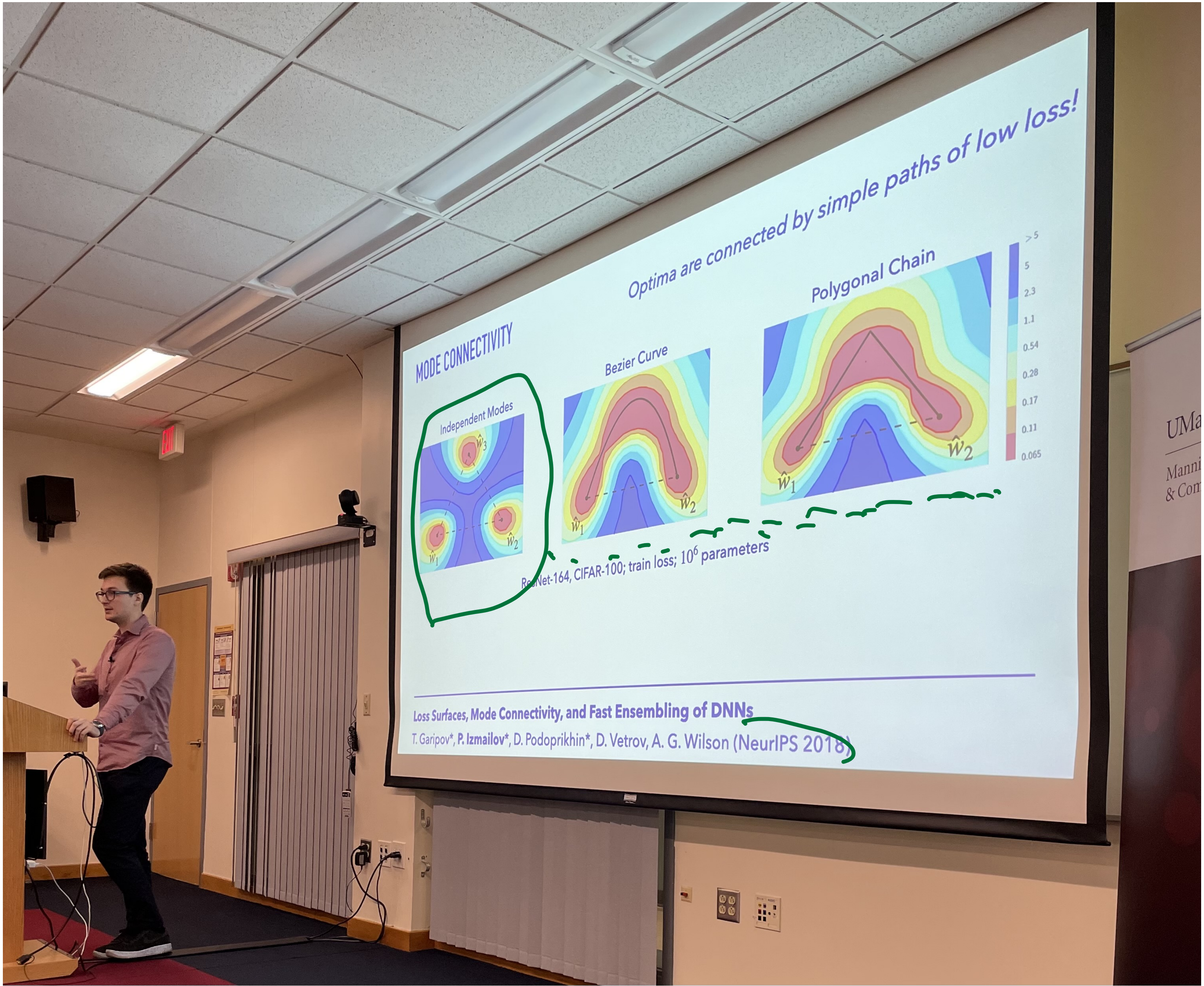
COMPSCI 689
Lecture 5: Numerical Optimization Foundations—Extra slides

Brendan O'Connor

College of Information and Computer Sciences
University of Massachusetts Amherst

Slides by Benjamin M. Marlin (marlin@cs.umass.edu).

Great talk yesterday!



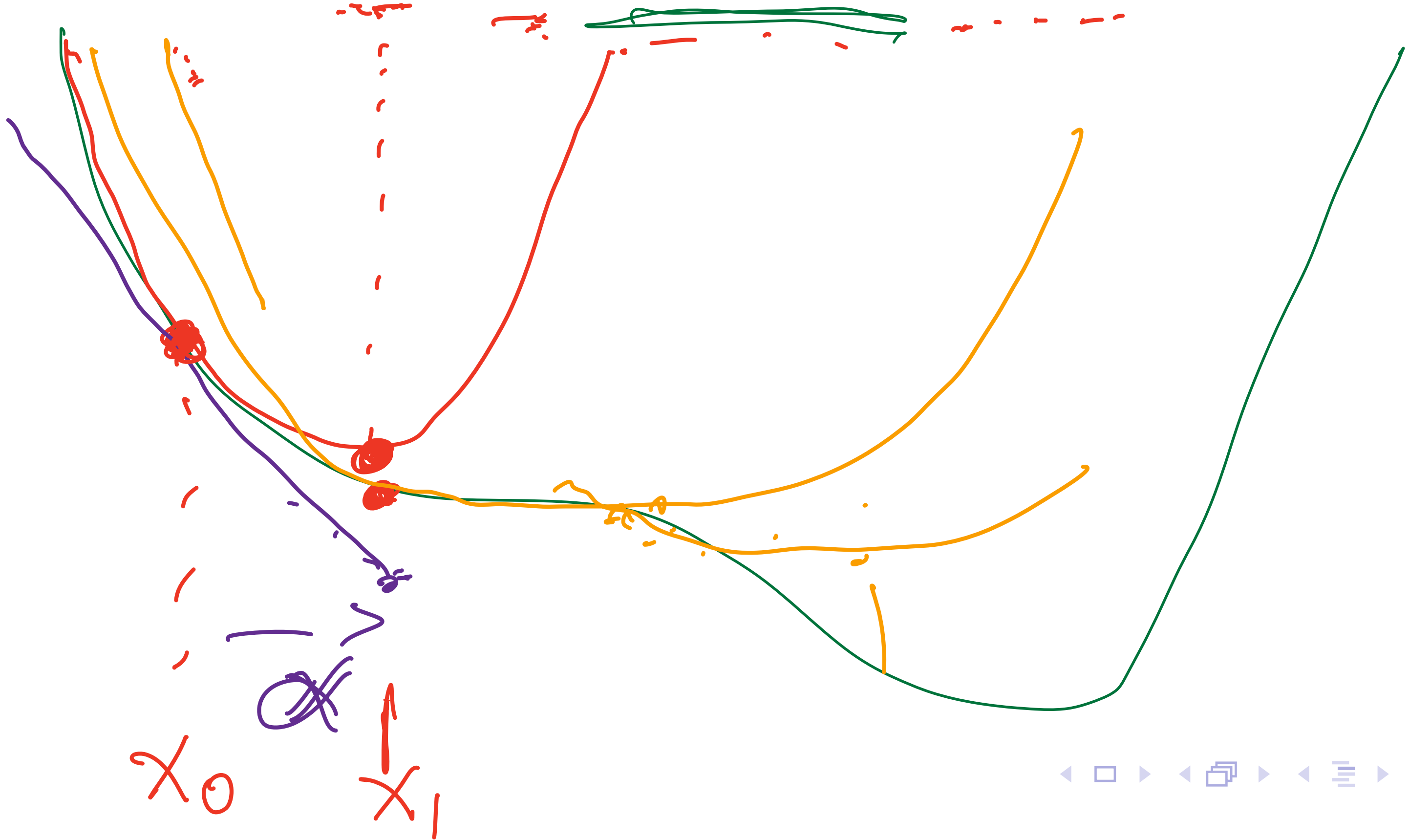
Newton-Raphson. Construct a 2nd order approx. at x_0: fit a parabola/bowl that matches the gradient and Hessian there

$$f(x) \approx f(x_0) + \nabla f(x_0)(x - x_0) + \nabla^2 f(x_0)(x - x_0)/2$$

If you believe the approximation, its minimum is at

$$x^* = x_0 - (\nabla^2 f(x_0))^{-1} \nabla f(x_0)$$

1D:



Newton-Raphson. Construct a 2nd order approx. at x_0: fit a parabola/bowl that matches the gradient and Hessian there

$$f(x) \approx f(x_0) + \nabla f(x_0)(x - x_0) + \nabla^2 f(x_0)(x - x_0)/2$$

If you believe the approximation, its minimum is at

$$x^* = x_0 - \overbrace{(\nabla^2 f(x_0))^{-1}}^{O(D^3)} \overbrace{\nabla f(x_0)}^{O(D^2)}$$

1D:

2D: Adjusts for curvature ($\approx$ how quickly gradient is changing)