ASSET PRICING MEETING 3: NOTES

ADEN CHEN

1. Gradient Descent

Proposition 1.1. *Fixing step size α (that is, given the constraint that $\|h\| = \alpha$), shifting the input in the direction of the gradient increases the function value by the most.*

Intuition: We have

$$f(x+h) - f(x) \approx \langle \nabla f(x), h \rangle,$$

where $\nabla f(x)$ is the gradient of f at x. Cauchy-Schwarz gives

$$\langle \nabla f(x), h \rangle \leq \|\nabla f(x)\| \|h\| = \alpha \|\nabla f(x)\|,$$

where equality is achieved when $h = \alpha \nabla f(x)$.

Example 1.2. Consider minimizing the differentiable function

$$f : \mathbb{R}^n \to \mathbb{R}.$$

0. Pick any input x_0.
1. Compute the gradient at the current input
$$\nabla f(x_k) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}.$$
2. Update the input:
$$x_{k+1} = x_k - \alpha \nabla f(x_k),$$
where α is the learning rate. (Pick for example $\alpha := e^{-x}$.)
3. Repeat steps 1 and 2 until the gradient is close to zero.

See exercise at: https://github.com/AdenChen27/Oeconomica2024-AssetPricing/blob/main/winter/l5.ipynb