Существует несколько способов нахождения коэффициентов логистической регрессии. На практике часто используют метод максимального правдоподобия. Он применяется в статистике для получения оценок параметров генеральной совокупности по данным выборки. Основу метода составляет функция правдоподобия (likehood function), выражающая плотность вероятности (вероятность) совместного появления результатов выборки.

$$L\,(Y_1,\,Y_2,\,\ldots,\,Y_k;\,\theta) = p\,(Y_1;\,\theta)\cdot\ldots\cdot p\,(Y_k;\,\theta)$$

Согласно методу максимального правдоподобия в качестве оценки неизвестного параметра принимается такое значение, которое максимизирует функцию L.

$$\theta = \theta(Y_1, \ldots, Y_k),$$

Нахождение оценки упрощается, если максимизировать не саму функцию L, а натуральный логарифм ln(L), поскольку максимум обеих функций достигается при одном и том же значении θ:

$$P_i = Prob(Y_i = 1).$$

В случае бинарной независимой переменной, которую мы имеем в логистической регрессии, выкладки можно продолжить следующим образом. Обозначим через Pi вероятность появления единицы:

$$L^*(Y;\,\theta) = \ln\left(L\,(Y;\,\theta)\,\right) \rightarrow \max$$

Эта вероятность будет зависеть от XiW, где Xi — строка матрицы регрессоров, W — вектор коэффициентов регрессии:

$$P_i = F\,(X_i W),\; F(z) = \frac{1}{1+e^{-z}}$$

Логарифмическая функция правдоподобия равна:

$$L^* = \sum_{i\epsilon\, I_1} lnP_i(W) + \sum_{i\epsilon\, I_0} ln(1 - P_i(W)) = \sum_{i=1}^{k} [Y_i \ln P_i(W) + (1 - Y_i)\ln\left(1 - P_i(W)\right)]$$

, где I0, I1— множества наблюдений, для которых Yi=0 и Yi=1 соответственно.

Можно показать, что градиент g и гессиан H функции правдоподобия равны: