6. (10 points) A concert venue is trying to predict how many people will attend an event based on features such as the date of the concert, the genre of the band, etc. They will use this model to decide which bands to book, but they would like the predictions to be conservative: if their model under-estimates the attendance it is better than if it overestimates. We can express this preference in a loss function, where g is the predicted ("guessed") value and y is the true target value:

$$L(g, y) = \begin{cases} c_1(g - y)^2 & \text{if } g > y \\ c_2(g - y)^2 & \text{otherwise} \end{cases}$$

(a) For what values of c_1 and c_2 is this loss function equivalent to squared error?

(c) Consider a linear model in which $g = \theta^T x + \theta_0$. Derive a stochastic gradient descent update rule for θ and θ_0, with step size η. (If you are having trouble thinking about θ and x as vectors, start by figuring it out for the scalar case.)