A Closer Look At The Hessian

Let's Build Some Intuition About Positive-Definiteness

Consider the 1D quadratic function

$$f(x) = f^* + g(x - x^*) + \frac{1}{2}(x - x^*)h(x - x^*)$$

We Can Perform A Similar Analysis For A Simple 2D Quadratic

Next, we consider the 2D quadratic function

$$f(x_1, x_2) = f^* + \frac{1}{2} \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

We Can Perform A Similar Analysis For A Simple 2D Quadratic (cont.)

The contours of this quadratic are aligned with the coordinate axes, because its Hessian is

$$\nabla^2 f = H = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix},$$

however, this is exceptionally rare.

Symmetric Matrices Have A Useful Decomposition

Theorem: Eigendecomposition for $H^T = H$

Every real symmetric matrix $H \in \mathbb{R}^{n \times n}$ can be decomposed as

$$H = Q \Lambda Q^T,$$

where $\Lambda = \mathrm{diag}(\lambda_i)$ is a diagonal matrix holding the eigenvalues of H and Q holds the corresponding orthogonal eigenvectors

Finally, Let's Consider A More General 2D Quadratic

Now we can consider the 2D quadratic function

$$f(x_1, x_2) = f^* + \frac{1}{2} \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Finally, Let's Consider A More General 2D Quadratic (cont.)

By rotating into the reference frame defined by the eigenvectors Q, the contours are aligned with the new coordinates $\hat{x}_1$ and $\hat{x}_2$.

Positive Definiteness Is Directly Related To The Eigenvalues

Bottom line: the important information we need to test for optimality are the eigenvalues of H.

Stationary Points Can Often Be Fully Characterized By The Hessian

H **is positive definite:** means $p^T Hp > 0$ for all p and $\lambda_i > 0 \; \forall i$

$\Rightarrow x^*$ is a local minimizer

H **is negative definite:** means $p^T Hp < 0$ for all p and $\lambda_i < 0 \; \forall i$

$\Rightarrow x^*$ is a local maximizer

H **is indefinite:** means $p^T Hp > 0$ and $r^T Hr < 0$ for some p, r and $\lambda_i > 0, \lambda_j < 0$ for some i, j

$\Rightarrow x^*$ is a saddle

H **is semi-definite:** means none of the above applies; typically some of the $\lambda_i = 0$

$\Rightarrow$???