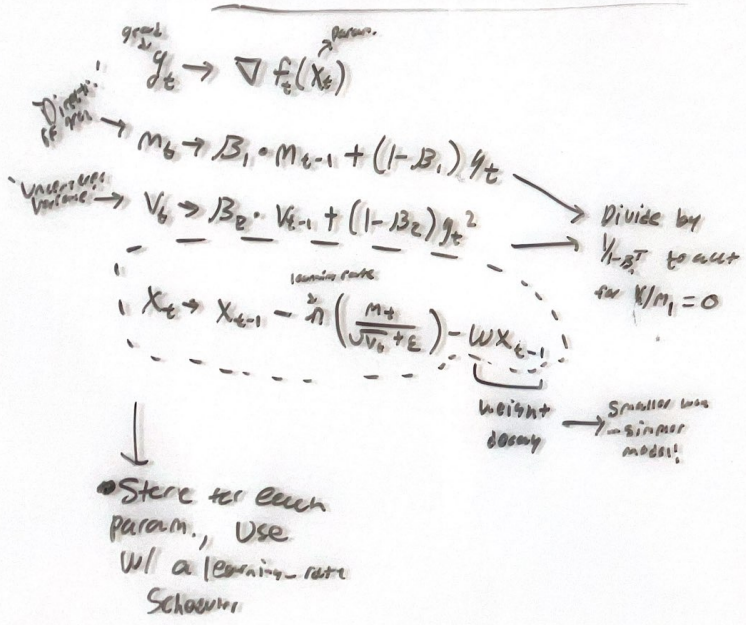
⛤ AdamW GD. Optimizer

- For each param. Geeps track of how steep/ gradient, guides step sizes accordingly



- Store for each param., use w/ a learning rate scheduler