Lern-Nugget: Kettenregel verstehen

🎯 Ziel

Die Kettenregel der Differentialrechnung verstehen - das Herzstück von Backpropagation in neuronalen Netzen!

📐 Konkrete Funktionen definieren

Gegeben: Eine zusammengesetzte Funktion f(x) = g(h(x))

Konkrete Beispiele:

- h(x) = 2x + 3 (eine Gerade)
- g(h) = 0.5h + 1 (eine andere Gerade, mit h als Argument)

Die zusammengesetzte Funktion: f(x) = g(h(x)) = g(2x + 3) = 0.5(2x + 3) + 1 = x + 2.5

🔍 Was ist f'(x)?

📝 **Notation:** f'(x) bedeutet "die Ableitung der Funktion f nach der Variable x"

- Andere Schreibweise: f'(x) = df/dx
- Gesprochen: "f-Strich von x" oder "die Ableitung von f nach x"

Direkte Ableitung: Da f(x) = x + 2.5, ist f'(x) = 1

Aber wie kommt man darauf mit der Kettenregel?

⛓️ Die Kettenregel Schritt für Schritt

Schritt 1: Die einzelnen Ableitungen

Ableitung der inneren Funktion h(x):

- h(x) = 2x + 3
- h'(x) = 2

Ableitung der äußeren Funktion g(h):

- g(h) = 0.5h + 1
- dg/dh = 0.5

Schritt 2: Die Kettenregel anwenden

Kettenregel: f'(x) = (dg/dh) · (dh/dx)

Einsetzen:

- dg/dh = 0.5
- dh/dx = 2

Ergebnis: $f'(x) = 0.5 \cdot 2 = 1$ ✓

☑ Verifikation durch direktes Ableiten:

- $y = g(h(x)) = 0.5(2x + 3) + 1 = x + 2.5$
- $dy/dx = 1$ ✓

💡 Intuitive Erklärung

Die Kettenregel sagt uns:

> "Die Änderungsrate der gesamten Funktion = (Änderungsrate der äußeren Funktion) × (Änderungsrate der inneren Funktion)"

In unserem Beispiel:

- Wenn x um 1 steigt → h steigt um 2 (wegen $dh/dx = 2$)
- Wenn h um 2 steigt → y steigt um $0.5 \times 2 = 1$ (wegen $dg/dh = 0.5$)
- **Gesamt:** Wenn x um 1 steigt → y steigt um 1

🧠 Warum ist das für neuronale Netze wichtig?

In neuronalen Netzen haben wir:

```
Input → Hidden Layer → Output → Loss
  x   →      h       →    ŷ    →  J
```

Die Kettenregel erlaubt uns:

- $\partial J/\partial x = (\partial J/\partial \hat{y}) \cdot (\partial \hat{y}/\partial h) \cdot (\partial h/\partial x)$
- **Backpropagation** = Kettenregel rückwärts durch das Netz!

📝 Zusammenfassung

Kettenregel in drei Formen:

1. **Klassisch:** $(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$

2. **Mit Variablen:** Wenn $y = f(u)$ und $u = g(x)$, dann $dy/dx = (dy/du) \cdot (du/dx)$

3. **Für Mehrfachabhängigkeiten:** Bei $z = f(x,y)$ mit $y = g(x)$ wird $\partial z/\partial x = (\partial z/\partial y) \cdot (dy/dx) + (\partial z/\partial x)$

Unser Fall: Eindimensionale Kette → **totale Ableitungen** (d) sind korrekt!

Kernidee: Änderungsraten multiplizieren sich entlang der Abhängigkeitskette!

🎯 Übung

Probiere selbst:

- $h(x) = 3x - 1$

- $g(h) = h^2$
- $f(x) = g(h(x)) = (3x - 1)^2$

Berechne f'(x) mit der Kettenregel:

- dh/dx = ?
- dg/dh = ?
- f'(x) = (dg/dh) · (dh/dx) = ?

Lösung:

- dh/dx = 3
- dg/dh = 2h = 2(3x - 1)
- f'(x) = 2(3x - 1) · 3 = 6(3x - 1) = 18x - 6

💡 **Tipp:** Die Kettenregel ist der Schlüssel zum Verständnis von Backpropagation. Wenn du sie hier verstehst, verstehst du auch, wie neuronale Netze lernen!