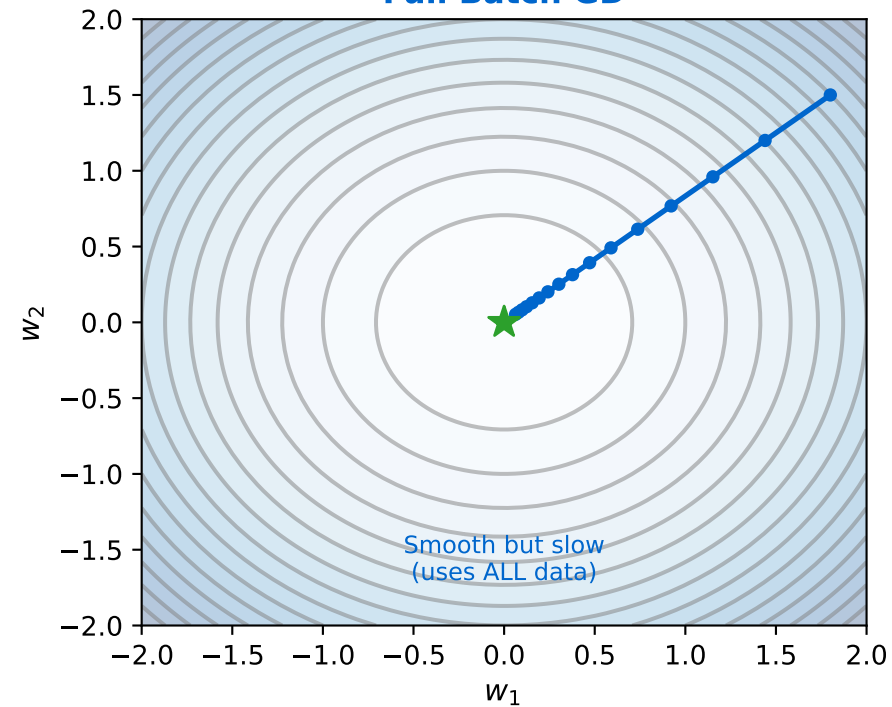
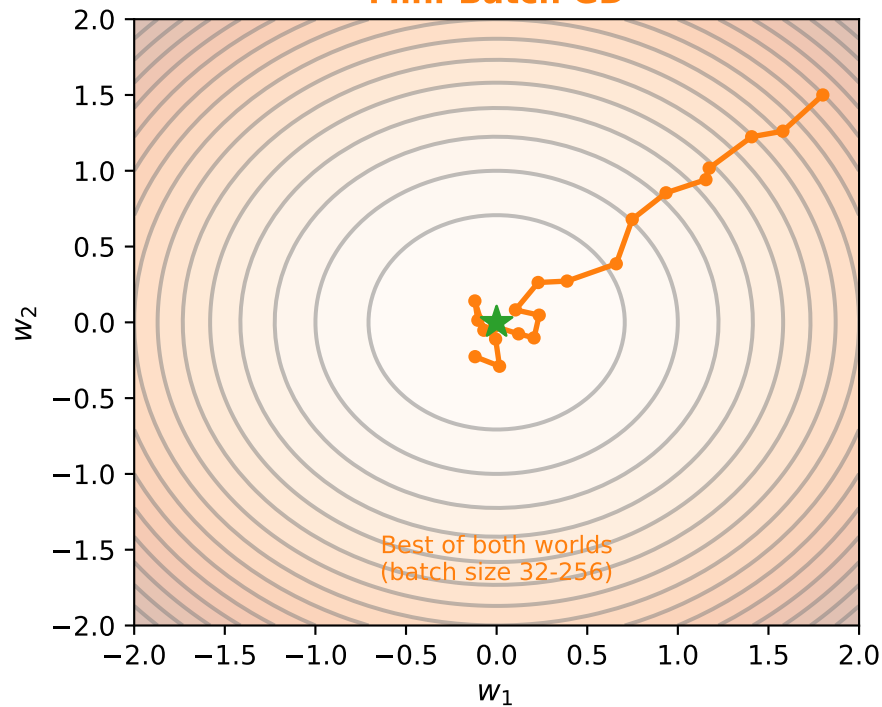


Gradient Descent Variants: Trading Off Speed vs Stability

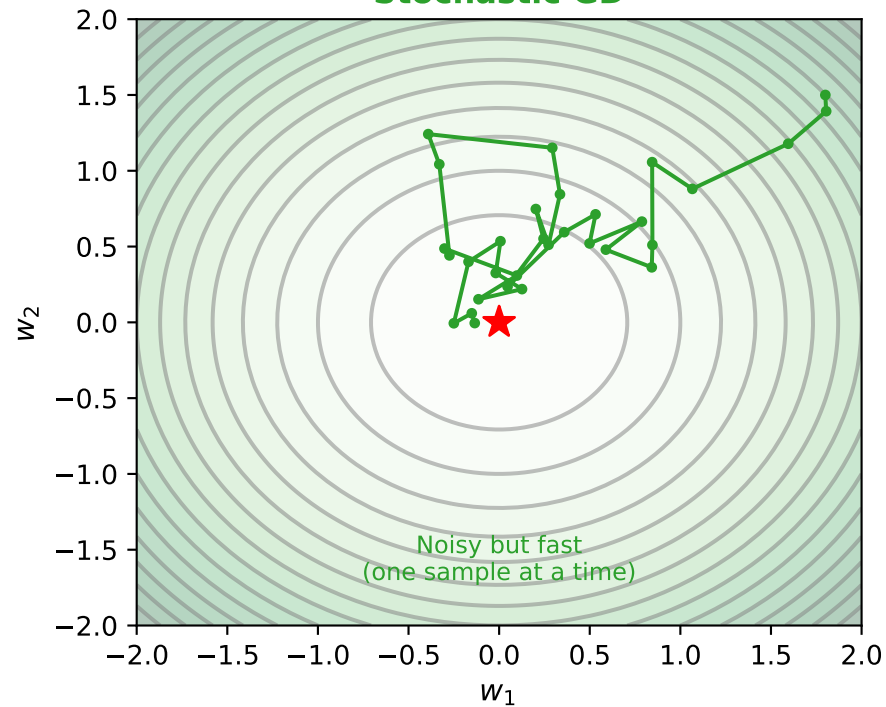
Full Batch GD



Mini-Batch GD



Stochastic GD



Full Batch: Stable gradients, slow updates | Mini-Batch: Good balance, standard choice | SGD: Fast updates, noisy gradients