A mysterious getDir() function is used in the fitting part, we have the objective function $y = f(x)$ and gradient $g = grad(x)$, then the direction is solved by solving the linear equation $Ax = b$.

$$\begin{pmatrix} g \\ 0.1yI \end{pmatrix} dx = \begin{pmatrix} y \\ 0 \end{pmatrix} \tag{1}$$

Note the coefficient matrix is $n+1$ by n, so the linear equations are overdetermined, and solved with least square $\min_x ||b - Ax||^2$. We are not going to discuss why this direction is chosen, but when large impurity is used, g can be quite large, and the least square routine takes forever. We are here to discuss how to compute dx efficiently.

The analytic expression for dx is

$$dx = (A^T A)^{-1} A^T b \tag{2}$$

Due to the simple form of A and b, we can evaluate the expression. Since

$$A^T A = \begin{pmatrix} g^T & 0.1yI \end{pmatrix} \begin{pmatrix} g \\ 0.1yI \end{pmatrix} = g^T g + 0.01y^2 I = 0.01y^2(100g^T g/y^2 + I) \tag{3}$$

And $A^T b = g^T y$, we have

$$dx = 100y^{-2}(100g^T g/y^2 + I)^{-1} g^T y = 100(100g^T g/y^2 + I)^{-1} g^T/y = 10(h^T h + I)^{-1} h^T \tag{4}$$

where $h = 10g/y$. Now we just need to evaluate $(h^T h + I)^{-1}$. Generally, if we can diagonalize a matrix $M = U\Lambda U^T$, then

$$(M + I)^{-1} = U(I + \Lambda)^{-1} U^T \tag{5}$$

can be easily evaluated. In our case, diagonalizing $h^T h$ is particularly straightforward,

$$h^T h = \begin{pmatrix} \tilde{h}^T & V^T \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} h \\ V \end{pmatrix} \tag{6}$$

where $\tilde{h} = h/||h||$ is simply the normalized vector of h, and V the complimentary space which forms a orthogonal matrix with $\tilde{h}$. And $\lambda = ||h||^2 = hh^T$. Therefore

$$\begin{aligned}(h^T h + I)^{-1} &= \begin{pmatrix} \tilde{h}^T & V^T \end{pmatrix} \begin{pmatrix} (\lambda+1)^{-1} & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} h \\ V \end{pmatrix} \\ &= \begin{pmatrix} \tilde{h}^T & V^T \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} h \\ V \end{pmatrix} - \begin{pmatrix} \tilde{h}^T & V^T \end{pmatrix} \begin{pmatrix} 1-(1+\lambda)^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} h \\ V \end{pmatrix} \\ &= I - \tilde{h}^T[1-(1+\lambda)^{-1}]\tilde{h} = I - (1+\lambda)^{-1} h^T h \end{aligned} \tag{7}$$

Further, we have

$$dx = 10[I - (1+\lambda)^{-1} h^T h] h^T = \frac{10}{1+\lambda} h^T \tag{8}$$