Aprendizado por Reforço

Aproximação estocástica – parte 2

Descida do Gradiente Estocástico

- Os algoritmos de descida do gradiente estocástico (SGD – do inglês *stochastic gradient descent*) são amplamente utilizados em aprendizado de máquina.
- Estimativa da média → caso especial do SGD → caso especial do RM.
- Considere o problema de otimização

$$\min_{w} J(w), \qquad J(w) = \mathbb{E}_X[f(w, X)]$$

  - w: parâmetro a ser otimizado
  - X: variável aleatória
  - $f(\cdot)$: escalar

Descida do Gradiente Estocástico

- Solução: algoritmo da descida do gradiente (GD – do inglês *gradient descent*)

$$w_{k+1} = w_k - \alpha_k \nabla_w J(w_k)$$

$$J(w) = \mathbb{E}_X[f(w,X)] \rightarrow \nabla_w J(w) = \nabla_w \mathbb{E}_X[f(w,X)] = \mathbb{E}_X[\nabla_w f(w,X)]$$

$$\boxed{w_{k+1} = w_k - \alpha_k \mathbb{E}_X[\nabla_w f(w_k,X)]}$$

- O GD exige $\mathbb{E}_X[\nabla_w f(w,X)]$. (**Como calcular?**)

Descida do Gradiente Estocástico

- O cálculo de $\mathbb{E}_X[\nabla_w f(w, X)]$ pode ser feito:

1. Utilizando a distribuição de X, mas ela é **desconhecida** na prática.
2. Utilizando n amostras i.i.d. de X e fazer uso da aproximação:

$$\mathbb{E}_X[\nabla_w f(w, X)] \approx \frac{1}{n}\sum_{i=1}^{n} \nabla_w f(w_k, x_i)$$

E dessa forma

$$w_{k+1} = w_k - \alpha_k \left(\frac{1}{n}\sum_{i=1}^{n} \nabla_w f(w_k, x_i)\right)$$

- **Qual o problema da formulação acima?**

Descida do Gradiente Estocástico

- Problema
  - Requer todas as amostras por iteração.
- Solução
  - Descida do gradiente estocástico (SGD) → atualização de w a cada amostra x.

$$w_{k+1} = w_k - \alpha_k \nabla_w f(w_k, x_k)$$

- GD vs. SGD
  - GD: gradiente verdadeiro $\mathbb{E}_X[\nabla_w f(w_k, X)]$.
  - SGD: gradiente estocástico $\nabla_w f(w_k, x_k)$.
    - **Ainda podemos assegurar que $\boldsymbol{w_k \to w^*}$ quando $\boldsymbol{k \to \infty}$?**

Descida do Gradiente Estocástico

- Observe que:

$$\nabla_w f(w_k, x_k) = \mathbb{E}[\nabla_w f(w_k, X)] + \underbrace{(\nabla_w f(w_k, x_k) - \mathbb{E}[\nabla_w f(w_k, X)])}$$

$$\nabla_w f(w_k, x_k) = \mathbb{E}[\nabla_w f(w_k, X)] + \eta_k$$

Descida do Gradiente Estocástico

- O algoritmo do SGD pode ser reescrito como:

$$w_{k+1} = w_k - \alpha_k \nabla_w f(w_k, x_k)$$

$$w_{k+1} = w_k - \alpha_k (\mathbb{E}[\nabla_w f(w_k, X)] + \eta_k)$$

$$\boxed{w_{k+1} = w_k - \alpha_k \mathbb{E}[\nabla_w f(w_k, X)] - \alpha_k \eta_k}$$

  - Fórmula é idêntica ao GD + um termo de perturbação.

Descida do Gradiente Estocástico

- η_k tem média zero $\longrightarrow$ intuitivamente indica que pode não comprometer a convergência.

$$\mathbb{E}[\eta_k] = \mathbb{E}\left[\nabla_w f(w_k, x_k) - \mathbb{E}[\nabla_w f(w_k, X)]\right]$$

$$\mathbb{E}[\eta_k] = \mathbb{E}[\nabla_w f(w_k, x_k)] - \mathbb{E}\left[\mathbb{E}[\nabla_w f(w_k, X)]\right]$$

$$\mathbb{E}[\eta_k] = \mathbb{E}_{x_k}[\nabla_w f(w_k, x_k)] - \mathbb{E}_X[\nabla_w f(w_k, X)]$$

$\{x_k\}$ é i.i.d.

$$\mathbb{E}[\eta_k] = \mathbb{E}_X[\nabla_w f(w_k, x_k)] - \mathbb{E}_X[\nabla_w f(w_k, X)]$$

$$\boxed{\mathbb{E}[\eta_k] = 0}$$

Descida do Gradiente Estocástico

- Problema de estimação da Média como um problema de otimização

$$\min_{w} J(w), \qquad J(w) = \mathbb{E}[f(w, X)]$$

  - $f(w, X) = \frac{1}{2}\|w - X\|^2$
  - $\nabla_w f(w_k, X) = w - X$
  - A solução ótima para $\nabla_w J(w) = 0$ é $w^* = \mathbb{E}[X]$.

Descida do Gradiente Estocástico

- Problema de estimação da Média como um problema de otimização
  - Algoritmo GD

$$w_{k+1} = w_k - \alpha_k \nabla_w J(w)$$

$$w_{k+1} = w_k - \alpha_k \mathbb{E}[\nabla_w f(w_k, X)]$$

$$w_{k+1} = w_k - \alpha_k \boxed{\mathbb{E}[w_k - X]}$$

Não aplicável pois $\mathbb{E}[X]$ é desconhecido.

  - Algoritmo SGD

$$w_{k+1} = w_k - \alpha_k \nabla_w f(w_k, x_k)$$

$$w_{k+1} = w_k - \alpha_k (w_k - x_k)$$

- Estimativa da média incremental → caso especial do SGD, isto é, um SGD projetado para estimar a média.

Descida do Gradiente Estocástico

- Padrão de convergência do SGD
  - Erro relativo:

Gradiente estocástico

Gradiente verdadeiro

$$\delta_k = \frac{|\nabla_w f(w, x_k) - \mathbb{E}[\nabla_w f(w_k, X)]|}{|\mathbb{E}[\nabla_w f(w_k, X)]|}$$

  - Longe de w^*: trajetória se comporta com GD → convergência rápida.
  - Perto de w^*: convergência mais aleatória.

Descida do Gradiente Estocástico

- Exemplo: Estimativa da Média

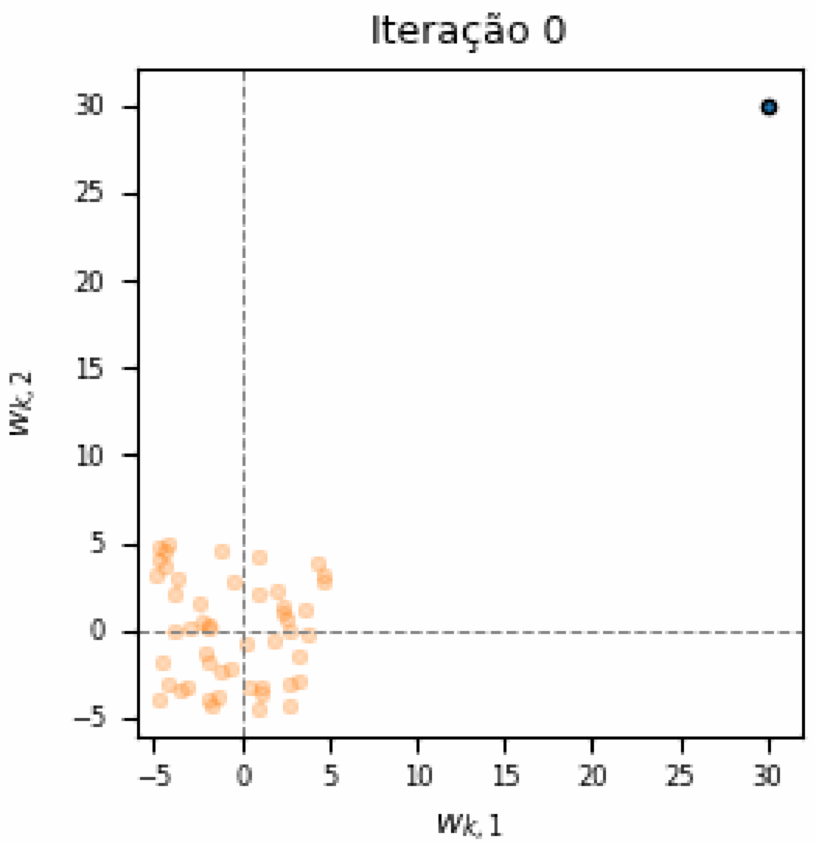
$$\min_w J(w) = \mathbb{E}\left[\frac{1}{2}\|w - X\|^2\right] = \mathbb{E}[f(w, X)]$$

  - $\nabla_w f(w, x_k) = w - x_k$
  - $\mathbb{E}[\nabla_w f(w, X)] = \mathbb{E}[w - X] = w - \mathbb{E}[X] = w - w^*$

- Erro relativo (δ_k):

$$\delta_k = \frac{|\nabla_w f(w, x_k) - \mathbb{E}[\nabla_w f(w_k, X)]|}{|\mathbb{E}[\nabla_w f(w_k, X)]|} = \frac{|(w - x_k) - (w - \mathbb{E}[X])|}{|w - w^*|} = \frac{|\mathbb{E}[X] - x_k|}{|w - w^*|}$$

- w_k longe de w^*: erro pequeno $\longrightarrow$ SGD $\approx$ GD (rápido).
- w_k perto de w^*: erro pode ser maior $\longrightarrow$ trajetória mais aleatória.
- $\mathbb{E}[\delta_k] \propto Var[X]$



Descida do Gradiente Estocástico

- Formulação determinística do SGD
  - Considere o conjunto $\{x_i\}_{i=1}^{n}$ (não necessariamente amostras de uma variável aleatória).
  - Problema otimização

$$\min_{w} J(w), \qquad J(w) = \frac{1}{n}\sum_{i=1}^{n} f(w, x_i)$$

    - $f(w, x_i)$: função parametrizada por w
    - w: parâmetro a ser otimizado

Descida do Gradiente Estocástico

- Formulação determinística do SGD
  - Algoritmo GD

$$w_{k+1} = w_k - \alpha_k \nabla_w J(w_k) = w_k - \alpha_k \left(\frac{1}{n} \sum_{i=1}^{n} \nabla_w f(w, x_i) \right)$$

  - Formulação incremental (se n é grande)

$$w_{k+1} = w_k - \alpha_k \nabla_w J(w_k) = w_k - \alpha_k \nabla_w f(w, x_k)$$

    - x_k: elemento obtido no tempo k (não necessariamente o k-ésimo elemento).
    - **O algoritmo incremental acima é SGD?**
    - **Como usar o conjunto finito $\{x_i\}_{i=1}^{n}$? Usar na ordem ou amostrar aleatoriamente?**

  - **Como podemos converter a formulação determinística em uma formulação estocástica?**

Descida do Gradiente Estocástico

- Conversão: formulação determinística → formulação estocástica
  - Introdução de uma variável aleatória X definida sobre $\{x_i\}_{i=1}^n$ com distribuição uniforme:

$$p(X = x_i) = \frac{1}{n}$$

  - Novo problema otimização

$$\min_w J(w), \qquad J(w) = \frac{1}{n}\sum_{i=1}^{n} f(w, x_i) = \mathbb{E}[f(w, X)]$$

  - O algoritmo $w_{k+1} = w_k - \alpha_k \nabla_w f(w, x_k)$ é SGD se x_k for amostrado uniformemente e independentemente $\{x_i\}_{i=1}^n$ (o mesmo elemento pode ser repetido na amostragem).

Descida do Gradiente Estocástico

- Batch GD, SGD e Mini-batch GD
  - Problema de otimização

$$\min_{w} J(w), \qquad J(w) = \mathbb{E}[f(w, X)]$$

    - Dado o conjunto de amostras $\{x_i\}_{i=1}^{n}$ da variável aleatória X.

Descida do Gradiente Estocástico

- Batch GD, SGD e Mini-batch GD
  - Batch GD (BGD): usa **todas** as n amostras por iteração.

$$w_{k+1} = w_k - \alpha_k \left(\frac{1}{n} \sum_{i=1}^{n} \nabla_w f(w_k, x_i) \right)$$

  - Mini-batch GD (MBGD): usa m amostras por iteração ($m < n$).

$$w_{k+1} = w_k - \alpha_k \left(\frac{1}{m} \sum_{j \in \mathcal{J}_k} \nabla_w f(w_k, x_j) \right), \qquad \mathcal{J}_k \subseteq \{1, \dots, n\}, \qquad |\mathcal{J}_k| = m$$

  - SGD: usa 1 amostra por iteração.

$$w_{k+1} = w_k - \alpha_k \nabla_w f(w_k, x_k)$$

Descida do Gradiente Estocástico

- MBGD
  - Versão intermediária entre SGD e BGD.
  - Menos aleatório que SGD.
  - Mais flexível que BGD por usar menos amostras.
  - Se $m = 1$, então MBGD = SGD.
  - **Se $\boldsymbol{m = n}$, então MBGD = BGD?**

Descida do Gradiente Estocástico

- MBGD
  - Versão intermediária entre SGD e BGD.
  - Menos aleatório que SGD.
  - Mais flexível que BGD por usar menos amostras.
  - Se $m = 1$, então MBGD = SGD.
  - **Se $m = n$, então MBGD = BGD?**
    - Não necessariamente: MBGD usa amostras aleatórias, podendo repetir o mesmo elemento.
  - Convergência geralmente mais rápida que SGD.

Descida do Gradiente Estocástico

- Exemplo: Estimativa da Média
  - Dado um conjunto $\{x_i\}_{i=1}^{n}$, temos o seguinte problema de otimização:

$$\min_{w} J(w), \qquad J(w) = \frac{1}{2n}\sum_{i=1}^{n} \|w - x_i\|^2$$

- Solução ótima

$$w^* = \bar{x}$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Descida do Gradiente Estocástico

- Algoritmos para estimar a média
  - BGD:

$$w_{k+1} = w_k - \alpha_k \left(\frac{1}{n} \sum_{i=1}^{n} (w_k - x_i) \right) = w_k - \alpha_k (w_k - \bar{x})$$

  - MBGD

$$w_{k+1} = w_k - \alpha_k \left(\frac{1}{m} \sum_{j \in \mathcal{J}_k} (w_k - x_j) \right) = w_k - \alpha_k \left(w_k - \bar{x}_k^{(m)} \right), \qquad \bar{x}_k^{(m)} = \frac{1}{m} \sum_{j \in \mathcal{J}_k} x_j$$

  - SGD:

$$w_{k+1} = w_k - \alpha_k (w_k - x_k)$$

Descida do Gradiente Estocástico

- Algoritmos para estimar a média ($\alpha_k = 1/k$):
  - BGD:

$$w_{k+1} = \frac{1}{k}\sum_{j=1}^{k} \bar{x} = \bar{x}$$

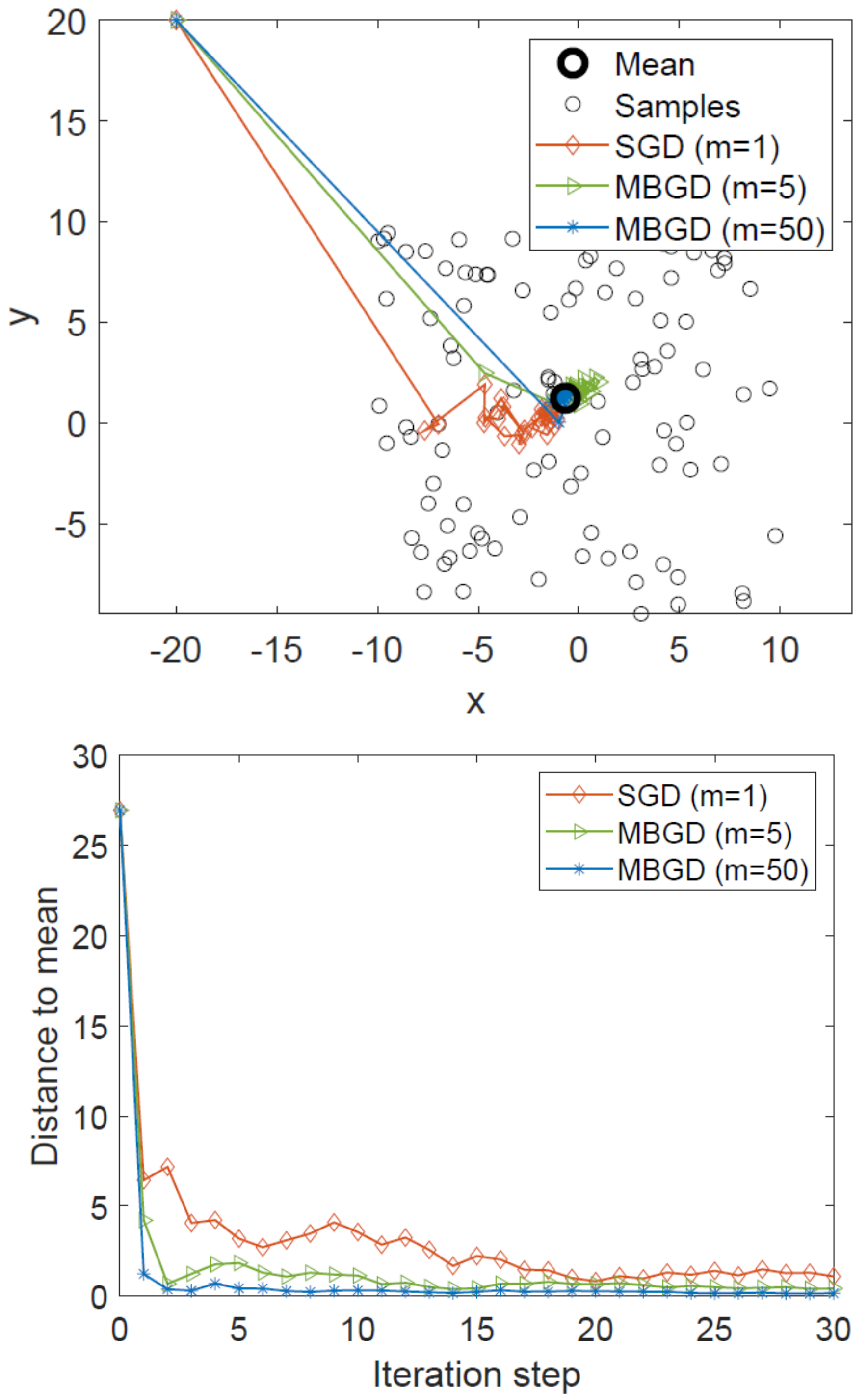
  - MBGD

$$w_{k+1} = \frac{1}{k}\sum_{j=1}^{k} \bar{x}_j^{(m)}$$

  - SGD:

$$w_{k+1} = \frac{1}{k}\sum_{j=1}^{k} x_j$$

Observações:
BGD: convergência imediata.
MBGD: convergência mais rápida do que SGD.



Revisão

- O que é aproximação estocástica?

- Por que precisamos estudar aproximação estocástica?

- Por que discutimos frequentemente o problema de estimação da média?

- Qual é a vantagem do algoritmo de RM sobre outros algoritmos para encontrar raízes?

Revisão

- O que é aproximação estocástica?
  - Aproximação estocástica refere-se a uma ampla classe de algoritmos iterativos estocásticos para resolver problemas de encontrar raízes ou de otimização.
- Por que precisamos estudar aproximação estocástica?
  - Porque os algoritmos de aprendizado por reforço com diferença temporal podem ser vistos como algoritmos de aproximação estocástica.
- Por que discutimos frequentemente o problema de estimação da média?
  - Porque os valores de estado e de ação são definidos como médias de variáveis aleatórias. Os algoritmos de aprendizado por diferença temporal são semelhantes a algoritmos de aproximação estocástica para estimação de média.
- Qual é a vantagem do algoritmo de RM sobre outros algoritmos para encontrar raízes?
  - O algoritmo RM é poderoso no sentido de que não requer a expressão da função objetivo nem de sua derivada. Como resultado, é um método de caixa-preta que exige apenas a entrada e a saída da função objetivo. O famoso algoritmo SGD é uma forma especial do algoritmo de RM.

Revisão

- Qual é a ideia básica do SGD?
- O SGD pode convergir rapidamente?
- O que é MBGD? Quais são suas vantagens em relação ao SGD e ao BGD?

Revisão

- Qual é a ideia básica do SGD?
  - O SGD tem como objetivo resolver problemas de otimização que envolvem variáveis aleatórias. Quando as distribuições de probabilidade das variáveis aleatórias fornecidas não são conhecidas, o SGD pode resolver os problemas de otimização apenas utilizando amostras. Matematicamente, o algoritmo SGD pode ser obtido substituindo o gradiente verdadeiro — expresso como um valor esperado — no algoritmo de descida do gradiente, por um **gradiente estocástico**.
- O SGD pode convergir rapidamente?
  - Padrão de convergência do SGD:
    - Se a estimativa estiver longe da solução ótima, então o processo de convergência é rápido.
    - Se a estimativa estiver próxima da solução, a aleatoriedade do gradiente estocástico torna-se significativa, e a taxa de convergência diminui.
- O que é MBGD? Quais são suas vantagens em relação ao SGD e ao BGD?
  - O MBGD pode ser visto como uma versão intermediária entre o SGD e o BGD.
  - Vantagens:
    - MBGD vs. SGD: MBGD tem menos aleatoriedade, pois utiliza mais amostras ao invés de apenas uma.
    - MBGD vs. BGD: MBGD não exige o uso de todas as amostras, tornando-o mais flexível.

Referências

- S. Zhao. *Mathematical Foundations of Reinforcement Learning*. Springer Singapore, 2025. [capítulo 6]
  - **disponível em:** https://github.com/MathFoundationRL/Book-Mathematical-Foundation-of-Reinforcement-Learning

Slides construídos com base nos livros supracitados, os quais estão disponibilizados publicamente pelos seus respectivos autores.