

POLYNOMIAL FIT MODELS

Why

We can cast various common probabilistic regression models into the probabilistic errors linear model by mentioning the input space and feature maps. This unifies our analysis.

Definition

A line fit model has input space \mathbf{R} and output space \mathbf{R} . We use a regression function $\phi : \mathbf{R} \to \mathbf{R}^2$ defined by $\phi(t) = (1, t)^{\top}$.

We think of $t \in T \subset \mathbf{R}$ as a "dose level" (T is an interval). Given dose levels t_1, \ldots, t_ℓ and repetitions n_1, \ldots, n_ℓ we obtain the design matrix. Here the regression function generates a line segment embedded in the plane \mathbf{R}^2 . We call the parameters the *intercept parameter* and *slope parameter*.

A parabola fit model has input space \mathbb{R} and output space \mathbb{R} . We use a regression function $\phi : \mathbb{R} \to \mathbb{R}^3$ defined by $\phi(t) = (1, t, t^2)^{\top}$. Here the regression space is a segment of a parabola embedded in space \mathbb{R}^3 (since $t \in T$ an interval).

These two are instance of polynomial fit models of degree $d \geq 1$, in which the regression function becomes $\phi : \mathbf{R} \to \mathbf{R}^{d+1}$ defined by $\phi(t) = (1, t, t^2, \dots, t^d)^{\top}$. In this case, the regression range $\phi(T)$ is a one-dimensional curve embedded in \mathbf{R}^{d+1} . In cases in which it is clear that the input space is a single real variable t, a linear model for a line fit (parabola fit, polynomial fit of degree d) is called a first-degree model (second-degree model, dth degree model).

m-way models

We can generalize to m-way dth degree polynomial fit models in which the input space is $X \subset \mathbb{R}^m$ and the regression function $\phi: \mathbb{R}^m \to \mathbb{R}^k$ (k is d+m choose d) is the vector of all monomials of degree d in m variables.

For example, a two-way third-degree model has a regression function

$$\phi(t_1, t_2) = \begin{bmatrix} 1 & t_1 & t_2 & t_1^2 & t_1t_2 & t_2^2 & t_1^3 & t_1^2t_2 & t_1t_2^2 & t_2^3 \end{bmatrix}^{\top}.$$

Or consider a three way second-degree model with regression function

$$\phi(t_1, t_2, t_3) = \begin{bmatrix} 1 & t_1 & t_2 & t_3 & t_1^2 & t_1 t_2 & t_1 t_3 & t_2^2 & t_2 t_3 & t_3^2 \end{bmatrix}^\top.$$

Both models will result in parameter vectors of size ten. We call these models *saturated* because they have every possible dth degree power or cross product of variables. In generally, a m-way dth degree model has d+m choose d mean parameters.

In contrast to saturated models we can talk about *nonsaturated* models. For example, a nonsaturated two-way second-degree model has $\phi : \mathbb{R}^2 \to \mathbb{R}^4$ where $\phi(t_1, t_2) = (1, t_1, t_2, t_1^2)^\top$.

