ЛЕКЦИЯ 9

Метод наискорейшего спуска

$G(\omega) = (\nabla E(\omega))^T = (I^T(\omega)e(\omega))^T$, I - якобиан, $E = \frac{1}{2}e^T e$ - ошибки выходного слоя.

$\omega_i[t+1] = \omega_{ij}[t] + (-\eta G(\omega_{ij}))$, здесь $G(\omega_{ij})$ - скалярное значение.

$E(\omega + p) = E(\omega) + G(\omega)p + \frac{1}{2}p^T H(\omega)p + o(p^2)$.

Здесь должна быть положительно определенная квадратичная форма ($H(\omega) > 0$), чтобы был минимум. Это необходимое и достаточное условие существования минимума.

$G(\omega) + p^T H(\omega) = 0$.

Находим p:

$p^T H(\omega) = -G(\omega)$,

$H(\omega)p = -G^T(\omega)$,

$p = -H^{-1}(\omega)G^T(\omega)$.

p - вектор, на который мы должны прирастить веса, чтобы достичь точки минимума.

Методы переменной метрики (алгоритмы переменной метрики)

Выпишем один алгоритм такого типа:

$s_k = \omega[t] - \omega[t-1]$,

$r_k = [G(\omega[t]) - G(\omega[t-1])]^T$,

V_k - аппроксимация $H^{-1}(\omega)$.

$$V_k = V_{k-1} + [1 + \frac{r_k^T V_{k-1} r_k}{r_k}]\frac{s_k s_k^T}{s_k^T r_k} - \frac{s_k r_k^T V_{k-1} r_k s_k^T}{s_k^T r_k}.$$

Эта формула называется BFGS или DFP.

$V_o = \mathrm{I}$ - единичная матрица.

Это один из методов нахождения матрицы, обратной гессиану. Это есть метод 2го порядка, у него сходимость быстрее, чем у метода 1го порядка (метод градиентного спуска).

Метод Левенберга-Марквардта

Это метод 2го порядка.

$H(\omega) = I^T(\omega)I(\omega) + s(\omega)$.

$\Gamma(\omega) = I^T(\omega)I(\omega) + \nu\mathrm{I}$ - метод аппроксимации гессиана.

ν - фактор.

2 случая:

1. далеко от точки минимума: $\Gamma(\omega) = \nu\mathrm{I}_0$;

2. близко к точке минимума $\Gamma(\omega) = I^T(\omega)I(\omega), \nu = 0$ (тут метод Гаусса-Ньютона и ν обнуляется).