# Διάλεξη 7

Βελτιστοποίηση συναρτήσεων χωρίς περιορισμούς, Τρίτη 13/3/18, Πέμπτη 15/3/18

Σε αυτή τη διάλεξη θα ασχοληθούμε με το πρόβλημα $\min_{\vec{x} \in A} f(\vec{x})$ όπου $f : \mathbb{R}^n \to \mathbb{R}$ και $A$ ένα ανοικτό υποσύνολο του $\mathbb{R}^n$.

Στην προηγούμενη διάλεξη ασχοληθήκαμε με την περίπτωση $n = 1$. Εδώ θα κοιτάξουμε την περίπτωση $n \geq 2$. Το πλήθος $n$ των μεταβλητών απόφασης σε ρεαλιστικά προβλήματα είναι από μερικές δεκάδες έως χιλιάδες.

## Περιεχόμενα



## 1. Διάνυσμα κλίσης

***Ορισμός:*** ***Η <u>κατά κατεύθυνση παράγωγος της</u>*** $f : \mathbb{R}^n \to \mathbb{R}$ ***<u>στο σημείο</u>*** $\vec{x}$ ***<u>στη διεύθυνση</u>*** $\vec{\delta} \in \mathbb{R}^n$ ***είναι*** $f'(\vec{x} + t\vec{\delta})\Big|_{t=0} = \lim_{t \to 0} \frac{f(\vec{x} + t\vec{\delta}) - f(\vec{x})}{t}$

Φαίνεται ότι η γραφική παράσταση μιας συνάρτησης πολλών μεταβλητών μπορεί να έχει άπειρα πολλές δυνατές κλίσεις, μια για κάθε δυνατή διεύθυνση $\vec{\delta}$. Θα δούμε ότι για τον υπολογισμό τους αρκεί να γνωρίζουμε την κλίση κατά μήκος των διευθύνσεων των αξόνων, δηλαδή τις μερικές παραγώγους $\frac{\partial f(\vec{x})}{\partial x_i}$ για $i = 1, \ldots, n$.

***Ορισμός:*** ***Το <u>διάνυσμα κλίσης της συνάρτησης</u>*** $f$ ***<u>στο σημείο</u>*** $\vec{x}$ ***είναι το διάνυσμα των μερικών παραγώγων:*** $\left(\frac{\partial f(\vec{x})}{\partial x_1}, \ldots, \frac{\partial f(\vec{x})}{\partial x_n}\right)$ ***και συμβολίζεται με*** $\nabla f(\vec{x})$.

Οι κατά κατεύθυνση παράγωγοι είναι γραμμικοί συνδυασμοί των μερικών παραγώγων:

***ΘΕΩΡΗΜΑ 7.1:*** $f'(\vec{x}+t\vec{\delta})\Big|_{t=0} = \nabla f(\vec{x}) \cdot \vec{\delta}$

***ΘΕΩΡΗΜΑ 7.2 [Αναγκαία συνθήκη βέλτιστης λύσης (για προβλήματα χωρίς περιορισμούς)]:***
***Εάν $\vec{x}^*$ βέλτιστη λύση τότε ισχύει $\nabla f(\vec{x}^*) = 0$.***

**Απόδειξη:** Εάν $\vec{x}^* = (x_1^*, \ldots, x_n^*)$ βέλτιστη λύση τότε ο περιορισμός $g(t) = f(\vec{x}^* + t\vec{\delta})$ της $f$ κατά μήκος της διεύθυνσης $\vec{\delta}$ ελαχιστοποιείται για $t = 0$. Άρα από την αναγκαία συνθήκη για βέλτιστη λύση στην περίπτωση συναρτήσεων μιας μεταβλητής (Θεώρημα 6.1), $g'(0) = 0$. Όμως $g'(0) = f'(\vec{x}^* + t\vec{\delta})\Big|_{t=0}$ και άρα $\nabla f(\vec{x}) \cdot \vec{\delta} = 0$ από το παραπάνω θεώρημα. Θέτοντας $\vec{\delta} = (0, \ldots, 0, 1, 0, \ldots, 0)$ όπου το 1 βρίσκεται στην $i$-οστή συνιστώσα, συμπεραίνουμε ότι $\dfrac{\partial f(\vec{x}^*)}{\partial x_i} = 0$ για κάθε $i = 1, \ldots, n$. ■

# 2. Κυρτές συναρτήσεις

Στην προηγούμενη διάλεξη δώσαμε τον ορισμό μιας κυρτής συνάρτησης μιας μεταβλητής.
Ακολουθεί ο ορισμός για συναρτήσεις πολλών μεταβλητών:

***Ορισμός: Η συνάρτηση $f : \mathbb{R}^n \to \mathbb{R}$ όπου $n \geq 2$ είναι κυρτή εάν και μόνο εάν ο περιορισμός της, $g : \mathbb{R} \to \mathbb{R}$ όπου $g(t) = f(\vec{x} + t\vec{\delta})$ προς οποιαδήποτε διεύθυνση $\vec{\delta} \in \mathbb{R}^n$ είναι κυρτή συνάρτηση της μεταβλητής $t$.***

Ο ορισμός αυτός δεν είναι κυκλικός αφού ορίζεται η κυρτότητα για $n \geq 2$ χρησιμοποιώντας τον ορισμό για $n = 1$ που δόθηκε στην προηγούμενη διάλεξη.
Για τον έλεγχο της κυρτότητας μιας συνάρτησης είναι πιο εύκολο αντί του ορισμού να χρησιμοποιηθεί το κριτήριο στο Θεώρημα 7.3 που ακολουθεί. Όπως η κυρτότητα στη μια διάσταση εξαρτάται από τη δεύτερη παράγωγο, έτσι και το κριτήριο αυτό εξετάζει τις παραγώγους δεύτερης τάξης.

***Ορισμός: Η <u>Εσσιανή της συνάρτησης</u> $f:\mathbb{R}^n \to \mathbb{R}$ <u>στο σημείο</u> $\vec{x} \in \mathbb{R}^n$ είναι ο $n \times n$ πίνακας***

$$H(f,\vec{x}) = \begin{bmatrix} \frac{\partial^2 f(\vec{x})}{\partial x_1^2} & \cdots & \frac{\partial^2 f(\vec{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\vec{x})}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(\vec{x})}{\partial x_n^2} \end{bmatrix}$$

***όλων των μερικών παραγώγων δεύτερης τάξης.***

***Ορισμός: ο συμμετρικός τετραγωνικός πίνακας $A$ είναι <u>θετικά ημιορισμένος</u> εάν και μόνο εάν $\vec{\delta}^T A \vec{\delta} \geq 0$ για κάθε διάνυσμα στήλη $\vec{\delta} \in \mathbb{R}^n$.***

***ΘΕΩΡΗΜΑ 7.3: [Κριτήριο κυρτότητας]***
***Η συνάρτηση $f:\mathbb{R}^n \to \mathbb{R}$ είναι κυρτή εάν και μόνο εάν η Εσσιανή της σε κάθε σημείο $\vec{x}$ είναι θετικά ημιορισμένος πίνακας.***

**Απόδειξη:** Εάν η $f$ είναι κυρτή τότε (από τον ορισμό) το ίδιο ισχύει για τη $g(t) = f(\vec{x} + t\vec{\delta})$ για κάθε $\vec{x}, \vec{\delta} = (\delta_1, \ldots, \delta_n)$ και άρα $g''(t) \geq 0$ για κάθε $t$. Όμως,

$$g''(t) = \sum_{i,j} \delta_i \delta_j \frac{\partial^2 f(\vec{x})}{\partial x_i x_j} = \vec{\delta}^T H(f,\vec{x}) \vec{\delta}$$

, δηλαδή η Εσσιανή είναι θετικά ημιορισμένος πίνακας. ■

## Παράδειγμα 1: γραμμικές συναρτήσεις

Κάθε γραμμική συνάρτηση $f(x_1, \ldots, x_n) = \sum_i a_i x_i + b$ είναι κυρτή αφού η Εσσιανή είναι ο μηδενικός πίνακας (γιατί;)

## Παράδειγμα 2

Η συνάρτηση $f(x_1, x_2) = e^{x_1 + x_2}$ είναι κυρτή αφού η Εσσιανή της $H(f,\vec{x}) = e^{x_1+x_2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ είναι θετικά ημιορισμένος πίνακας:

$$\begin{bmatrix} \delta_1 & \delta_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix} = (\delta_1 + \delta_2)^2 \geq 0$$

για κάθε $\delta_1, \delta_2 \in \mathbb{R}$.