**Linear Regression vs Multiple Regression**

| **Tiêu chí** | **Linear Regression** | **Multiple Regression** |
| --- | --- | --- |
| **Definition** | Mô hình hồi quy tuyến tính đơn giản, mô tả mối quan hệ giữa **một biến độc lập** (X) và **một biến phụ thuộc** (Y). | Mô hình hồi quy tuyến tính tổng quát hơn, mô tả mối quan hệ giữa **nhiều biến độc lập** (X₁, X₂, …, Xₙ) và một biến phụ thuộc (Y). |
| **Equation** |  |  |
| **Complexity** | Đơn giản, dễ hiểu và dễ tính toán. | Phức tạp hơn do phải xử lý nhiều biến độc lập và mối quan hệ giữa chúng. |
| **Use Cases** | Khi chỉ có **một yếu tố chính** ảnh hưởng đến Y, ví dụ: **dự đoán giá nhà dựa trên diện tích**. | Khi nhiều yếu tố ảnh hưởng đến Y, ví dụ: **dự đoán giá nhà dựa trên diện tích, số phòng, vị trí…**. |
| **Assumptions** | - Quan hệ giữa X và Y là tuyến tính.  - Phân phối của lỗi (ε) là chuẩn.  - Không có tự tương quan và phương sai của lỗi là đồng nhất. | - Giống Linear Regression, nhưng cần thêm giả định về **không có đa cộng tuyến** giữa các biến độc lập. |
| **Visualization** | Biểu đồ **2D** (đường thẳng trên đồ thị có 1 trục X và 1 trục Y). | Khó trực quan hóa hơn, thường sử dụng **biểu đồ 3D** hoặc **ma trận hệ số tương quan**. |
| **Risk of Overfitting** | Thấp hơn vì chỉ có một biến độc lập. | Cao hơn nếu có quá nhiều biến độc lập hoặc biến không liên quan. Cần kỹ thuật **Feature Selection** hoặc **Regularization** để giảm overfitting. |
| **Multicollinearity Concern** | Không có vì chỉ có một biến độc lập. | Có thể gặp nếu các biến độc lập tương quan mạnh với nhau. Có thể kiểm tra bằng **VIF (Variance Inflation Factor)**. |
| **Applications** | - Dự đoán tăng trưởng doanh thu dựa trên một yếu tố.  - Phân tích tác động của giá sản phẩm đến doanh số. | - Phân tích dữ liệu kinh doanh với nhiều yếu tố (marketing, giá, thời tiết…).  - Dự đoán kết quả học tập dựa trên nhiều yếu tố (thời gian học, điểm số trung bình, số buổi học thêm…). |