- Recall definition of eigenvalues and eigenvectors.

matrix A, eigenvalue $\lambda \in \mathbb{R}$, eigenvector $x \in \mathbb{R}^n$

$$\Rightarrow \quad Ax = \lambda x$$

- Normally there are $\underline{\underline{n}}$ eigs.
- zero eig means there is a vector $x \neq 0$ s.t. $Ax = 0$
- eigs are complex in general.
- eigs of a symmetric matrix are real.
- Hessian is symmetrix.
- Example: Find eigs of $A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$

$\Rightarrow$ Look for $\lambda \in \mathbb{R}$ and $x \in \mathbb{R}^2$ s.t.

$$\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{or} \quad \det\left(\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right) = 0$$

$$(\det(A - \lambda I) = 0)$$

- Consider $x^T A x$, where $A \in \mathbb{R}^{n\times n}$ is a symmetric matrix.

- Note that $x^T A x = 0$ if $x = 0$

- If $x^T A x > 0$ for all $x \neq 0 \Rightarrow A \succ 0$ positive definite

- If $x^T A x \geq 0$ for all $x \neq 0 \Rightarrow A \succeq 0$ positive semidefinite

- Example: $A = \begin{bmatrix} 1 & 2 \\ 2 & -4 \end{bmatrix}$ is not positive definite

because: $\begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & -4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -4 < 0$

- Assume that x is an eigenvector of A associated with the eigenvector λ: $Ax = \lambda x$

$\Rightarrow$ $x^T A x = x^T(\lambda x) = \lambda \|x\|^2 \Rightarrow$ sign of $x^T A x$ depends on λ.

$\Rightarrow$ 1- $A \succ 0 \iff$ eigs of $A > 0$

2- $A \succeq 0 \iff$ eigs of $A \geq 0$

- Back to $\min_{x \in \mathbb{R}^n} f(x)$: $x_* =$ Local min

$$\Rightarrow \nabla f(x_*) = 0$$

$$\Rightarrow f(x_* + \Delta x) = f(x_*) + \frac{1}{2} \Delta x^T H(f(x_*)) \Delta x + \cdots$$

since $f(x_* + \Delta x) > f(x_*)$ for a small Δx (x_* is a local min)

$$\Rightarrow \Delta x^T H(f(x_*)) \Delta x \geq 0 \quad \text{for small } \Delta x$$

$\Rightarrow$ 1- x_* : Local min $\Rightarrow H(f(x_*)) \geq 0$

2- $H(f(x_*)) > 0 \Rightarrow \Delta x^T H(f(x_*)) \Delta x > 0$ for $\Delta x \neq 0$
and
$\nabla f(x_*) = 0 \Rightarrow x_* =$ Local min

3- $H(f(x_*))$ has a negative eig

$\Rightarrow \Delta x^T H(f(x_*)) \Delta x < 0$ for some Δx

$\Rightarrow x_*$: not a local min

4- $H(f(x_*))$ is positive semidefinite but not positive definite $\Rightarrow \Delta x^T H(f(x_*)) \Delta x = 0$ for some $\Delta x \neq 0$ $\Rightarrow$ high order terms should be checked.

- Numerical Algorithm:

$\min_{x \in \mathbb{R}^n} f(x) \longrightarrow$ Find a stationary point (min or saddle)

- strategy:
    - Guess a solution $x^{(0)}$
    - perform some computation and improve it to $x^{(1)}$
    - perform some computation and improve it to $x^{(2)}$

    ⋮

$\Rightarrow$ sequence: $x^{(0)} \xrightarrow[\text{iteration}]{} x^{(1)} \xrightarrow[\text{iteration}]{} x^{(2)} \longrightarrow \cdots$

Hope: sequence converges to x_*

- How to ensure there is an improvement at every iteration?

$$f(x^{(0)}) > f(x^{(1)}) > f(x^{(2)}) > \cdots$$

- The value of the function must reduce.

- Assume we are at iteration $K \in \{1, 2, \ldots\}$

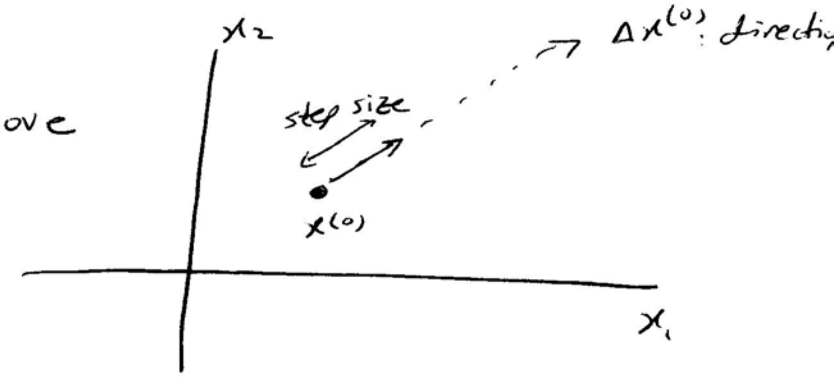
$x^{(K-1)} \longrightarrow x^{(K)}$ need a rule

- strategy:

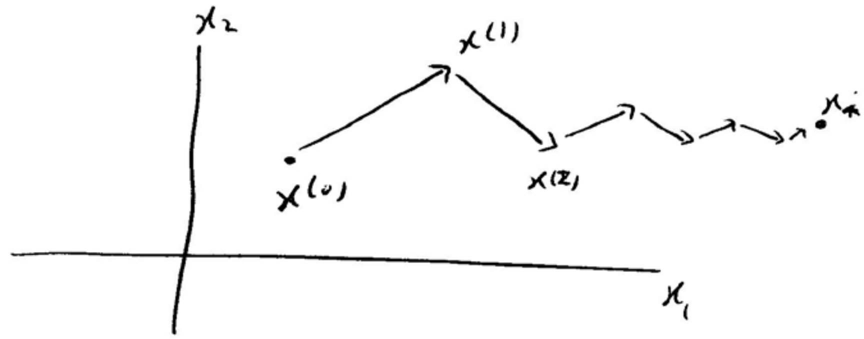
$$x^{(k)} = \underbrace{x^{(k-1)} + t^{(k-1)}}_{\substack{\text{step size} \\ \in \mathbb{R}^+}} \underbrace{\Delta x^{(k-1)}}_{\text{direction} \in \mathbb{R}^n}$$

- Example: $\min \; x_1^2 + e^{x_1 - x_2} + x_2^2$

- Direction: where to move
- step size: how far to go in that direction



$\Rightarrow$



- Recall that:

$$f(x^{(k)}) = f(x^{(k-1)} + t^{(k-1)} \Delta x^{(k-1)})$$

$$= f(x^{(k-1)}) + t^{(k-1)} \underbrace{\nabla f(x^{(k-1)}) \Delta x^{(k-1)}} + \underbrace{\cdots}_{\substack{\text{high-order} \\ \text{terms}}}$$

- Thm: If $\nabla f(x^{(k-1)}) \Delta x^{(k-1)} < 0 \Rightarrow f(x^{(k)}) < f(x^{(k-1)})$

for $\underbrace{\text{a small}}$ number $t^{(k-1)}$.

- Thm: If $\nabla f(x^{(k-1)}) \Delta x^{(k-1)} > 0$, then

$$f(x^{(k)}) > f(x^{(k-1)}) \text{ for all small numbers } t^{(k-1)}.$$

$\Rightarrow$ Need to design $\Delta x^{(k-1)}$ s.t. $\nabla f(x^{(k-1)}) \Delta x^{(k-1)} < 0$.

- A direction Δx at a point x is called descent if $\nabla f(x) \Delta x < 0$

- Example, $f(x) = (x_1 - 1)^2 + (x_2 - 2)^2$

  Consider the point $\bar{x} = (2, -1)$.

  Descent directions are: $\begin{bmatrix} 2(2-1) & 2(-1-2) \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} < 0$

  $\Rightarrow 2\Delta x_1 - 6\Delta x_2 < 0$

- How to choose Δx? $\rightarrow$ Gradient Algorithm
  - $\rightarrow$ Newton
  - $\rightarrow$ Steepest descent