Math Insight

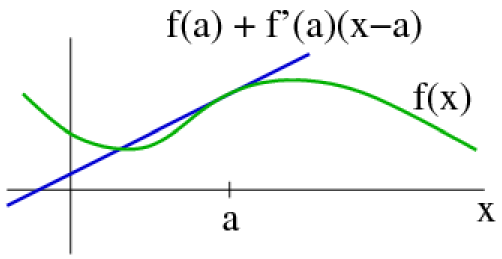
Introduction to Taylor's theorem for multivariable functions

Remember one-variable calculus Taylor's theorem. Given a one variable function $f(x)$, you can fit it with a polynomial around $x = a$.

For example, the best linear approximation for $f(x)$ is

$$f(x) \approx f(a) + f'(a)(x - a).$$

This linear approximation fits $f(x)$ (shown in green below) with a line (shown in blue) through $x = a$ that matches the slope of f at a.



We can add additional, higher-order terms, to approximate $f(x)$ better near a. The best quadratic approximation is

$$f(x) \approx f(a) + f'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2$$

We could add third-order or even higher-order terms:

$$f(x) \approx f(a) + f'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2 + \frac{1}{6}f'''(a)(x - a)^3 + \cdots .$$

The important point is that this *Taylor polynomial* approximates $f(x)$ well for x near a.

We want to generalize the Taylor polynomial to (scalar-valued) functions of multiple variables:

$$f(\mathbf{x}) = f(x_1, x_2, \ldots, x_n).$$

We already know the best linear approximation to f. It involves the derivative,

$$f(\mathbf{x}) \approx f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{x} - \mathbf{a}).$$

where $Df(\mathbf{a})$ is the matrix of partial derivatives. The linear approximation is the first-order Taylor polynomial.

What about the second-order Taylor polynomial? To find a quadratic approximation, we need to add quadratic terms to our linear approximation. For a function of one-variable $f(x)$, the quadratic term was

$$\frac{1}{2}f''(a)(x - a)^2.$$

For a function of multiple variables $f(\mathbf{x})$, what is analogous to the second derivative?

Since $f(\mathbf{x})$ is scalar, the first derivative is $Df(\mathbf{x})$, a $1 \times n$ matrix, which we can view as an n-dimensional vector-valued function of the n-dimensional vector $\mathbf{x}$. For the second derivative of $f(\mathbf{x})$, we can take the matrix of partial derivatives of the function $Df(\mathbf{x})$. We could write it as $DDf(\mathbf{x})$ for the moment. This second derivative matrix is an $n \times n$ matrix called the **Hessian matrix** of f. We'll denote it by $Hf(\mathbf{x})$,

$$Hf(\mathbf{x}) = DDf(\mathbf{x}).$$

When f is a function of multiple variables, the second derivative term in the Taylor series will use the Hessian $Hf(\mathbf{a})$. For the single-variable case, we could rewrite the quadratic expression as

$$\frac{1}{2}(x-a)f''(a)(x-a).$$

The analog of this expression for the multivariable case is

$$\frac{1}{2}(\mathbf{x}-\mathbf{a})^T Hf(\mathbf{a})(\mathbf{x}-\mathbf{a}).$$

We can add the above expression to our first-order Taylor polynomial to obtain the second-order Taylor polynomial for functions of multiple variables:

$$f(\mathbf{x}) \approx f(\mathbf{a}) + Df(\mathbf{a})(\mathbf{x}-\mathbf{a}) + \frac{1}{2}(\mathbf{x}-\mathbf{a})^T Hf(\mathbf{a})(\mathbf{x}-\mathbf{a}).$$

The second-order Taylor polynomial is a better approximation of $f(\mathbf{x})$ near $\mathbf{x} = \mathbf{a}$ than is the linear approximation (which is the same as the first-order Taylor polynomial). We'll be able to use it for things such as finding a local minimum or local maximum of the function $f(\mathbf{x})$.

You can read some examples here.

See also

Multivariable Taylor polynomial example

Cite this as

Nykamp DQ, "Introduction to Taylor's theorem for multivariable functions." From *Math Insight*. http://mathinsight.org/taylors_theorem_multivariable_introduction

Keywords: derivative, linear approximation, partial derivative, Taylor polynomial, Taylor's theorem