Optimality conditions for unconstrained optimization

Wednesday, August 23, 2017 14:50

For the glory of God

Optimality conditions for unconstrained optimization

- Let's say that you have found a minimum. How would you guarantee that you found the minimum?
  - If the function is of unknown form, we would have little or no knowledge about its shape.
- First-order Necessary Condition — We will have details at the end of this hand-written note.
  - Let $f(x)$ be a function that is differentiable at x^*
  - $\nabla f(x^*) = 0$ is necessary condition for x^* to be a local solution to the unconstrained problem.

$$x^* \text{ is a local solution } \underset{\text{Not always}}{\overset{\text{Always}}{\rightleftarrows}} \nabla f(x^*) = 0$$

Definition of a local solution

- a) x^* is a weak local solution if $f(x^*) \le f(x)$ for all x in a neighborhood around x^*
- b) x^* is a strong local solution if $f(x^*) < f(x)$ for all x in a neighborhood around x^*

- Second-order optimality condition (It would be further investigated) e.g. 1st order → $x = 1, 2, 3$ satisfied; 2nd order → $x=1$ and $x=3$ only
  - Let $f(x)$ be a function that is twice differentiable at x^*
  - $\nabla f(x^*) = 0$ and $H(x^*)$ is positive semi-definite is necessary condition for x^* to be a local solution.
  - $\nabla f(x^*) = 0$ and $H(x^*)$ is positive definite is sufficient condition for x^* to be a (strong) local solution.

    Hessian matrix ↵ — After $\det(H - \lambda I) = 0$

    It will be a global minimum if the function is convex

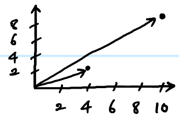
No line lies below the graph at any points

Matrix is
- positive definite if all eigenvalues are positive
- positive semi-definite if all eigenvalues are non-negative

Let $A \in R^{n\times n}$ we say that v is an Eigenvector of A matrix if there exist a number λ such that $Av = \lambda v$ and $v \ne 0$. Here, λ is called as an Eigenvalue of A. (It would be stretch/shrink in the same line.)

- For example,

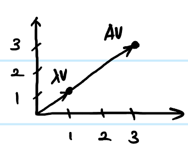
$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 10 \\ 8 \end{bmatrix}$$; $\begin{bmatrix} 4 \\ 2 \end{bmatrix}$ is not an Eigenvector of A because of no relation between $\begin{bmatrix} 4 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 10 \\ 8 \end{bmatrix}$



These vectors are not in the same line

What about this case?

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$



; These are in the same line.

$$\Leftrightarrow \underset{A}{\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}}\underset{v}{\begin{bmatrix} 1 \\ 1 \end{bmatrix}} = \underset{\lambda}{3}\underset{v}{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}$$

$\therefore$ Eigenvalue $= 3$ and Eigenvector $= \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

It tells us things about the matrix that are useful in systems of linear equations

- Steps to find Eigenvectors and Eigenvalues (So, where does it come from?) e.g. Find A^{-1}

  ↳ please see the end of this note.

  a) Solve $\det(A - \lambda I) = 0$; $\lambda_1, \lambda_2, \ldots, \lambda_n$ = Eigenvalues

- Steps to find Eigenvectors and Eigenvalues (So, where does it come from ?)   e.g. Find A^{-1}

  ↳ please see the end of this note.

  a) Solve $\det(A - \lambda I) = 0$; $\lambda_1, \lambda_2, \cdots \lambda_n$ = Eigenvalues

  b) For each eigenvalue λ, Solve $(A - \lambda I)V = 0$; V = Eigenvector

  → Except for $V = 0$

  ↳ Identity matrix is a square matrix in which all the elements of the principal diagonal are ones and all other elements are zero. ($AI = A$)

  ↳ The effect

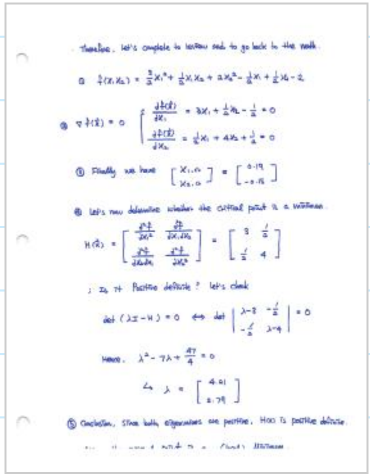
cf. Hessian matrix

· In mathematics, the Hessian matrix is a square matrix of second-order partial derivatives of scalar field.

· This matrix is useful for unconstrained optimization problem.

$$H = \begin{bmatrix} \dfrac{d^2 f(x)}{dx_1^2} & \dfrac{d^2 f(x)}{dx_1\, dx_2} & \cdots & \dfrac{d^2 f(x)}{dx_1\, dx_n} \\ \dfrac{d^2 f(x)}{dx_2 dx_1} & \dfrac{d^2 f(x)}{dx_2^2} & \cdots & \\ \vdots & \vdots & \ddots & \\ \dfrac{d^2 f(x)}{dx_n\, dx_1} & & & \dfrac{d^2 f(x)}{dx_n^2} \end{bmatrix}$$

→ Example of how to use the Hessian matrix


③ Finally we have $\begin{bmatrix} x_{1,0} \\ x_{2,0} \end{bmatrix} = \begin{bmatrix} 0.19 \\ -0.15 \end{bmatrix}$

$H(\hat{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 3 & \frac{1}{2} \\ \frac{1}{2} & 4 \end{bmatrix}$

; Is it positive definite ? let's check

$\det(\lambda I - H) = 0 \iff \det \begin{vmatrix} \lambda - 3 & -\frac{1}{2} \\ -\frac{1}{2} & \lambda - 4 \end{vmatrix} = 0$

Hence, $\lambda^2 - 7\lambda + \frac{47}{4} = 0$

↳ $\lambda = \begin{bmatrix} 4.21 \\ 2.79 \end{bmatrix}$

⑤ Conclusion, since both eigenvalues are positive, H is positive definite.

Differentiability and Continuity

· Generally speaking, a real function is said to be differentiable at a point (say c) if ;

  - Its derivative exist at that point

  - Its function value exist at that point

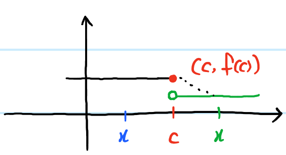
  - Its derivative for both left and right should be same.

· What about the relationship between differentiability and continuity ?

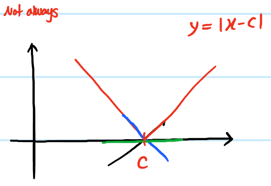
  ⇒ If f is differentiable at $x = c$, then f is continuous at $x = c$ ⇔ if f isn't continuous, it's not differentiable

  ↳ However, if f is continuous, it hasn't to be differentiable.

· For example,   ↳ Hence, Differentiable at $x = c$ $\underset{\text{Not always}}{\overset{\text{Always}}{\rightleftarrows}}$ Continuous at $x = c$



vs.



$f'(c) = \lim\limits_{x \to c} \dfrac{f(x) - f(c)}{x - c}$

For LHS, $f'(c)$ = negative

For RHS, $f'(c)$ = positive

for left, $f'(c) = 0$ (with x)

Obviously, this is continuous

for right, $f'(c)$ → approaching negative ∞

but not differentiable at c

for x, $f'(c) = \dfrac{f(x) - f(c)}{x - c}$

$= \dfrac{2-4}{5-2} = \ominus$

also, there are infinite slopes at c,

↳ all are possible.

Hence, it had derivative at c
it had function value at c
but LHS $\neq$ RHS
also, it's not continuous

. Review Linear Algebra

a) Determinant

: In general, there has been no definition in terms of the Determinant. However, it has been called as determinant because the result from Determinant can be considered as very important things in Matrix calculation. For instance,

det (A) = 0 : Matrix A doesn't have an inverse matrix
det (A) $\neq$ 0 : " does have inverse matrix so that we can use a couple of methods such as Gauss Elimination method to find a solution.

b) Characteristic Equation

: It is an equation to find out eigen values of certain Matrix A. In other words, the roots of the characteristic equation could be referred to as an eigen value. If we obtain the eigen value, we could estimate eigen vectors.

$$\det(\lambda I - A) = 0$$; where I is a unit matrix

Then, how could we obtain the formulation of the Equation ?

In order to answer the question, we need to understand a concept of both Eigen value and Eigen vector first.

c) Eigenvalues and Eigenvectors

: Suppose you have A matrix and hundreds vectors. Imagine that almost all vectors change direction when they are multiplied by A.

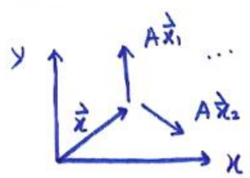
e.g. $\vec{x} = (1,1)$

$\downarrow$

multiplied by A

$\downarrow$

$A\vec{x}_1 = (0.3, 2)$



To explain eigenvalues, we first explain eigenvectors. From the assumption as mentioned above, certain exceptional vectors $\vec{x}$ are in the same direction as $A\vec{x}$. Those are called as Eigenvectors. It might be showing the certain characteristics of the matrix A.

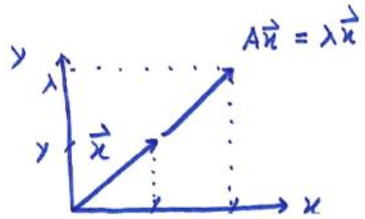
e.g. $\vec{x} = (1,1)$

$\downarrow$

Multiplied by A

$\downarrow$

$A\vec{x}_1 = (1,1)$ or



· Then, what about Eigenvalues ?

: The eigenvalue (λ) tells whether the special vector $\vec{x}$ is stretched or shrunk or reversed or left unchanged when it is multiplied by A. We may find $\lambda = 2$ or $\lambda = -1$; however, it could be zero.

· In summary,

The basic equation of Eigenvectors and Eigenvalues is ;

$$Ax = \lambda x$$

; where the number λ is an eigenvalue of A.

· Let's go back to the characteristic equation. We asked how could we get the formation ?

* Spring 2017

: From the basic equation, we have

$$Ax = \lambda x \iff (A - \lambda I)x = 0$$

If we assume that we have a solution(s),

$\begin{cases} x \text{ shouldn't be zero. (It has to be a solution)} \\ (A-\lambda I) \text{ of determinant will be zero. } \det(A-\lambda I) = 0 \end{cases}$

Example of Eigenvector and Eigenvalue

let's say that we have arbitrary matrix ;

$$\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ -3 \end{pmatrix}$$

× (It couldn't be related)

what about this case ?

$$\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix}$$

o

$$= 3 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

; we get the same vector out

Eigenvalue

Eigenvector

$\therefore \quad A x = \lambda x$ (General Equation)