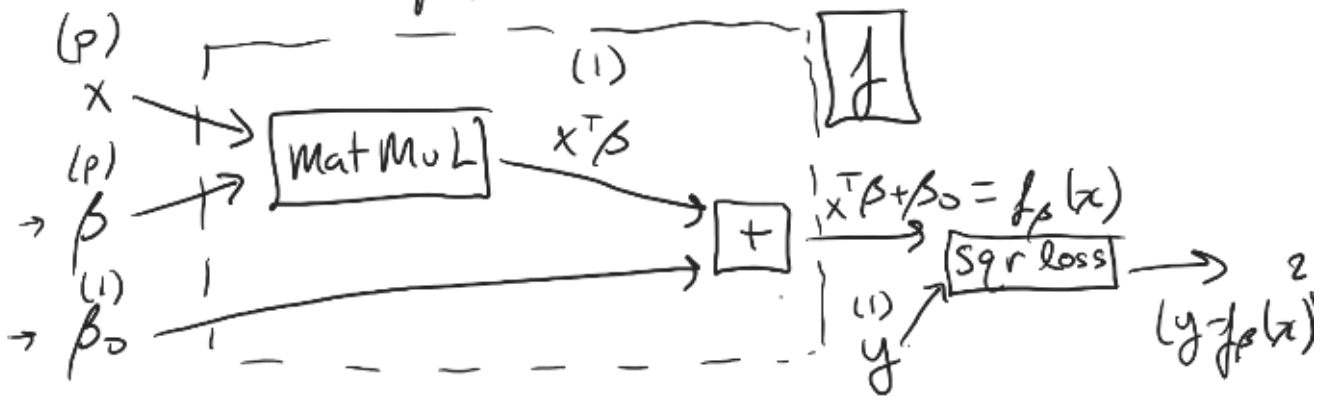
Computation Graphs

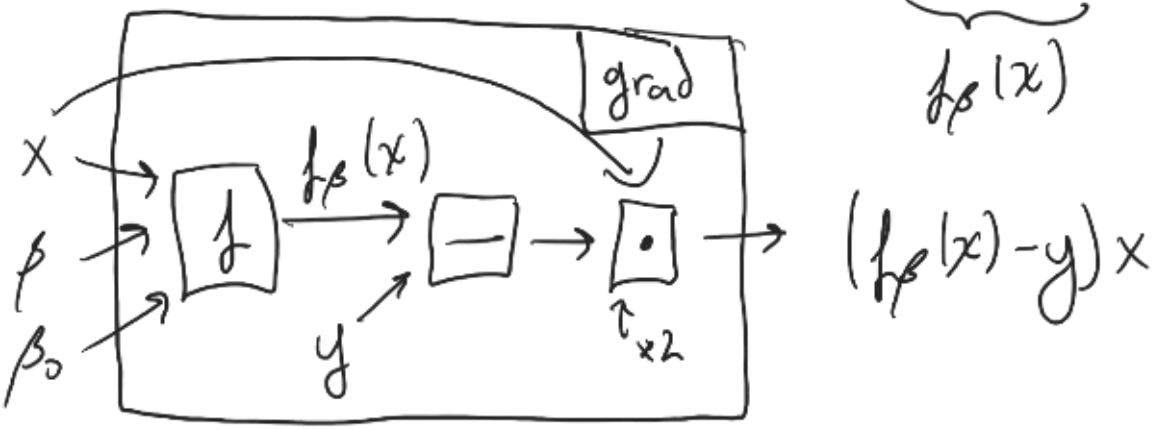
predict: $f_\beta(x) = x^T\beta + \beta_0$

input ↗ ↑ variables

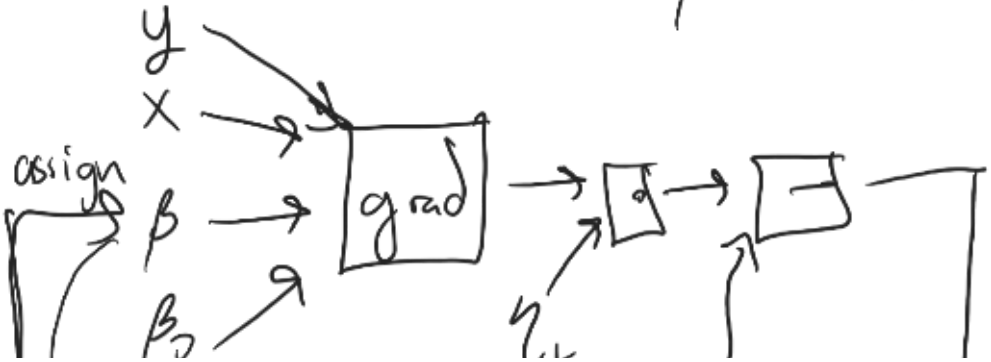


Computation graph for predict and loss

$$\frac{\partial}{\partial\beta}\left(y-(x^T\beta+\beta_0)\right)^2 = -2\left(y-\underbrace{(x^T\beta+\beta_0)}_{f_\beta(x)}\right)^{(p)} x$$



$$\beta \leftarrow \beta - \eta_t \frac{\partial}{\partial\beta} \ell(y, f_\beta(x))$$



Exercise 10.1

Outline the compute graph for Stoch. Grad Descent for SVM

$$(1 - y(x^T\beta + \beta_0))_+ + \lambda \|\beta\|_2^2$$