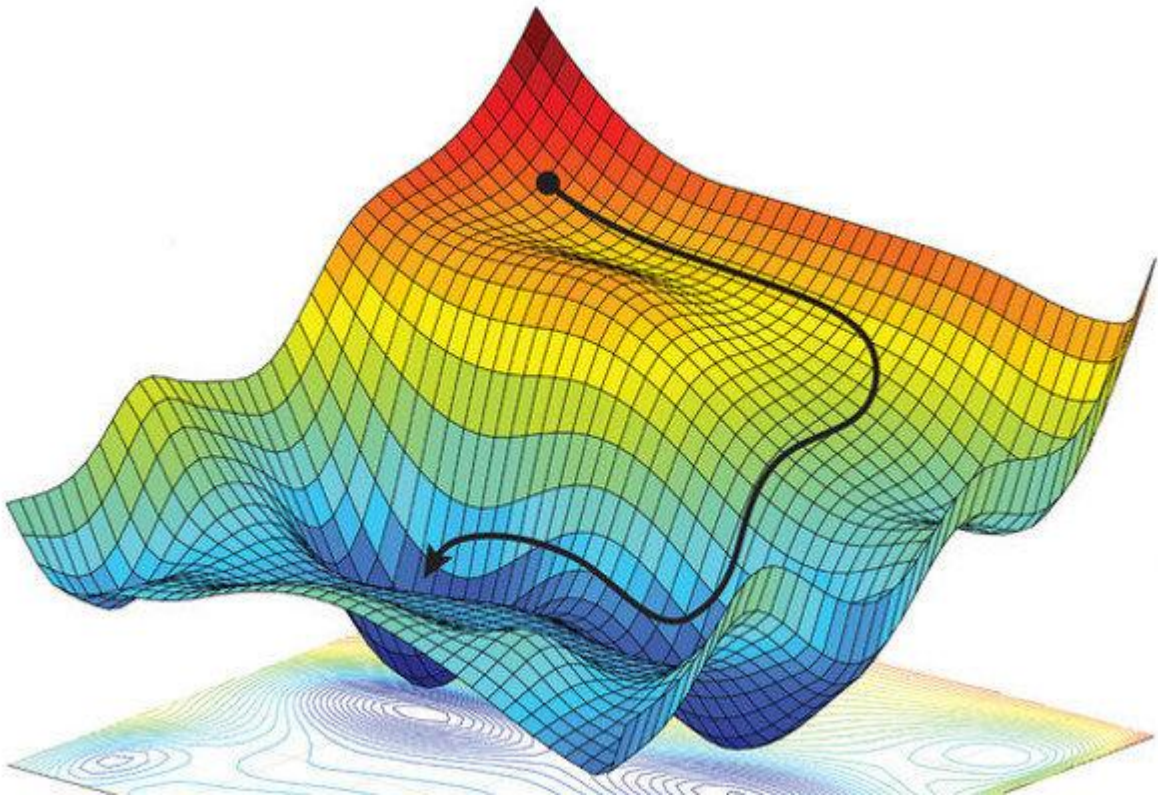
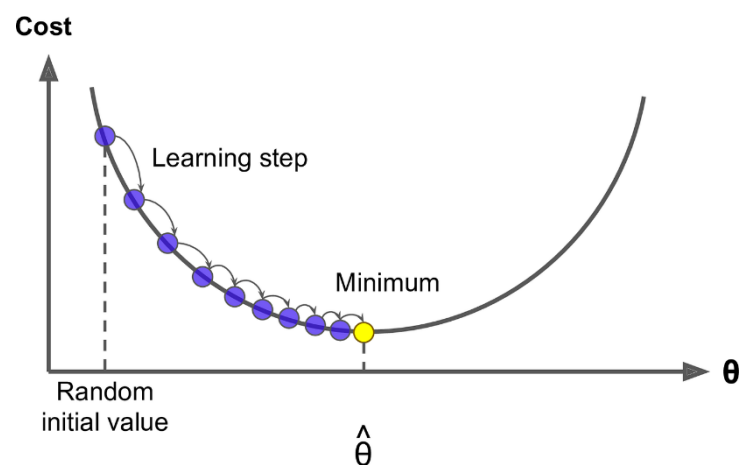
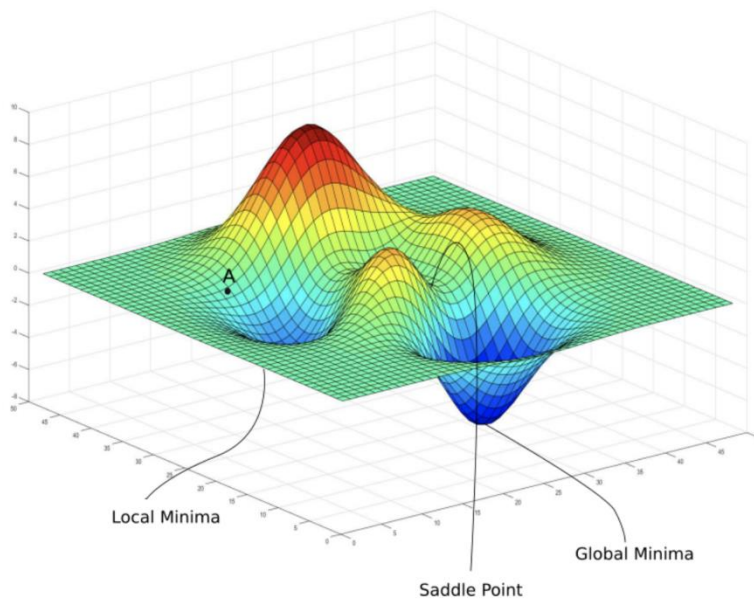


Loss



เราเทรนโมเดลเพื่อพยายามหา loss ที่ต่ำที่สุด

Learning rate and decay



Learning rate เป็นพารามิเตอร์ที่กำหนดว่าโมเดลควรปรับค่าน้ำหนักของเครือข่ายประสาทเทียมเท่าไรในแต่ละขั้นตอนการอัปเดตเพื่อลดค่า Loss function. หาก Learning rate สูงเกินไป, โมเดลอาจจะ "ข้าม" ค่า Loss ต่ำสุดไปได้ ในทางกลับกัน, หาก Learning

rate ต่ำเกินไป, การเรียนรู้ของโมเดลอาจช้ามากหรือติดอยู่ใน local minima ทำให้ไม่สามารถหาค่า Loss ต่ำสุดที่เป็น global minimum ได้ .

Decay เป็นการปรับลด Learning rate ลงเรื่อย ๆ ตามเงื่อนไขหรือกำหนดเวลาที่กำหนดไว้ การใช้ Decay สามารถช่วยให้โมเดล "ลงหลุม" ค่า Loss ที่ต่ำกว่าได้อย่างละเอียดยิ่งขึ้นเมื่อเข้าใกล้ค่าต่ำสุด (Learning rate คือเราจะให้โมเดลเรียนรู้เยอะหรือน้อยแค่ไหน ถ้าเยอะไปก็อาจจะทำให้ ไม่เจอ loss ที่ต่ำที่สุด แต่ถ้าน้อยเกินไปก็จะทำให้ขึ้นมาจากหลุมพรางไม่ได้ และ decay คือความเสื่อมถอย ที่จะไปลด lr ลงเรื่อย ๆ )