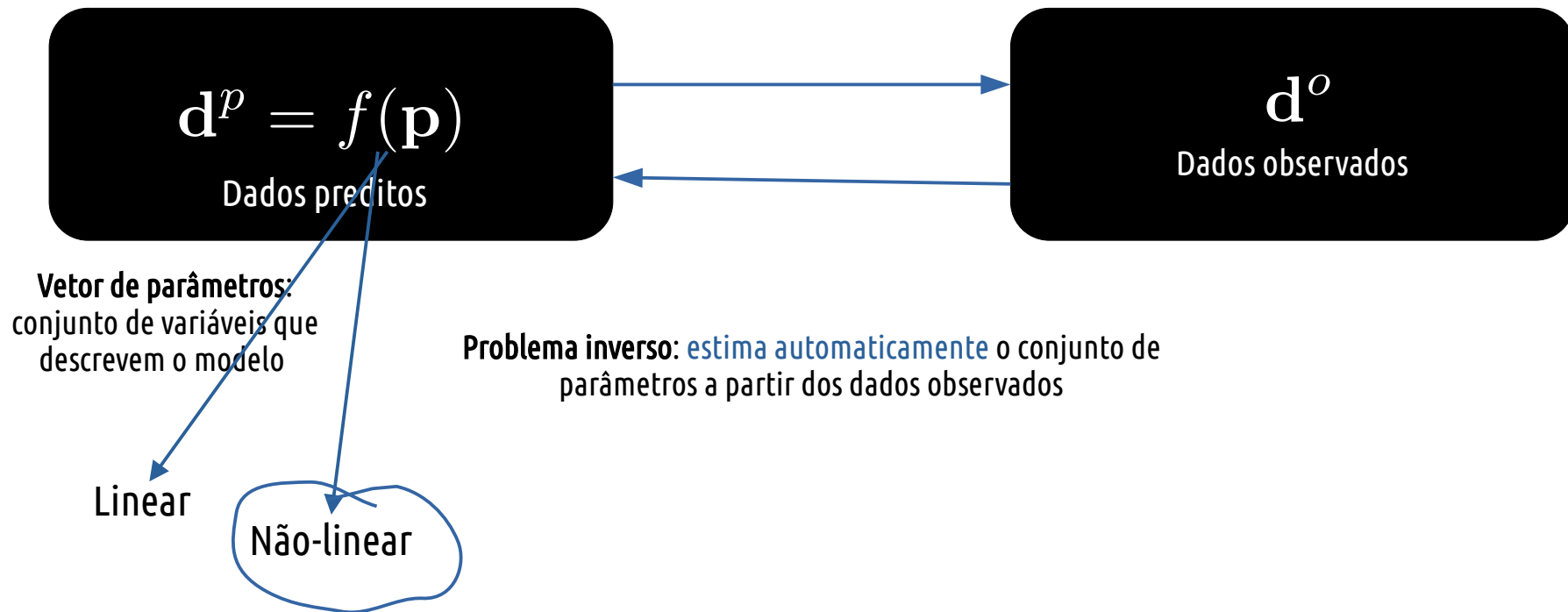


Formulação matemática

Dados preditos: dados gerados por um conjunto de parâmetros que descrevem o nosso modelo, que nos traz a relação física e matemática do problema

Modelagem direta: *sintoniza manualmente* o conjunto de parâmetros e compara com os dados observados

Dados observados: medições realizadas na superfície da Terra



Introdução ao problema não-linear

Minimizar a norma euclidiana entre o vetor de dados observados e o vetor de dados preditos!

$$\varphi(\mathbf{p}) = \| \mathbf{d}^o - \mathbf{d}^p \|_2$$

Função de ajuste

$$\nabla \varphi(\mathbf{p}) = \mathbf{0}$$

Gradiente da função de ajuste

Problema linear

$$\mathbf{d}^p = \mathbf{G}\mathbf{p}$$

$$(\mathbf{G}^T \mathbf{G}) \mathbf{p}^\# = \mathbf{G}^T \mathbf{d}^o$$

Estimador de mínimos quadrados

Problema não-linear

Introdução ao problema não-linear

Minimizar a norma euclidiana entre o vetor de dados observados e o vetor de dados preditos!

$$\varphi(\mathbf{p}) = \| \mathbf{d}^o - \mathbf{d}^p \|_2$$

Função de ajuste

$$\nabla \varphi(\mathbf{p}) = \mathbf{0}$$

Gradiente da função de ajuste

Problema linear

$$\mathbf{d}^p = \mathbf{G}\mathbf{p}$$

$$(\mathbf{G}^T \mathbf{G}) \mathbf{p}^\# = \mathbf{G}^T \mathbf{d}^o$$

Estimador de mínimos quadrados

Problema não-linear

Introdução ao problema não-linear

Minimizar a norma euclidiana entre o vetor de dados observados e o vetor de dados preditos!

$$\varphi(\mathbf{p}) = \| \mathbf{d}^o - \mathbf{d}^p \|_2$$

Função de ajuste

$$\nabla \varphi(\mathbf{p}) = \mathbf{0}$$

Gradiente da função de ajuste

Problema linear

$$\mathbf{d}^p = \mathbf{G}\mathbf{p}$$

$$(\mathbf{G}^T \mathbf{G}) \mathbf{p}^\# = \mathbf{G}^T \mathbf{d}^o$$

Estimador de mínimos quadrados

Problema não-linear

$$\mathbf{d}^p \neq \mathbf{G}\mathbf{p}$$

??????

Introdução ao problema não-linear

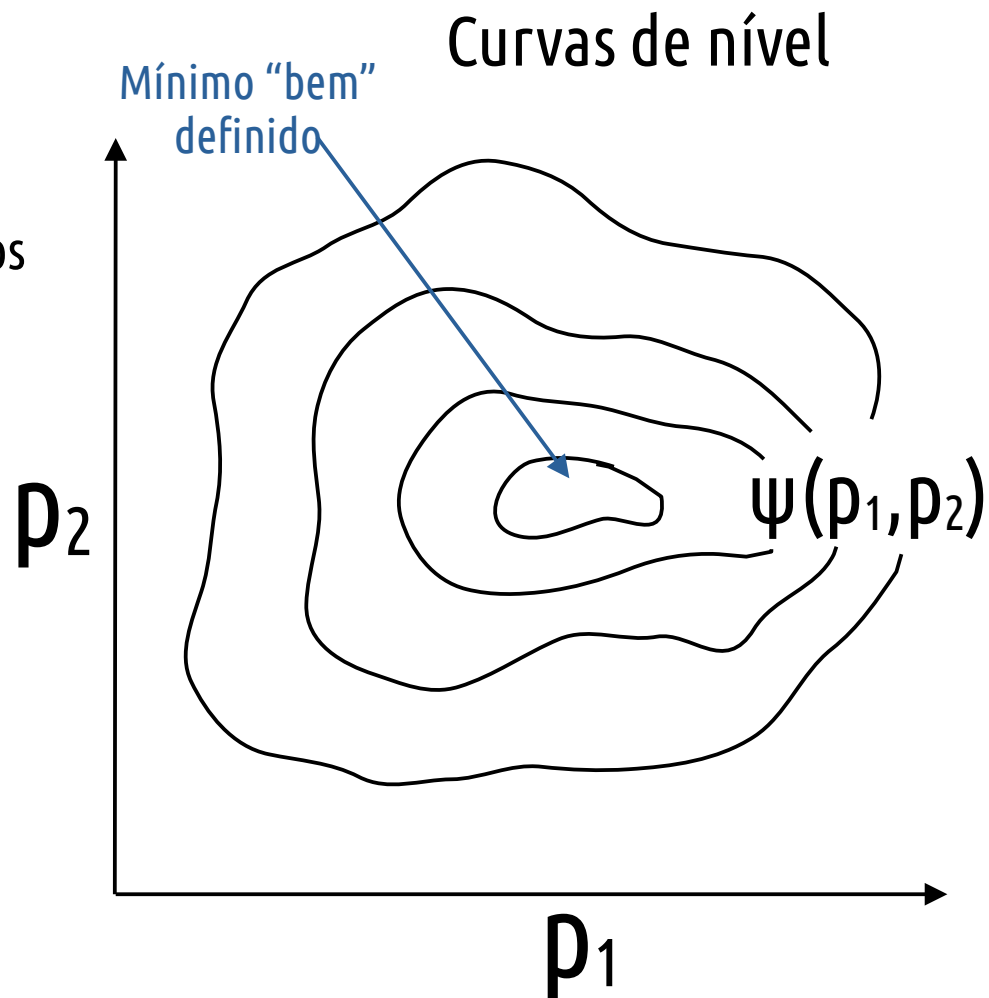
Minimizar a norma euclidiana entre o vetor de dados observados e o vetor de dados preditos!

$$\varphi(\mathbf{p}) = \| \mathbf{d}^o - \mathbf{d}^p \|_2$$

Função de ajuste

$$\nabla \varphi(\mathbf{p}) = \mathbf{0}$$

Gradiente da função de ajuste



Introdução ao problema não-linear

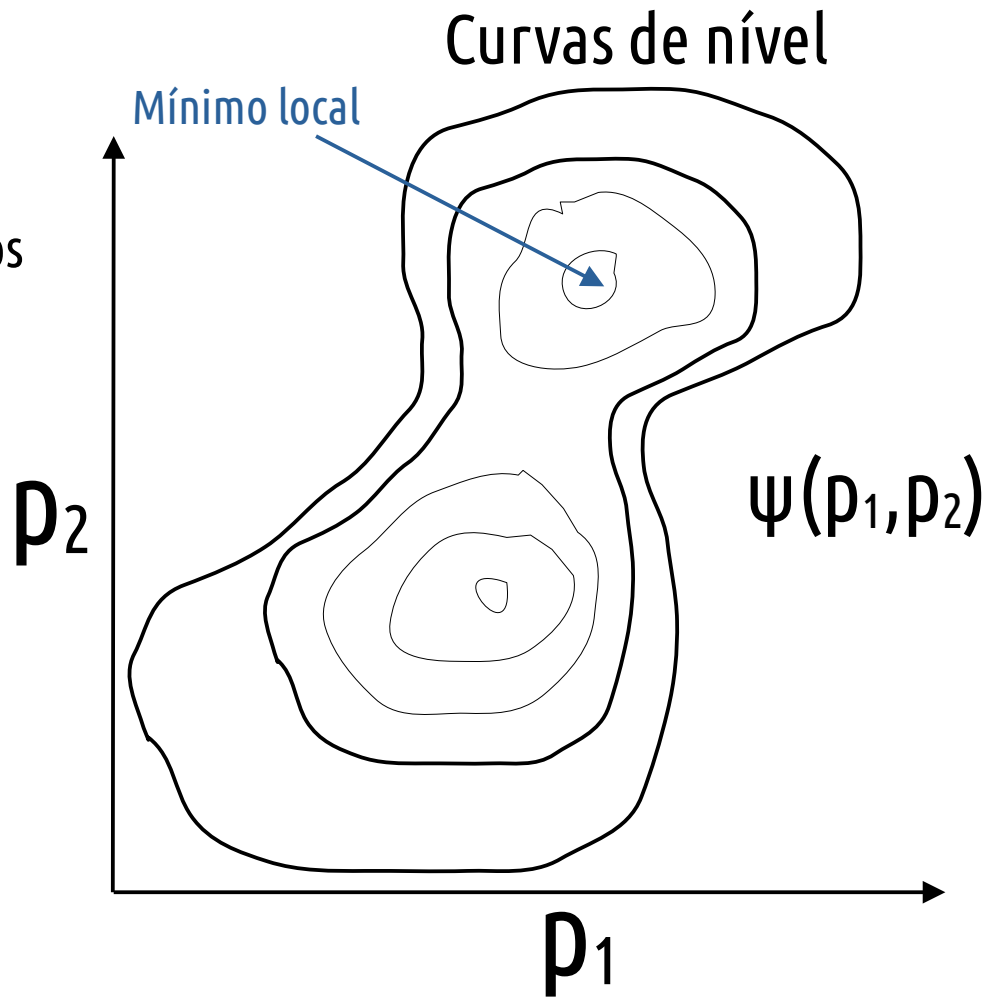
Minimizar a norma euclidiana entre o vetor de dados observados e o vetor de dados preditos!

$$\varphi(\mathbf{p}) = \| \mathbf{d}^o - \mathbf{d}^p \|_2$$

Função de ajuste

$$\nabla \varphi(\mathbf{p}) = \mathbf{0}$$

Gradiente da função de ajuste



Introdução ao problema não-linear

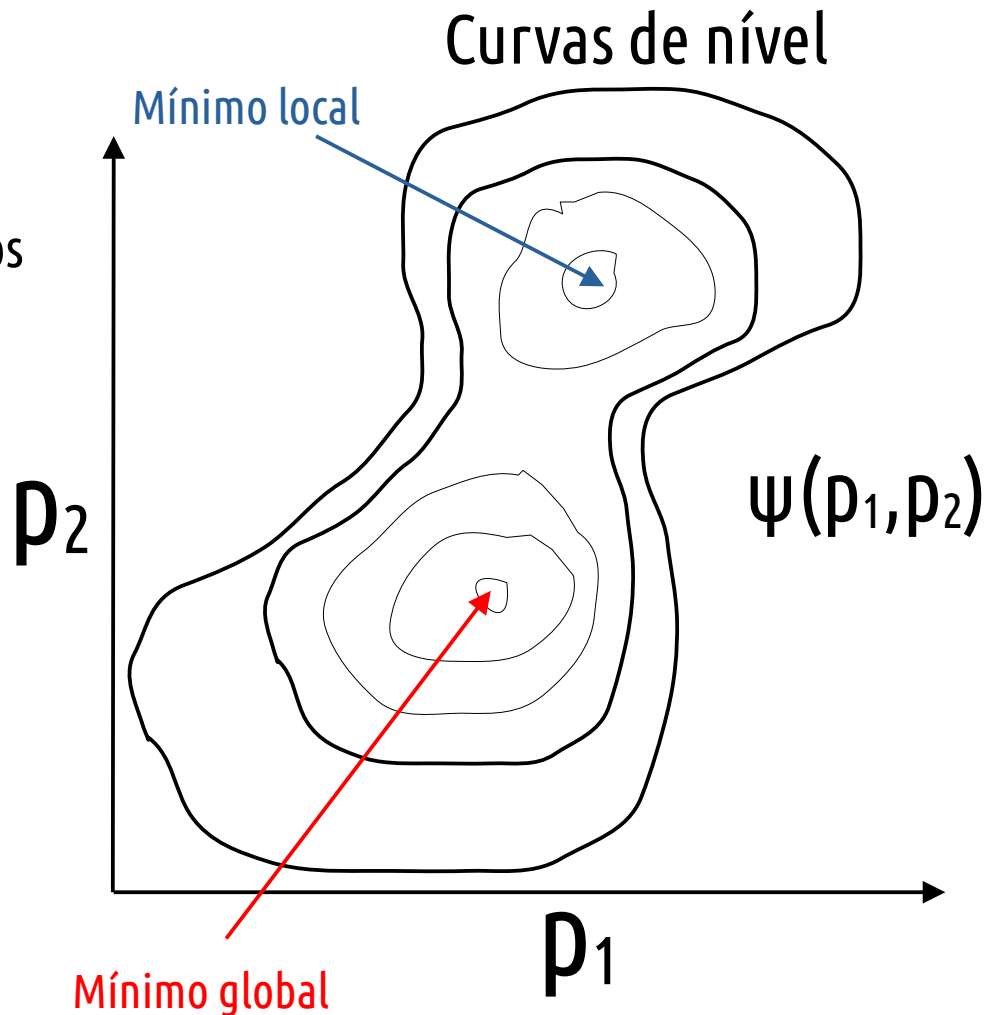
Minimizar a norma euclidiana entre o vetor de dados observados e o vetor de dados preditos!

$$\varphi(\mathbf{p}) = \| \mathbf{d}^o - \mathbf{d}^p \|_2$$

Função de ajuste

$$\nabla \varphi(\mathbf{p}) = \mathbf{0}$$

Gradiente da função de ajuste



Introdução ao problema não-linear

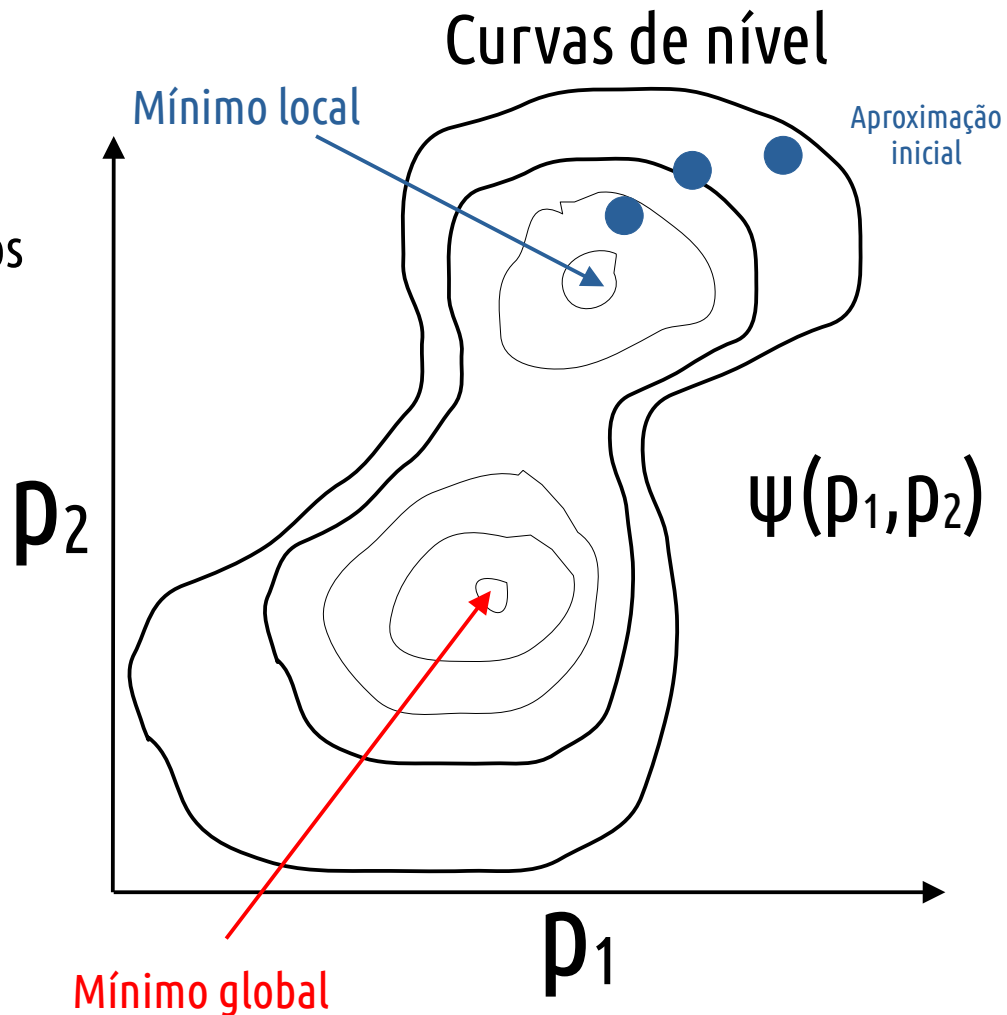
Minimizar a norma euclidiana entre o vetor de dados observados e o vetor de dados preditos!

$$\varphi(\mathbf{p}) = \| \mathbf{d}^o - \mathbf{d}^p \|_2$$

Função de ajuste

$$\nabla \varphi(\mathbf{p}) = \mathbf{0}$$

Gradiente da função de ajuste



Introdução ao problema não-linear

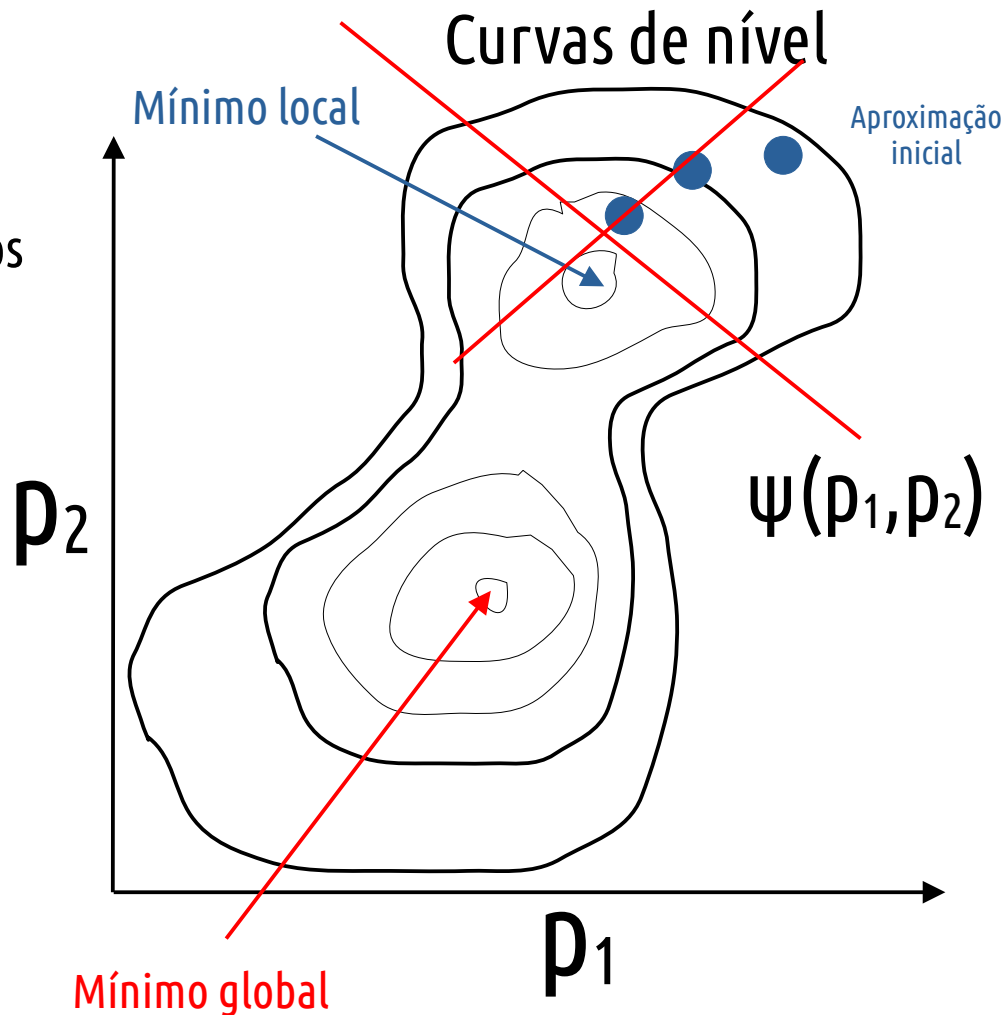
Minimizar a norma euclidiana entre o vetor de dados observados e o vetor de dados preditos!

$$\varphi(\mathbf{p}) = \| \mathbf{d}^o - \mathbf{d}^p \|_2$$

Função de ajuste

$$\nabla \varphi(\mathbf{p}) = \mathbf{0}$$

Gradiente da função de ajuste



Introdução ao problema não-linear

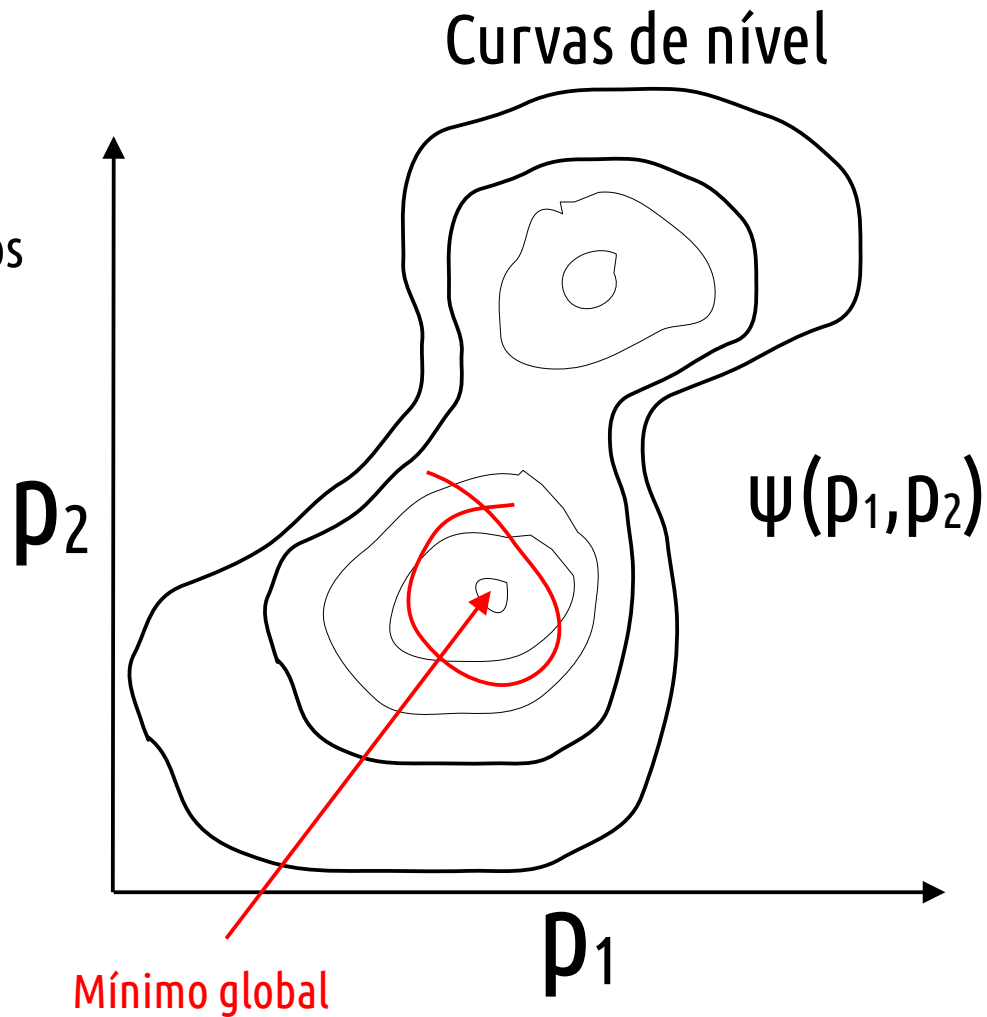
Minimizar a norma euclidiana entre o vetor de dados observados e o vetor de dados preditos!

$$\varphi(\mathbf{p}) = \| \mathbf{d}^o - \mathbf{d}^p \|_2$$

Função de ajuste

$$\nabla \varphi(\mathbf{p}) = \mathbf{0}$$

Gradiente da função de ajuste



Introdução ao problema não-linear

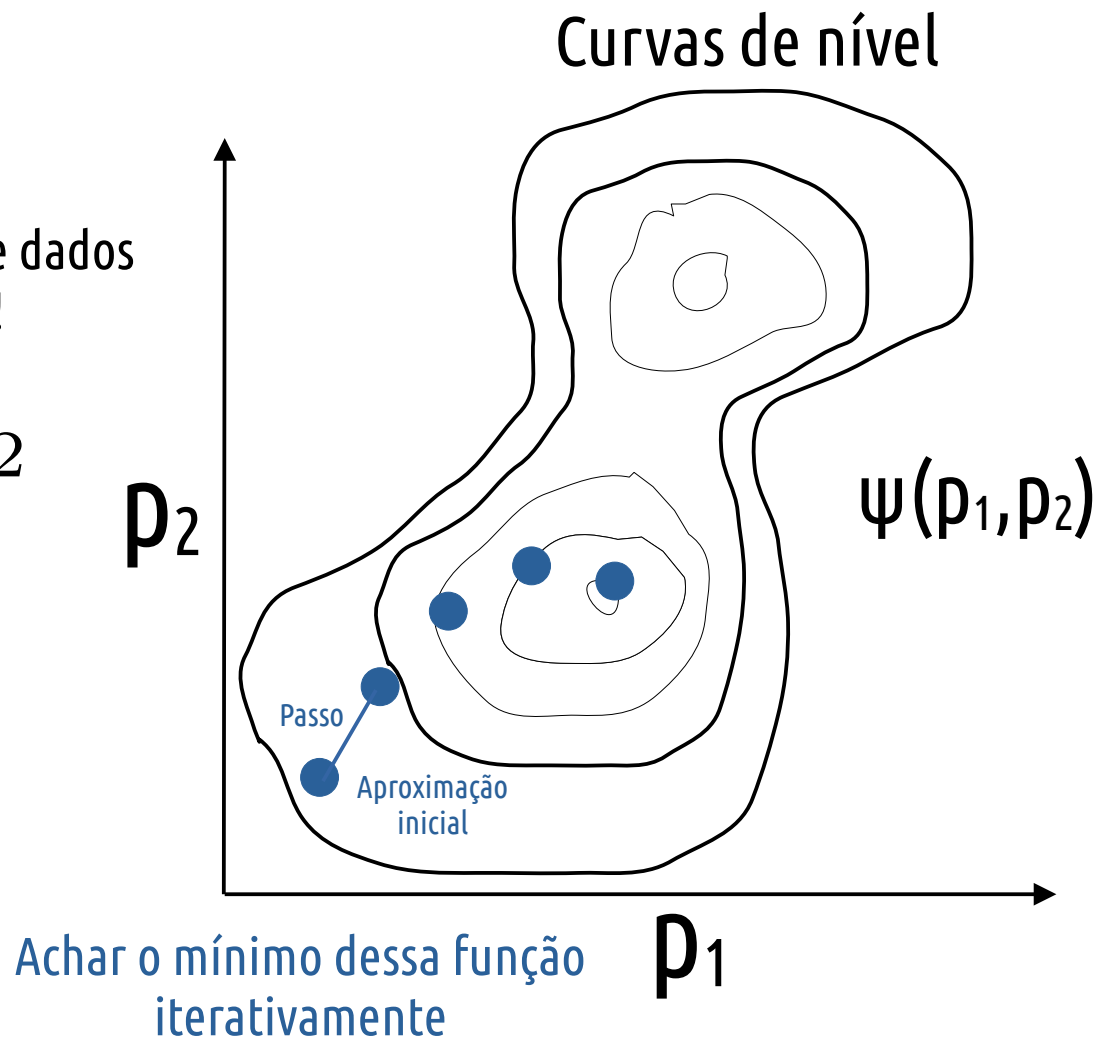
Minimizar a norma euclidiana entre o vetor de dados observados e o vetor de dados preditos!

$$\varphi(\mathbf{p}) = \| \mathbf{d}^o - \mathbf{d}^p \|_2$$

Função de ajuste

$$\nabla \varphi(\mathbf{p}) = \mathbf{0}$$

Gradiente da função de ajuste

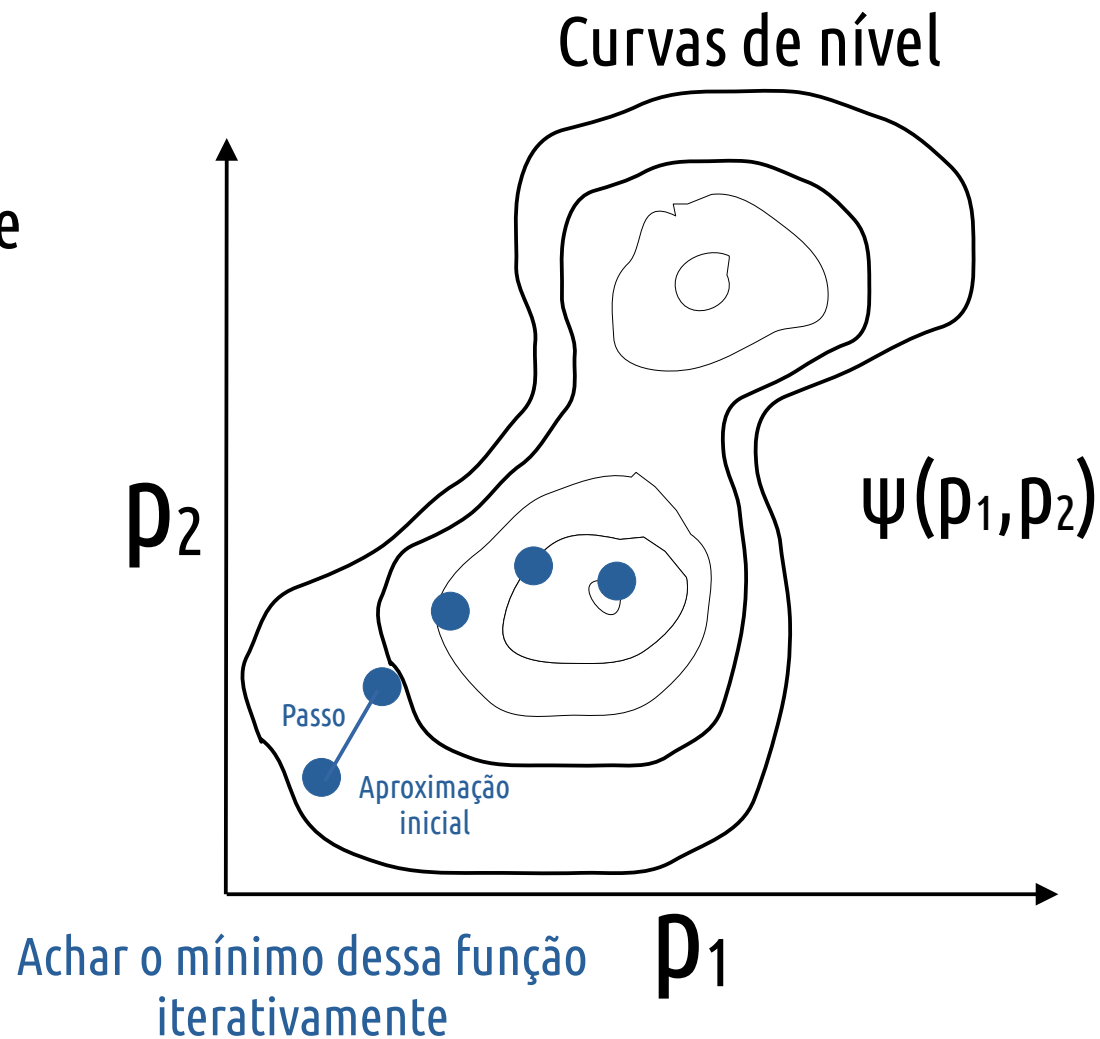


Introdução ao problema não-linear

Existem dois tipos de métodos que minimizam estas funções:

Métodos por gradiente
(Determinísticos)

Métodos Heurísticos



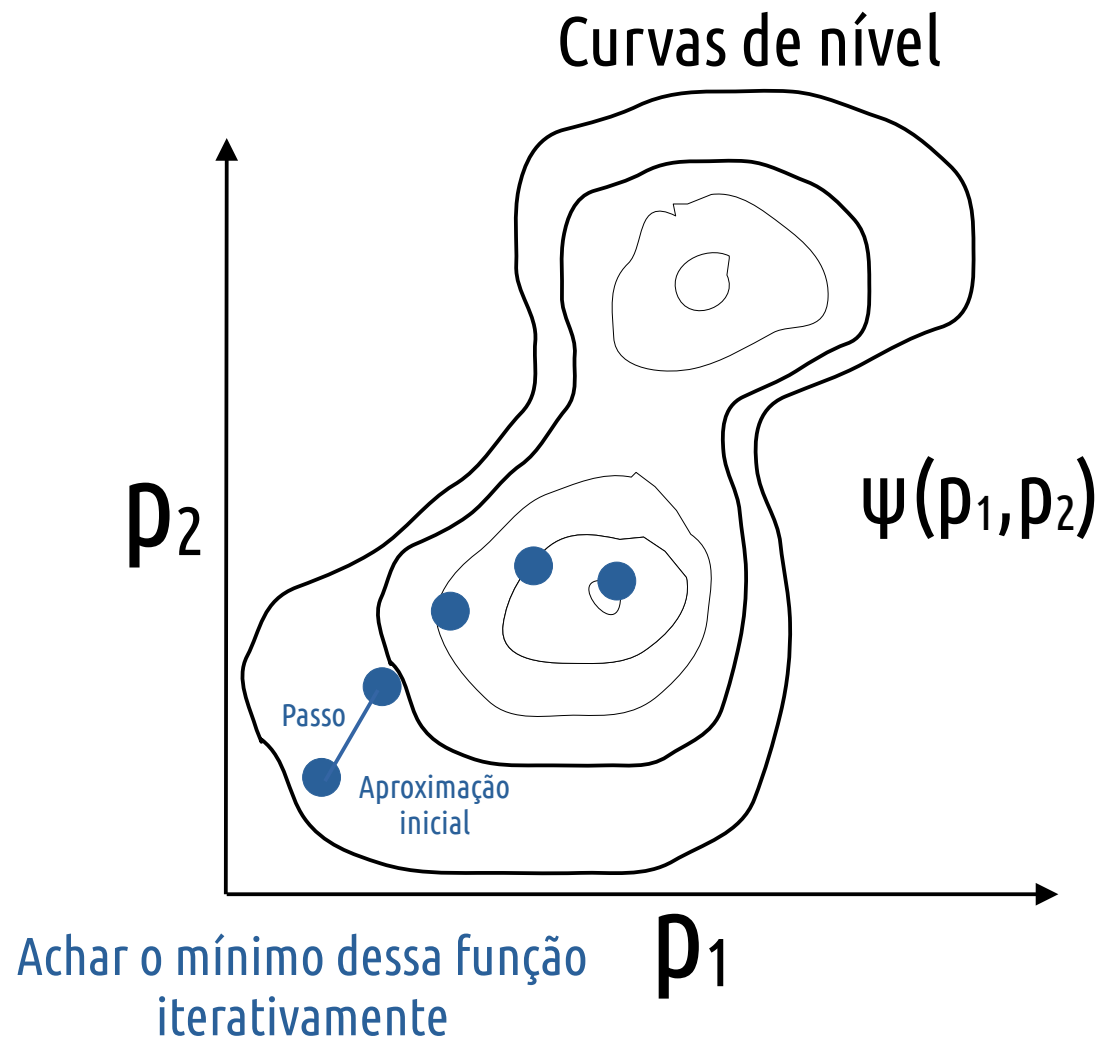
Introdução ao problema não-linear

Métodos por gradiente :

- Steepest descent
- Newton
- Gauss-Newton
- Levenberg-Marquardt

Métodos Heurísticos :

- Simulated Annealing
- Ant Colony
- Algoritmo Genético



Métodos determinísticos

Matematicamente, este é um processo de otimização no qual queremos **minimizar a norma euclidiana** de uma função que mede a distância entre **os dois vetores!**

$$\varphi(\mathbf{p}) = \| \mathbf{d}^o - \mathbf{d}^p \|_2 = [\mathbf{d}^o - \mathbf{d}^p]^T [\mathbf{d}^o - \mathbf{d}^p]$$

Função de ajuste

$$\nabla \varphi(\mathbf{p}) = \mathbf{0}$$

Quando falamos em minimizar uma função, significa que queremos tomar o gradiente desta função e igualarmos a zero.

Métodos determinísticos

Primeiro expandindo a função de ajuste em **série de Taylor (até segunda ordem)** teremos:

$$\varphi(\mathbf{p}_0 + \Delta \mathbf{p}) \approx \varphi(\mathbf{p}_0) + \Delta \mathbf{p}^T \mathbf{J}(\mathbf{p}_0) + \frac{1}{2} \Delta \mathbf{p}^T \mathbf{H}(\mathbf{p}_0) \Delta \mathbf{p}$$

Encontraremos que **o sistema linear que teremos que resolver** será igual a :

$$\mathbf{H}(\mathbf{p}^k) \Delta \mathbf{p}^k = -\mathbf{J}(\mathbf{p}^k)$$

Resolveremos este sistema linear a cada k-ésima iteração!

Métodos determinísticos

Minimizar a norma euclidiana entre o vetor de dados observados e o vetor de dados preditos!

$$\varphi(\mathbf{p}) = \| \mathbf{d}^o - \mathbf{d}^p \|_2$$

Função de ajuste

$$\nabla \varphi(\mathbf{p}) = \mathbf{0}$$

Gradiente da função de ajuste

Problema linear

$$\mathbf{d}^p = \mathbf{G}\mathbf{p}$$

$$(\mathbf{G}^T \mathbf{G}) \mathbf{p}^\# = \mathbf{G}^T \mathbf{d}^o$$

Estimador de mínimos quadrados

Problema não-linear

$$\mathbf{d}^p \neq \mathbf{G}\mathbf{p}$$

$$\mathbf{H}(\mathbf{p}^k) \Delta \mathbf{p}^k = -\mathbf{J}(\mathbf{p}^k)$$

Sistema de equações

Diferença entre métodos

Problema não-linear

$$\mathbf{d}^p \neq \mathbf{G}\mathbf{p}$$

$$\mathbf{H}(\mathbf{p}^k) \Delta \mathbf{p}^k = -\mathbf{J}(\mathbf{p}^k)$$

Sistema de equações

Diferença entre métodos

Problema não-linear

$$\mathbf{d}^p \neq \mathbf{G}\mathbf{p}$$

$$\mathbf{H}(\mathbf{p}^k) \Delta \mathbf{p}^k = -\mathbf{J}(\mathbf{p}^k)$$

Sistema de equações



O que irá diferenciar os métodos é a aproximação da matriz Hessiana!

Diferença entre métodos

Problema não-linear

$$\mathbf{d}^p \neq \mathbf{G}\mathbf{p}$$

$$\mathbf{H}(\mathbf{p}^k) \Delta \mathbf{p}^k = -\mathbf{J}(\mathbf{p}^k)$$

Sistema de equações



O que irá diferenciar os métodos é a aproximação da matriz Hessiana!

Steepest Descent : $1/\eta$

Newton : $\mathbf{H}(\mathbf{p}^k)$

Gauss-Newton : $\mathbf{J}(\mathbf{p}^k)^T \mathbf{J}(\mathbf{p}^k)$

Levenberg-Marquardt : $\mathbf{J}(\mathbf{p}^k)^T \mathbf{J}(\mathbf{p}^k) + \lambda \mathbf{I}$

Diferença entre métodos

Problema não-linear

$$\mathbf{d}^p \neq \mathbf{G}\mathbf{p}$$

$$\mathbf{H}(\mathbf{p}^k)\Delta\mathbf{p}^k = -\mathbf{J}(\mathbf{p}^k)$$

Sistema de equações

Método	Convergência
Steepest Descent	0
Levenberg-Marquardt	1
Gauss-Newton	2
Newton	3

Lenta = 0

Rápida = 3

Diferença entre métodos

Problema não-linear

$$\mathbf{d}^p \neq \mathbf{G}\mathbf{p}$$

$$\mathbf{H}(\mathbf{p}^k) \Delta \mathbf{p}^k = -\mathbf{J}(\mathbf{p}^k)$$

Sistema de equações

Método	Chute inicial
Steepest Descent	Pode ser longe
Levenberg-Marquardt	Pode ser longe
Gauss-Newton	Deve ser próxima
Newton	Deve ser próxima

Diferença entre métodos

Problema não-linear

$$\mathbf{d}^p \neq \mathbf{G}\mathbf{p}$$

$$\mathbf{H}(\mathbf{p}^k) \Delta \mathbf{p}^k = -\mathbf{J}(\mathbf{p}^k)$$

Sistema de equações

Método	Passo
Steepest Descent	Depende de um parâmetro e do gradiente
Levenberg-Marquardt	Depende do gradiente, da Hessiana e de um parâmetro
Gauss-Newton	Depende do gradiente e da Hessiana
Newton	Depende do gradiente e da Hessiana

Diferença entre métodos

Problema não-linear

$$\mathbf{d}^p \neq \mathbf{G}\mathbf{p}$$

$$\mathbf{H}(\mathbf{p}^k)\Delta\mathbf{p}^k = -\mathbf{J}(\mathbf{p}^k)$$

Sistema de equações

Método	Custo computacional
Steepest Descent	0
Levenberg-Marquardt	2
Gauss-Newton	1
Newton	3

Baixo = 0

Alto = 3

Como regularizar um problema inverso não-linear? E os exemplos?

Até breve!