Normal Equation: A Direct Solution

*Main Points:*

* Normal equation is an alternative to gradient descent.
* Solves for optimal θ directly by taking derivatives.

*Equation:*

* Normal equation formula: θ = (XTX)^(-1)XTy.

**Comparison: Gradient Descent vs. Normal Equation**

*Gradient Descent:*

* Requires choosing alpha.
* Needs many iterations.
* Complexity O(kn^2).
* Works well for large n.

*Normal Equation:*

* No need to choose alpha.
* No iterations required.
* Complexity O(n^3).
* Slows down with large n.
* **Note:** For very large n, consider an iterative process.

**Handling Noninvertibility**

*Main Points:*

* Use 'pinv' function for normal equation in Octave.
* 'pinv' provides θ even when XTX is noninvertible.

*Causes of Noninvertibility:*

* Redundant features: closely related, linearly dependent.
* Too many features (m ≤ n).

*Solutions:*

* Remove redundant features or use regularization.
* **Examples:**
  + Delete linearly dependent features.
  + Delete features if too many present.

**Key Takeaway:**

The normal equation directly calculates optimal θ without iterations, providing an alternative to gradient descent. It's efficient for small-to-moderate n but can slow with large n. Handling noninvertibility involves identifying and addressing redundant or excessive features.