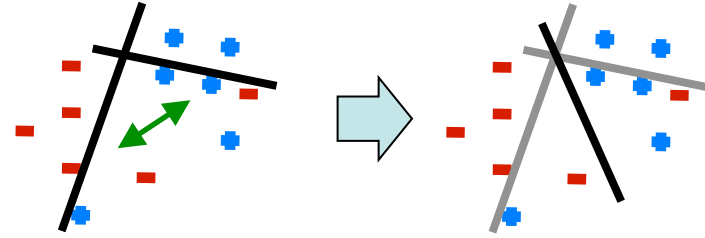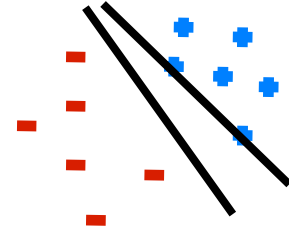# Noise: if the data isn't separable, weights might thrash

- Averaging weight vectors over time can help (averaged perceptron)

Mediocre generalization: finds a "barely" separating solution

Overtraining: test / held-out accuracy usually rises, then falls

- Overtraining is a kind of overfitting