Κεφάλαιο 6

Εφαρμογές της Παραγώγου

Λόγω της φυσικής της ερμηνείας ως ρυθμός μεταβολής, η παράγωγος εμφανίζεται πολύ συχνά στις φυσικές επιστήμες και τα εφαρμοσμένα μαθηματικά. Στο κεφάλαιο αυτό θα εξετάσουμε εν συντομία μερικές χαρακτηριστικές εφαρμογές της.

Στην Παράγραφο 6.1 δείχνουμε πώς μπορούμε με λίγο κόπο να υπολογίσουμε προσεγγιστικά τις τιμές μιας συνάρτησης γύρω από ένα σημείο. Η εφαρμογή δεν είναι πολύ σημαντική, αλλά μας βοηθά να ξεκαθαρίσουμε τι σημαίνει η παράγωγος μιας συνάρτησης σε ένα σημείο και η εφαπτομένη ευθεία. Στην Παράγραφο 6.2 αναφέρουμε εν συντομία τον Κανόνα του L'Hôpital και ορισμένες παγίδες στη χρήση του. Στην Παράγραφο 6.3 κάνουμε μια σύντομη εισαγωγή στις κυρτές συναρτήσεις, δείχνοντας κριτήρια κυρτότητας που βασίζονται στο Θεώρημα Μέσης Τιμής. Ολοκληρώνουμε το κεφάλαιο με την Παράγραφο 6.4 και ακόμα ένα θέμα Αριθμητικής Ανάλυσης, και συγκεκριμένα με μια σύντομη εισαγωγή στη Μέθοδο του Νεύτωνα για τον υπολογισμό των ριζών μιας συνάρτησης.

6.1 Προσεγγιστικοί Υπολογισμοί

Θυμηθείτε ότι, σύμφωνα με τον ορισμό της παραγώγου, ισχύει

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

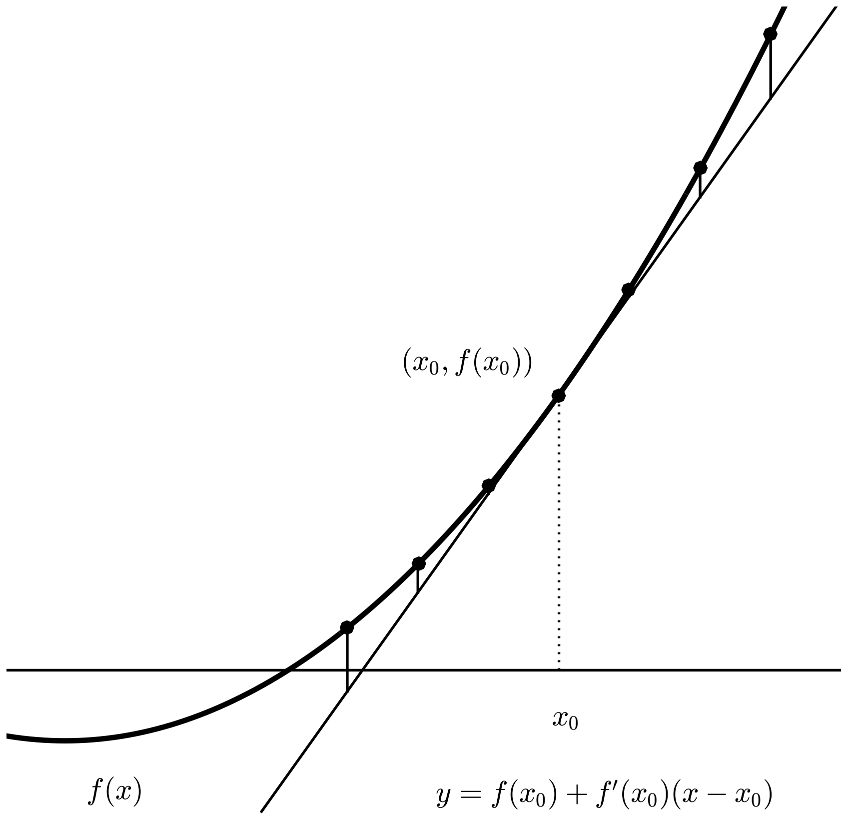
Επομένως, και σύμφωνα με τη διαισθητική ερμηνεία του ορίου, όταν το x είναι κοντά στο x_0, θα έχουμε, κατά προσέγγιση, ότι

$$f'(x_0) \simeq \frac{f(x) - f(x_0)}{x - x_0} \Leftrightarrow f(x) \simeq f(x_0) + f'(x_0)(x - x_0). \tag{6.1}$$

Η προσέγγιση αυτή, στη δεύτερη μορφή της, έχει μια απλή γεωμετρική ερμηνεία. Συγκεκριμένα, μας δείχνει ότι όταν το x είναι κοντά στο x_0, τότε και η τιμή της συνάρτησης στο $f(x)$, δηλαδή το αριστερό μέλος της προσέγγισης, είναι κοντά στην τιμή της γραμμικής συνάρτησης που περιγράφει την εφαπτομένη ευθεία, δηλαδή το δεξί μέλος. Δείτε το Σχήμα 6.1. Στο σχήμα έχουμε σχεδιάσει, για διάφορες τιμές του x, ευθύγραμμα τμήματα που ενώνουν τη συνάρτηση με την εφαπτομένη ευθεία σε αυτά τα σημεία.Το μήκος των ευθύγραμμων τμημάτων ισούται με το σφάλμα

$$E(x) = f(x) - f(x_0) - f'(x_0)(x - x_0) \tag{6.2}$$

της προσέγγισης. Παρατηρήστε ότι καθώς το x πλησιάζει το x_0, το σφάλμα γίνεται ολοένα και μικρότερο.



Σχήμα 6.1: Προσέγγιση της f περί το $(x_0, f(x_0))$ από την ευθεία $y = f(x_0) + f'(x_0)(x - x_0)$.

Επιπλέον, η προσέγγιση αυτή μας επιτρέπει να προσδιορίζουμε προσεγγιστικά τις τιμές μιας συνάρτησης με λιγότερο κόπο από ό,τι αν χρησιμοποιούσαμε την ίδια την τιμή της συνάρτησης. Δείτε το ακόλουθο παράδειγμα.

Παράδειγμα 6.1. (Προσδιορισμός τρίτης ρίζας) Σαν μια απλή εφαρμογή της προσέγγισης (6.1), θα προσδιορίσουμε προσεγγιστικά την τιμή της τρίτης ρίζας $f(x) = x^{\frac{1}{3}}$ των αριθμών x που εμφανίζονται στην πρώτη στήλη του ακόλουθου πίνακα. Η επακριβής τιμή εμφανίζεται στη δεύτερη στήλη. Η προσεγγιστική τιμή εμφανίζεται στην τρίτη στήλη. Οι τιμές βρίσκονται όλες γύρω από την τιμή $x_0 = 1000$, επομένως η προσέγγιση σε αυτή την περίπτωση είναι η ακόλουθη:

$$f(x) \simeq f(x_0) + f'(x_0)(x - x_0) = x_0^{\frac{1}{3}} + \frac{x - x_0}{3x_0^{\frac{2}{3}}} = 1000^{\frac{1}{3}} + \frac{x - 1000}{3 \times 1000^{\frac{2}{3}}} = 10 + \frac{1}{300}(x - x_0). \quad (6.3)$$

Στην τέταρτη στήλη εμφανίζουμε την τιμή $x - x_0$, και τέλος στην πέμπτη στήλη εμφανίζουμε το σφάλμα στην προσέγγιση, δηλαδή την τιμή $E(x)$ της (6.2).

x	$f(x) = x^{\frac{1}{3}}$	$f(x_0) + f'(x_0)(x - x_0)$	$x - x_0$	$E(x)$
900	9.6549	9.6667	-100	0.0118
990	9.9666	9.9667	-10	1.1×10^{-4}
999	9.9967	9.9967	-1	1.1×10^{-6}
1000	10	10	0	0
1001	10.0033	10.0033	1	1.1×10^{-6}
1010	10.0332	10.0333	10	1.1×10^{-4}
1100	10.3228	10.333	100	0.0115

Παρατηρήστε ότι, όσο πιο κοντά είναι το x στο x_0, τόσο πιο μικρό γίνεται το σφάλμα $E(x)$, όπως φαίνεται και από το Σχήμα 6.1.

Παρατηρήστε, επίσης, ότι οι πράξεις που χρειάστηκε να κάνουμε για να υπολογίσουμε τα στοιχεία της τρίτης στήλης ήταν πολύ λιγότερες. Πράγματι, παρατηρήστε ότι οποιαδήποτε και αν ήταν η συνάρτηση $f(x)$, σίγουρα θα αρκούσε να υπολογίσουμε μια φορά μόνο την παράγωγό της, στη θέση x_0, και μετά το κάθε σημείο υπολογίζεται με λίγες αριθμητικές πράξεις. Αντίθετα, για τον υπολογισμό των επακριβών τιμών, που εμφανίζονται στη δεύτερη στήλη, οι πράξεις που ενδέχεται να χρειαστούν είναι πολύ πιο σύνθετες. Στο παράδειγμά μας, χρειάζεται ο υπολογισμός μιας κυβικής ρίζας ανά στοιχείο του πίνακα.

Ασκήσεις

6.1. (Υπολογισμός των τιμών της εφαπτομένης) Επαναλάβετε το Παράδειγμα 6.1 για τη συνάρτηση $\tan x$. Συγκεκριμένα, υπολογίστε προσεγγιστικά τις τιμές $\tan x$ για τις ακόλουθες τιμές του x:

$$\pi/4 - 0.1, \quad \pi/4 - 0.01, \quad \pi/4 - 0.001, \quad \pi/4, \quad \pi/4 + 0.001, \quad \pi/4 + 0.01, \quad \pi/4 + 0.1.$$

Χρησιμοποιήστε μόνο την τιμή της $\tan x$ και της παραγώγου της στη θέση $x_0 = \pi/4$. Συγκρίνετε, με χρήση ενός αναλυτικού πίνακα, όπως και στο Παράδειγμα 6.1, τις τιμές που βρήκατε με τις ακριβείς.

6.2 Κανόνας L'Hôpital

Ορισμός 6.1. *(Απροσδιοριστίες)*

1. *Το όριο* $\lim_{x \to x_0} f(x)g(x)$ *εμφανίζει* ***απροσδιοριστία*** $0 \times \infty$ *αν* $\lim_{x \to x_0} f(x) = 0$ *και* $\lim_{x \to x_0} g(x) = \pm\infty$.
2. *Το όριο* $\lim_{x \to x_0} \frac{f(x)}{g(x)}$ *εμφανίζει* ***απροσδιοριστία*** $\frac{0}{0}$ *αν* $\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = 0$.
3. *Το όριο* $\lim_{x \to x_0} \frac{f(x)}{g(x)}$ *εμφανίζει* ***απροσδιοριστία*** $\frac{\infty}{\infty}$ *αν* $\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = \infty$.
4. *Το όριο* $\lim_{x \to x_0} (f(x) - g(x))$ *εμφανίζει* ***απροσδιοριστία*** $\infty - \infty$ *αν* $\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = \infty$.

Οι ίδιοι ορισμοί ισχύον και στην περίπτωση που τα παραπάνω όρια είναι πλευρικά ή στο $\pm\infty$.

Σε επόμενα κεφάλαια θα δούμε και άλλες μορφές απροσδιοριστίας.

Σε περιπτώσεις απροσδιοριστίας δεν μπορούμε να χρησιμοποιήσουμε τους κανόνες που έχουμε δει μέχρι τώρα για τον υπολογισμό ορίων. Δεν υπάρχει κάποιος γενικός κανόνας που να μας δίνει την τιμή ενός ορίου που εμφανίζει απροσδιοριστία, και κατά περίπτωση μπορεί το όριο να έχει οποιαδήποτε τιμή ή και να μην υπάρχει. Για παράδειγμα, δείτε την Άσκηση 6.2.

Παρατηρήστε, επίσης, ότι ορισμένα όρια εμφανίζουν διαφορετικά είδη απροσδιοριστίας, ανάλογα με τον τρόπο που τα γράφουμε. Για παράδειγμα, το όριο

$$\lim_{x \to 0} \frac{\sin^2}{x^2} = \lim_{x \to 0} \sin^2 x \times \frac{1}{x^2}$$

εμφανίζει απροσδιοριστία $\frac{0}{0}$ αν το γράψουμε με τον πρώτο τρόπο και απροσδιοριστία $0 \times \infty$ αν το γράψουμε σύμφωνα με τον δεύτερο τρόπο.

Μια τελευταία παρατήρηση που πρέπει να γίνει είναι ότι απροσδιοριστία έχουμε, για παράδειγμα, αν πολλαπλασιάσουμε κάτι που *τείνει* στο 0 με κάτι που *τείνει* στο άπειρο. Αν, όμως, πολλαπλασιάσουμε *το* 0 με κάτι που απλώς τείνει στο άπειρο, το γινόμενο είναι το 0, και δεν υπάρχει απροσδιοριστία.

Μπορούμε τώρα να δούμε τον αγαπημένο κανόνα όλων των φοιτητών, τον Κανόνα του L'Hôpital. Η χρησιμότητα του κανόνα είναι ότι μας επιτρέπει να αντιμετωπίζουμε απροσδιοριστίες της μορφής $\frac{0}{0}$ και της μορφής $\frac{\infty}{\infty}$.

Θεώρημα 6.1. ***(Κανόνας του L'Hôpital)*** *Έστω συναρτήσεις f, g για τις οποίες ισχύει*

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = 0, \qquad \lim_{x\to x_0} \frac{f'(x)}{g'(x)} = L.$$

Τότε έχουμε και

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = L.$$

Το αποτέλεσμα ισχύει και αν όλα τα όρια είναι πλευρικά ή είναι στο $\pm\infty$, αν το $L = \pm\infty$, και, τέλος, αν το κοινό όριο των f, g είναι το ∞ ή το $-\infty$.

Επειδή ο Κανόνας του L'Hôpital περιλαμβάνει πολλά σκέλη, δεν θα δούμε την απόδειξή του. Στη συνέχεια θα εξετάσουμε ορισμένα παραδείγματα σωστής και λάθος χρήσης του κανόνα. Σε επόμενα κεφάλαια θα δούμε αρκετά παραδείγματα ακόμα.

Παράδειγμα 6.2. **(Χρήση του Κανόνα του L'Hôpital)** Θα υπολογίσουμε το όριο

$$\lim_{x\to 0} \frac{\tan x - \sin x}{x + \sin x},$$

για το οποίο έχουμε απροσδιοριστία της μορφής $0/0$, με χρήση του Κανόνα του L'Hôpital. Πολύ απλά,

$$\lim_{x\to 0} \frac{\tan x - \sin x}{x + \sin x} = \lim_{x\to 0} \frac{\frac{1}{\cos^2 x} - \cos x}{1 + \cos x} = \frac{\frac{1}{1} - 1}{1 + 1} = 0.$$

Παράδειγμα 6.3. **(Λάθος χρήση του Κανόνα L'Hôpital)** Θα υπολογίσουμε το ακόλουθο όριο με δύο διαφορετικές μεθόδους. Παρατηρήστε ότι τόσο ο αριθμητής όσο και ο παρανομαστής τείνουν στο 0:

$$\lim_{x\to 1} \frac{x^3 + 3x - 4}{x^3 + x^2 - 2}.$$

Καταρχάς, έχουμε, με διπλή χρήση του Κανόνα του L'Hôpital:

$$\lim_{x\to 1} \frac{x^3 + 3x - 4}{x^3 + x^2 - 2} = \lim_{x\to 1} \frac{3x^2 + 3}{3x^2 + 2x} = \lim_{x\to 1} \frac{6x}{6x} = 1.$$

Εναλλακτικά, έχουμε, με μία μόνο χρήση του Κανόνα του L'Hôpital:

$$\lim_{x\to 1} \frac{x^3 + 3x - 4}{x^3 + x^2 - 2} = \lim_{x\to 1} \frac{3x^2 + 3}{3x^2 + 2x} = \frac{\lim_{x\to 1}(3x^2 + 3)}{\lim_{x\to 1}(3x^2 + 2x)} = \frac{6}{5}.$$

Το σωστό αποτέλεσμα είναι, βέβαια, το δεύτερο. Στην πρώτη λύση, ήταν λάθος η δεύτερη εφαρμογή του κριτηρίου, διότι τότε δεν υπήρχε απροσδιοριστία. Για να εφαρμόσουμε τον Κανόνα του L'Hôpital *πρέπει* να υπάρχει απροσδιοριστία.

Παράδειγμα 6.4. (Μη υπάρχον όριο) Σαν ένα άλλο παράδειγμα των προβλημάτων που μπορεί να κρύβει η χρήση του Κανόνα του L'Hôpital, θα υπολογίσουμε το ακόλουθο όριο, με δύο διαφορετικές μεθόδους:

$$\lim_{x\to 0}\frac{x^2\sin\left(\frac{1}{x}\right)}{\sin x}.$$

Παρατηρούμε, καταρχάς, πως

$$\lim_{x\to 0}\frac{x^2\sin\left(\frac{1}{x}\right)}{\sin x}=\lim_{x\to 0}\left(\frac{x}{\sin x}\right)\lim_{x\to 0}x\sin\left(\frac{1}{x}\right).$$

Σχετικά με το πρώτο όριο, παρατηρούμε πως είναι ίσο με 1, αφού $\lim\limits_{x\to 0}\frac{\sin x}{x}=1$ (δείτε το Παράδειγμα 3.12). Σχετικά με το δεύτερο όριο, στο Παράδειγμα 3.10 έχουμε υπολογίσει ότι είναι 0. Επομένως, με την πρώτη μέθοδο υπολογίζουμε πως το όριο είναι 0.

Σχετικά με τη δεύτερη μέθοδο, χρησιμοποιώντας τον Κανόνα του L'Hôpitalπροκύπτει πως

$$\begin{aligned}\lim_{x\to 0}\frac{x^2\sin\left(\frac{1}{x}\right)}{\sin x} &= \lim_{x\to 0}\frac{2x\sin\left(\frac{1}{x}\right)+x^2\cos\left(\frac{1}{x}\right)\left(-\frac{1}{x^2}\right)}{\cos x}\\ &= \lim_{x\to 0}\frac{2x\sin\left(\frac{1}{x}\right)}{\cos x}-\lim_{x\to 0}\frac{\sin\left(\frac{1}{x}\right)}{\cos x}.\end{aligned}$$

Από τα παραπάνω όρια, το πρώτο είναι εύκολο να δείξουμε ότι ισούται με 0, και πάλι χρησιμοποιώντας το Παράδειγμα 3.10. Το δεύτερο όμως όριο, δεν υπάρχει, διότι δεν υπάρχει το όριο της $\sin\left(\frac{1}{x}\right)$ (δείτε το Παράδειγμα 3.16). Επομένως, δεν υπάρχει και το όριο του αριστερού μέλους. Πράγματι, αν υποθέσουμε ότι το όριο αυτό υπάρχει, εύκολα μπορούμε να καταλήξουμε σε ένα άτοπο. Δείτε σχετικά την Άσκηση 3.24.

Επομένως, βρήκαμε δύο διαφορετικά αποτελέσματα. Το λάθος αποτέλεσμα είναι το δεύτερο: για να εφαρμόσουμε τον Κανόνα του L'Hôpital, *πρέπει* να υπάρχει το όριο του πηλίκου των παραγώγων. Αν δεν υπάρχει, ο κανόνας δεν εφαρμόζεται, και, όπως είδαμε και με αυτό το αντιπαράδειγμα, μπορεί κάλλιστα να υπάρχει το όριο του πηλίκου των αρχικών συναρτήσεων, όχι όμως και το όριο του πηλίκου των παραγώγων της.

Παράδειγμα 6.5. (Όριο $\lim\limits_{x\to 0}\frac{\sin x}{x}$) Τυπικά, μπορούμε να χρησιμοποιήσουμε τον Κανόνα του L'Hôpital για να γράψουμε

$$\lim_{x\to 0}\frac{\sin x}{x}=\lim_{x\to 0}\frac{\cos x}{1}=1.$$

Το συγκεκριμένο όριο έχει ήδη υπολογιστεί στο Παράδειγμα 3.12. Τυπικά, δεν έχουμε κάνει κάποιο λάθος. Όμως, στην εφαρμογή του Κανόνα, χρησιμοποιήσαμε την παράγωγο $(\sin x)'=\cos x$, που υπολογίσαμε στο Παράδειγμα 5.1 χρησιμοποιώντας το όριο που ζητούμε! Επομένως, αν και η συγκεκριμένη εφαρμογή του Κανόνα είναι σωστή, μπορεί να χρησιμεύσει μόνο ως επαλήθευση, και όχι για να αποδείξουμε ότι ισχύει το δοσμένο όριο.

Ασκήσεις

6.2. (Απροσδιοριστίες $0/0$) Παρατηρήστε ότι τα ακόλουθα όρια εμφανίζουν, και τα τρία, απροσδιοριστία $\frac{0}{0}$:

$$\lim_{x\to 0}\frac{\sin^2 x}{x},\qquad \lim_{x\to 0}\frac{\sin^2 x}{x^4},\qquad \lim_{x\to 0}\frac{\sin x}{x^2}.$$

Να υπολογίσετε τα παραπάνω όρια ή να δείξετε ότι δεν υπάρχουν. Να χρησιμοποιήσετε γνωστά τριγωνομετρικά όρια και τον ορισμό του ορίου, αλλά όχι παραγώγους.

6.3. (Χρήση Κανόνα L'Hôpital) Να υπολογίσετε το ακόλουθο όριο της Άσκησης 3.32 με χρήση του Κανόνα του L'Hôpital:

$$\lim_{x\to 0}\frac{\sin^3 x+\sin^2 x+\sin x}{x^3+x^2+x}.$$

6.3 Κυρτές Συναρτήσεις

Την έννοια της κυρτότητας μιας συνάρτησης την έχουμε ήδη συναντήσει στο Λύκειο. Σε αυτή την παράγραφο θα την ξαναδούμε σε μια διαφορετική μορφή, η οποία είναι πολύ πιο γενική. Συγκεκριμένα, ο ορισμό που θα δώσουμε, πρώτον, καλύπτει και μη παραγωγίσιμες συναρτήσεις και, δεύτερον, μπορεί να επεκταθεί εύκολα και σε συναρτήσεις πολλών μεταβλητών.

Ορισμός 6.2. ***(Κυρτές και κοίλες συναρτήσεις)*** *Έστω συνάρτηση $f : A \to \mathbb{R}$. Η συνάρτηση f καλείται* ***κυρτή στο διάστημα*** *I αν*

$$\forall x_0,\ x_1 \in I,\ \forall \theta \in [0,1], \qquad f((1-\theta)x_0+\theta x_1) \le (1-\theta)f(x_0)+\theta f(x_1).$$

Η συνάρτηση f καλείται ***κοίλη στο διάστημα*** *I αν η $-f$ είναι κυρτή στο διάστημα αυτό.*

Σύμβαση: *Στη συνέχεια, αν μια συνάρτηση ορίζεται ως κυρτή ή κοίλη σε κάποιο σύνολο I, θα υπονοείται ότι το σύνολο I είναι διάστημα.*

Ο ορισμός φαίνεται αρκετά σύνθετος, όμως η γεωμετρική του ερμηνεία είναι αρκετά απλή: μια συνάρτηση είναι κυρτή αν όλες οι χορδές (δηλαδή όλα τα ευθύγραμμα τμήματα που ξεκινούν από το γράφημα της συνάρτησης και καταλήγουν σ' αυτό) βρίσκονται πάνω από το γράφημα της συνάρτησης (δηλαδή για κάθε τετμημένη x η αντίστοιχη τεταγμένη επί της χορδής είναι μεγαλύτερη από την αντίστοιχη τεταγμένη επί του γραφήματος) ή το πολύ επί του γραφήματος (δηλαδή οι τεταγμένες είναι ίσες).

Στη συνέχεια, θα δείξουμε τη σχέση ανάμεσα στον παραπάνω ορισμό και τη γεωμετρική ερμηνεία που αναφέραμε. Δείτε το Σχήμα 6.2. Έστω δύο σημεία $(x_0, f(x_0))$, $(x_1, f(x_1))$ επί του γραφήματος, με $x_0, x_1 \in I$. Έστω η ευθεία που ενώνει τα δύο αυτά σημεία. Κατά τα γνωστά από τη γεωμετρία, η ευθεία αυτή έχει εξίσωση

$$y = f(x_0) + \frac{f(x_1)-f(x_0)}{x_1-x_0}(x-x_0).$$

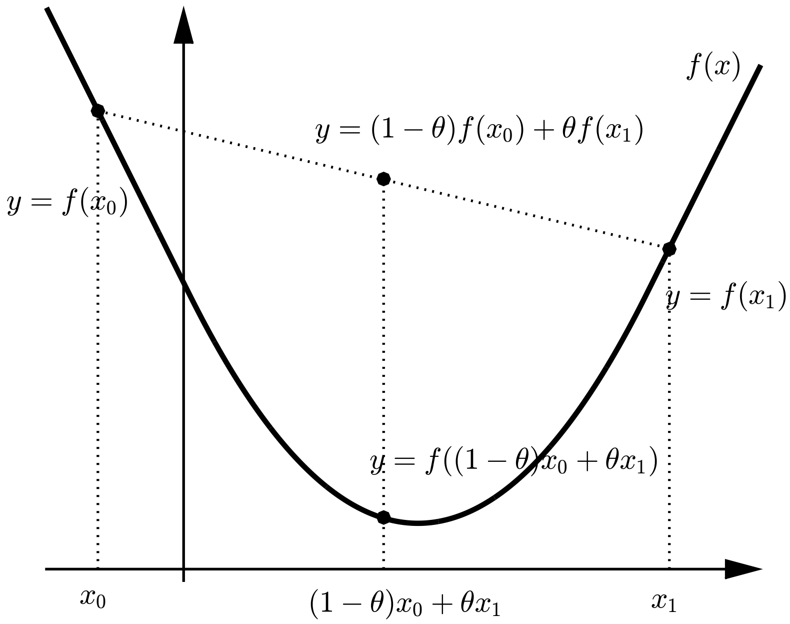
(Μπορείτε να θέσετε $x = x_1$ και $x = x_2$ για να επιβεβαιώσετε ότι πράγματι η ευθεία αυτή διέρχεται από τα $(x_0, f(x_0))$, $(x_1, f(x_1))$.)

Ακολούθως, παρατηρήστε πως τα σημεία επί του άξονα x που βρίσκονται μεταξύ του x_0 και του x_1 είναι ακριβώς τα $x = \theta x_0 + (1-\theta)x_0$ όπου $\theta \in [0,1]$. Πράγματι, όταν $\theta \in [0,1]$, τότε έχουμε $\theta \ge 0$ και $1-\theta \ge 0$, επομένως

$$x_0 = \theta x_0 + (1-\theta)x_0 \le \theta x_0 + (1-\theta)x_1 \le \theta x_1 + (1-\theta)x_1 = x_1.$$

Για $\theta = 0$ το $x = x_0$, για $\theta = 1$ το $x = x_1$, για $\theta = 1/2$ το $x = (x_0+x_1)/2$, κ.ο.κ.

Παρατηρήστε επίσης ότι η τεταγμένη y της ευθείας που αντιστοιχεί στην τετμημένη $x = (1 -$



Σχήμα 6.2: Γραφική απεικόνιση του ορισμού της κυρτότητας.

$\theta)x_0 + \theta x_1$ είναι η

$$\begin{aligned} y &= f(x_0) + \frac{f(x_1) - f(x_0)}{(1-\theta)x_0 + \theta x_1 - x_0}(x - x_0) \\ &= f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}((1-\theta)x_0 + \theta x_1 - x_0) \\ &= f(x_0) + \theta\frac{f(x_1) - f(x_0)}{x_1 - x_0}(x_1 - x_0) \\ &= (1-\theta)f(x_0) + \theta f(x_1). \end{aligned}$$

Επομένως, πράγματι η ανισότητα του ορισμού απαιτεί το ευθύγραμμο τμήμα, στο σημείο που αντιστοιχεί στο συγκεκριμένο θ, να είναι πάνω από το γράφημα της συνάρτησης στο αντίστοιχο σημείο ή το πολύ να συμπίπτει με αυτό. Ο ορισμός απαιτεί η ανισότητα να ισχύει για κάθε $\theta \in [0, 1]$, άρα πρέπει όλο το ευθύγραμμο τμήμα να είναι πάνω από το γράφημα ή, το πολύ, να εφάπτεται με αυτό. Πρέπει επίσης η ανισότητα να ισχύει και για κάθε $x_0, x_1 \in I$, άρα πράγματι ο ορισμός έχει τη γεωμετρική ερμηνεία που διατυπώσαμε.

Παρατηρήστε ότι στον ορισμό δεν γίνεται πουθενά αναφορά στην παράγωγο της f. Πράγματι, μπορεί μια συνάρτηση να είναι κυρτή σε ένα διάστημα ακόμα και αν δεν είναι παραγωγίσιμη παντού σε αυτό. Δείτε και μερικά από τα ακόλουθα παραδείγματα. Επομένως, ο ορισμός που δώσαμε είναι πιο γενικός από τον ορισμό που είδαμε στο Λύκειο.

Στη συνέχεια, θα δούμε μερικά παραδείγματα προσδιορισμού της κυρτότητας μιας συνάρτησης χρησιμοποιώντας τον ορισμό της κυρτότητας.

Παράδειγμα 6.6. (Οι γραμμικές συναρτήσεις είναι κυρτές και κοίλες) Έστω μια γραμμική συνάρτηση $f : \mathbb{R} \to \mathbb{R}$ με $f(x) = ax + b$ όπου οι σταθερές $a, b \in \mathbb{R}$. Παρατηρήστε πως

$$\begin{aligned} f((1-\theta)x_0 + \theta x_1) = a(1-\theta)x_0 + a\theta x_1 + b &= (1-\theta)(ax_0 + b) + \theta(ax_1 + b) \\ &= (1-\theta)f(x_0) + \theta f(x_1). \end{aligned}$$

Επομένως, η ανισότητα του ορισμού ισχύει με ισότητα, και η $f(x)$ είναι κυρτή. Επειδή όμως, αν η f είναι γραμμική, τότε είναι γραμμική και η $-f$, από τον ορισμό της κοιλότητας ισχύει ότι η f είναι *και* κοίλη.

Μπορούμε να δείξουμε (Άσκηση 6.5) ότι οι μόνες συναρτήσεις που είναι κυρτές *και* κοίλες είναι οι γραμμικές.

Εκ πρώτης όψεως το να είναι μια συνάρτηση και κυρτή και κοίλη φαίνεται λάθος. Το αποτέλεσμα όμως ισχύει διότι στον ορισμό της κυρτότητας χρησιμοποιήσαμε μια ανισοϊσότητα και όχι μια αυστηρή ανισότητα, διότι αυτό εξυπηρετούσε την ανάπτυξη της θεωρίας. Ουσιαστικά, οι γραμμικές συναρτήσεις έχουν όλες τις υπόλοιπες ιδιότητες των κυρτών συναρτήσεων, επομένως θέλουμε να τις συμπεριλάβουμε σε αυτές.

Παράδειγμα 6.7. (Η $f(x) = |x|$ είναι κυρτή) Έστω η συνάρτηση $f : \mathbb{R} \to \mathbb{R}$ με $f(x) = |x|$. Θα δείξουμε ότι η $f(x)$ είναι κυρτή. Πράγματι, η ανισότητα του ορισμού γίνεται

$$\begin{aligned} f((1-\theta)x_0 + \theta x_1) \leq (1-\theta)f(x_0) + \theta f(x_1) &\Leftrightarrow |(1-\theta)x_0 + \theta x_1| \leq (1-\theta)|x_0| + \theta|x_1| \\ &\Leftrightarrow |(1-\theta)x_0 + \theta x_1| \leq |(1-\theta)x_0| + |\theta x_1|, \end{aligned}$$

που ισχύει λόγω της τριγωνικής ανισότητας $|a + b| \leq |a| + |b|$.

Παράδειγμα 6.8. (Η x^2 είναι κυρτή) Θα δείξουμε ότι η $f(x) = x^2$ όπου $x \in \mathbb{R}$ είναι κυρτή παντού στο $\mathbb{R}$ με χρήση του ορισμού.

Η ανισότητα του ορισμού μας δίνει

$$\begin{aligned} & f((1-\theta)x_0 + \theta x_1) \leq (1-\theta)f(x_0) + \theta f(x_1) \\ & \Leftrightarrow (1-\theta)^2 x_0^2 + \theta^2 x_1^2 + 2\theta(1-\theta)x_0x_1 \leq (1-\theta)x_0^2 + \theta x_1^2 \\ & \Leftrightarrow x_0^2[(1-\theta)(1-(1-\theta))] + x_1^2[\theta(1-\theta)] + 2\theta(1-\theta)x_0x_1 \geq 0 \\ & \Leftrightarrow \theta(1-\theta)(x_0 + x_1)^2 \geq 0, \end{aligned}$$

που ισχύει για κάθε $x_0, x_1 \in \mathbb{R}$ και για κάθε $\theta \in [0, 1]$.

Όπως μπορείτε να διαπιστώσετε από τα παραπάνω παραδείγματα, ο ορισμός της κυρτότητας δεν είναι ιδιαίτερα εύχρηστος. Ευτυχώς, υπάρχει μια σειρά από αποτελέσματα που μας εξασφαλίζουν ότι μια συνάρτηση που έχει δημιουργηθεί χρησιμοποιώντας άλλες που είναι κυρτές, είναι και αυτή κυρτή. Δείτε τα παρακάτω.

Λήμμα 6.1. ***(Μέγιστο κυρτών συναρτήσεων)*** *Έστω συναρτήσεις $f, g : I \to \mathbb{R}$ κυρτές στο I. Τότε είναι κυρτή και η συνάρτηση $h = \max\{f, g\}$ που ορίζεται ως $h(x) = \max\{f(x), g(x)\}$.*

Απόδειξη: Θα επαληθεύσουμε τον ορισμό. Έστω, λοιπόν, δύο σημεία $x_0, x_1 \in I$ και ένα $\theta \in [0, 1]$. Έχουμε:

$$\begin{aligned} h((1-\theta)x_0 + \theta x_1) &= \max\{f((1-\theta)x_0 + \theta x_1), g((1-\theta)x_0 + \theta x_1)\} \\ &\leq \max\{(1-\theta)f(x_0) + \theta f(x_1), (1-\theta)g(x_0) + \theta g(x_1), \} \\ &\leq \max\{(1-\theta)f(x_0), (1-\theta)g(x_0)\} + \max\{\theta f(x_1), \theta g(x_1)\} \\ &= (1-\theta)\max\{f(x_0), g(x_0)\} + \max\{f(x_1), g(x_1)\} \\ &= (1-\theta)h(x_0) + \theta h(x_1). \end{aligned}$$

Στην πρώτη ανισότητα χρησιμοποιήσαμε το γεγονός ότι αν μεγαλώσουμε δύο αριθμούς, θα μεγαλώσει και το μέγιστό τους. Στη δεύτερη ανισότητα χρησιμοποιήσαμε το αποτέλεσμα της Άσκησης 6.7. ■

Πρόταση 6.1. ***(Γραμμικός συνδυασμός κυρτών συναρτήσεων)*** *Αν οι $f_1, \dots, f_n : I \to \mathbb{R}$ είναι κυρτές, τότε είναι κυρτή συνάρτηση και κάθε γραμμικός συνδυασμός τους με θετικούς συντελεστές*

$$g = \sum_{i=1}^{n} a_i f_i, \qquad a_i \geq 0.$$

Απόδειξη: Θα εφαρμόσουμε τον ορισμό της κυρτότητας. Παρατηρήστε πως

$$\begin{aligned} g((1-\theta)x_0 + \theta x_1) &= \sum_{i=1}^{n} a_i f_i((1-\theta)x_0 + \theta x_1) \\ &\leq \sum_{i=1}^{n} a_i((1-\theta)f_i(x_0) + \theta f(x_1)) \\ &= (1-\theta)\sum_{i=1}^{n} a_i f_i(x_0) + \theta \sum_{i=1}^{n} a_i f_i(x_i) \\ &= (1-\theta)g(x_0) + \theta \sum_{i=1}^{n} g(x_1), \end{aligned}$$

και η απόδειξη ολοκληρώθηκε. Παρατηρήστε ότι, για να ισχύει η ανισότητα, είναι απαραίτητο να έχουμε $a_i \geq 0$. ■

Λήμμα 6.2. ***(Κριτήρια κυρτότητας για παραγωγίσιμες συναρτήσεις)*** *Έστω $f : I \to \mathbb{R}$.*

1. *Αν η f είναι παραγωγίσιμη στο I και η f' είναι αύξουσα στο εσωτερικό του I, τότε η f είναι κυρτή στο I.*

2. *Αν η f είναι διπλά παραγωγίσιμη στο I και $f'' \geq 0$ στο εσωτερικό του I, τότε η f είναι κυρτή στο I.*

Απόδειξη:

1. Έστω δύο οποιαδήποτε $x_0, x_1 \in I$ και έστω $\theta \in [0,1]$ με $x_2 = (1-\theta)x_0 + \theta x_1$. Στα διαστήματα $[x_0, x_2]$ και $[x_2, x_1]$ ικανοποιούνται οι προϋποθέσεις του Θεωρήματος Μέσης Τιμής (Θεώρημα 5.5), επομένως υπάρχουν a, b τέτοια ώστε

$$f'(a) = \frac{f(x_2) - f(x_0)}{x_2 - x_0}, \qquad f'(b) = \frac{f(x_1) - f(x_2)}{x_1 - x_2}.$$

Επειδή όμως η f' είναι αύξουσα, έχουμε $f'(a) \leq f'(b)$, επομένως

$$\begin{aligned} &\frac{f(x_2) - f(x_0)}{x_2 - x_0} \leq \frac{f(x_1) - f(x_2)}{x_1 - x_2} \Leftrightarrow \frac{f(x_2) - f(x_0)}{\theta(x_1 - x_0)} \leq \frac{f(x_1) - f(x_0)}{(1-\theta)(x_1 - x_0)} \\ &\Leftrightarrow f(x_2) - f(x_0) \leq \theta f(x_1) - \theta f(x_0) \Leftrightarrow f(x_2) \leq (1-\theta)f(x_0) + \theta f(x_1) \\ &\Leftrightarrow f((1-\theta)x_0 + \theta x_1) \leq (1-\theta)f(x_0) + \theta f(x_1), \end{aligned}$$

και η απόδειξη ολοκληρώθηκε.

2. Αν η $f'' \geq 0$ στο εσωτερικό του I, τότε από την Πρόταση 5.5 η f' είναι αύξουσα, άρα από το προηγούμενο σκέλος η f είναι κυρτή. ■

Παράδειγμα 6.9. (Παράδειγμα προσδιορισμού κυρτότητας) Έστω η συνάρτηση

$$f(x) = x^3 - 10x^2 + 27x - 18,$$

ορισμένη στο διάστημα $[0, 7)$, των Παραδειγμάτων 5.9, 5.10, 5.11. Αφού

$$f'(x) = 3x^2 - 20x + 27, \quad f''(x) = 6x - 20,$$

προκύπτει πως η συνάρτηση είναι κυρτή στο διάστημα $[10/3, 7)$ και κοίλη στο διάστημα $[0, 10/3]$. Δείτε το Σχήμα 5.4.

Ασκήσεις

6.4. (Ορισμός κοιλότητας) Να δώσετε έναν ορισμό της κοίλης συνάρτησης χωρίς να χρησιμοποιήσετε την έννοια της κυρτότητας αλλά βάσει μιας ανισότητας ανάλογης του ορισμού της κυρτής συνάρτησης. Να δείξετε ότι ο νέος ορισμός είναι ισοδύναμος με αυτόν που έχει δοθεί.

6.5. (Κυρτές *και* κοίλες συναρτήσεις) Να δείξετε ότι οι μόνες συναρτήσεις που είναι και κυρτές και κοίλες σε ένα διάστημα I είναι οι γραμμικές, δηλαδή οι συναρτήσεις της μορφής $f(x) = ax + b$, όπου $a, b \in \mathbb{R}$.

6.6. Να δείξετε ότι αν μια συνάρτηση f είναι κυρτή σε ένα διάστημα I, τότε είναι αδύνατο να βρεθούν τρία σημεία $x_0 < x_1 < x_2$ εντός του I τέτοια ώστε $f(x_1) > f(x_0)$ και $f(x_1) > f(x_2)$.

6.7. (Ιδιότητες μέγιστου και ελάχιστου) Να δείξετε ότι ισχύουν οι παρακάτω ιδιότητες, όπου $a, b, c, d \in \mathbb{R}$.

$$\max\{a + b, c + d\} \leq \max\{a, c\} + \max\{b, d\},$$

$$-\min\{a, b\} = \max\{-a, -b\}.$$

6.8. [Σ/Λ, Π] (Μέγιστο κοίλων συναρτήσεων) Το μέγιστο δύο κοίλων συναρτήσεων είναι κοίλη συνάρτηση.

6.9. [Σ/Λ, Π] (Ελάχιστο κοίλων συναρτήσεων) Το ελάχιστο δύο κοίλων συναρτήσεων είναι κοίλη συνάρτηση.

6.10. [Σ/Λ, Π] (Ελάχιστο κυρτών συναρτήσεων) Το ελάχιστο δύο κυρτών συναρτήσεων είναι κυρτή συνάρτηση.

6.11. (Κυρτή σύνθεση) Έστω η συνάρτηση $h : [a, b] \to \mathbb{R}$, κυρτή στο $[a, b]$. Έστω η συνάρτηση $g : [c, d] \to \mathbb{R}$, αύξουσα και κυρτή στο $[c, d]$, με $h([a, b]) \subseteq [c, d]$. Να δείξετε ότι η σύνθεση $g \circ h$ είναι κυρτή παντού στο $[a, b]$. Μην υποθέσετε παραγωγισιμότητα των g, h!

6.12. [$\star\star$] (Κυρτότητα $\Rightarrow$ Συνέχεια) Να δείξετε ότι αν η g είναι κυρτή σε κάποιο ανοικτό διάστημα I τότε είναι και συνεχής σε αυτό. Να δείξετε επίσης ότι αν η g είναι κυρτή σε κάποιο όχι ανοικτό διάστημα I τότε ενδέχεται η g να μην είναι συνεχής.

6.13. (Κυρτά σύνολα) Ένα σύνολο S στο $\mathbb{R}^n$ λέγεται κυρτό όταν για οποιαδήποτε σημεία x_0, x_1 που ανήκουν στο σύνολο S, θα ανήκουν στο σύνολο και όλα τα σημεία που βρίσκονται επί του ευθυγράμμου τμήματος που ενώνει τα δύο αυτά σημεία, δηλαδή

$$\forall x_0, x_1 \in S,\ \forall \theta \in [0, 1], \qquad (1 - \theta)x_0 + \theta x_1 \in S.$$

Να δείξετε ότι, αν δύο σύνολα S_1, S_2 είναι κυρτά, τότε είναι και η τομή τους $S_1 \cap S_2$. Ποιες πατάτες καθαρίζονται πιο εύκολα, οι κυρτές ή οι μη κυρτές;

6.14. (Τοπικό ελάχιστο $=$ ολικό ελάχιστο) Να δείξετε ότι κάθε τοπικό ελάχιστο μιας κυρτής συνάρτησης είναι και ολικό ελάχιστο.

6.4 Μέθοδος του Νεύτωνα

Στην παράγραφο αυτή θα επανέλθουμε στο πρόβλημα του εντοπισμού των ριζών μιας συνάρτησης, που πρωτοείδαμε στην Παράγραφο 4.4, κάνοντας μια επιπλέον υπόθεση, και συγκεκριμένα ότι έχουμε στη διάθεσή μας την παράγωγο της συνάρτησης σε κάθε σημείο. Θα δούμε περιληπτικά τη λεγόμενη Μέθοδο του Νεύτωνα (Newton), γνωστή και ως μέθοδο Newton-Raphson.

Ας φανταστούμε, λοιπόν, ότι μας έχουν δοθεί δύο ρουτίνες, οι οποίες δέχονται ως είσοδο μια τιμή του x και δίνουν ως έξοδο τις αντίστοιχες τιμές της συνάρτησης $f(x)$ και της παραγώγου $f'(x)$ για εκείνο το x. Θέλουμε να εντοπίσουμε μια ρίζα της συνάρτησης, δηλαδή ένα σημείο x^* για το οποίο $f(x^*) = 0$. Για την ακρίβεια, επειδή οι υπολογιστές έχουν περιορισμένη ακρίβεια, αρκεί να βρούμε ένα σημείο x^* για το οποίο $|f(x^*)| < E_f$, όπου η ανοχή E_f είναι ένας πολύ μικρός θετικός αριθμός.

Θυμίζουμε ότι η αφελής προσέγγιση να σχεδιάσουμε το γράφημα της συνάρτησης και να εντοπίσουμε γραφικά το σημείο της ρίζας δεν είναι αποτελεσματική, διότι τότε θα πρέπει να υπολογίσουμε την τιμή της συνάρτησης σε πολλά σημεία. Το «παιχνίδι» που παίζουμε είναι πώς να εντοπίσουμε το σημείο της ρίζας με όσο το δυνατόν λιγότερες πράξεις.

Η βασική ιδέα της Μεθόδου του Νεύτωνα είναι η ακόλουθη. Έστω πως βρισκόμαστε σε ένα σημείο x_0 για το οποίο γνωρίζουμε την τιμή της συνάρτησης $f(x_0)$ καθώς και την τιμή της παραγώγου $f'(x_0)$. Φανταστείτε ένα μυρμήγκι που βρίσκεται πάνω σε ένα δρόμο (το γράφημα της συνάρτησης) του οποίου την κατεύθυνση σε σχέση με τον άξονα ανατολή-δύση (δηλαδή την παράγωγο) γνωρίζει, και θέλει μια εκτίμηση για το σημείο στο οποίο ο δρόμος αυτός θα διασταυρωθεί με μια μεγάλη λεωφόρο με κατεύθυνση ανατολή-δύση (τον άξονα των x) από την οποία ξέρει πόσο απέχει επί τον άξονα βορράς-νότος (αφού γνωρίζει την τιμή $f(x)$). Το μυρμήγκι είναι πολύ μικρό, και γι' αυτό δεν μπορεί να αντιληφθεί την καμπυλότητα του δρόμου (αυτό σημαίνει ότι δεν γνωρίζει την τιμή της δεύτερης παραγώγου $f''(x)$.) ή κάποιο άλλο στοιχείο πέραν αυτών που αναφέραμε. Η καλύτερη εκτίμηση που μπορεί να κάνει το μυρμήγκι είναι ότι το σημείο τομής των δύο δρόμων είναι αυτό όπου θα τέμνονταν οι δρόμοι αν ο δρόμος πάνω στον οποίο βρίσκεται το μυρμήγκι είναι *ευθύς*.

Η Μέθοδος του Νεύτωνα, λοιπόν, είναι η ακόλουθη: Ξεκινάμε από ένα σημείο x_0, για το οποίο γνωρίζουμε τα $f(x_0)$, $f'(x_0)$, και το οποίο αποτελεί την πρώτη μας εκτίμηση για το σημείο όπου βρίσκεται η ρίζα της συνάρτησης. Προσδιορίζουμε το σημείο x_1 στο οποίο θα έτεμνε η συνάρτηση τον άξονα των x αν ήταν ευθεία, δηλαδή το σημείο στο οποίο η εφαπτομένη στο x_0 τέμνει τον άξονα των x. Αυτό το σημείο είναι η νέα μας εκτίμηση για τη ρίζα της συνάρτησης. Επαναλαμβάνουμε τη διαδικασία από αυτό το σημείο, βρίσκοντας ένα νέο σημείο x_2, κ.ο.κ. Η διαδικασία θα μας δώσει τελικά μια ακολουθία η οποία (ελπίζουμε ότι) θα συγκλίνει στη ρίζα x^*.

Σχετικά με τον επακριβή υπολογισμό του σημείου x_{n+1} αν έχουμε το σημείο x_n, παρατηρήστε πως η εφαπτομένη ευθεία στο σημείο x_n είναι η

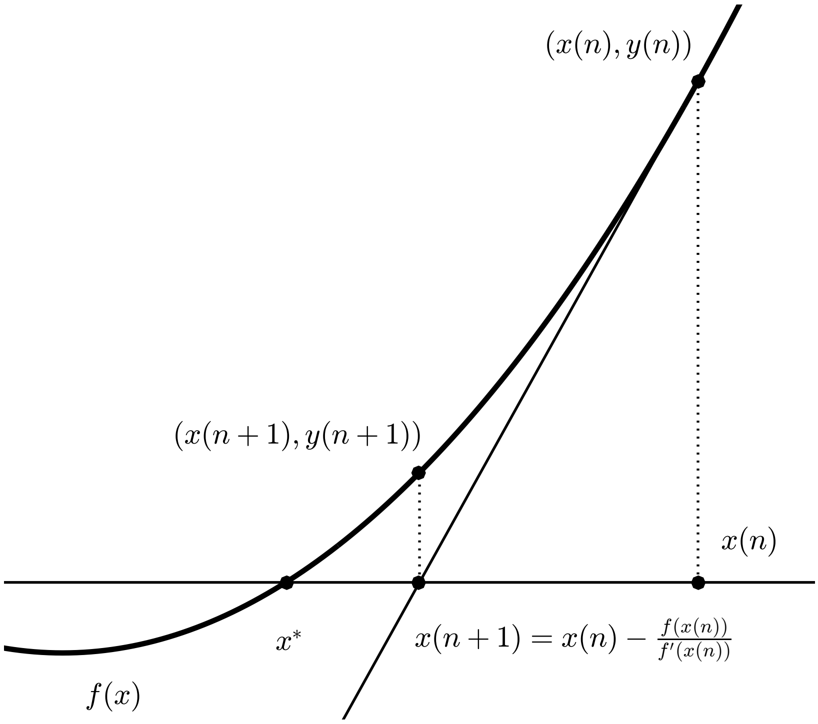
$$y = f(x_n) + f'(x_n)(x - x_n),$$

επομένως θέτοντας $y = 0$, $x = x_{n+1}$, και λύνοντας ως προς x_{n+1}, προκύπτει η βασική επανάληψη της μεθόδου

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Δείτε το Σχήμα 6.3.

Σε αντίθεση με τη Μέθοδο της Διχοτόμησης, η Μέθοδος του Νεύτωνα δεν εγγυάται ότι θα βρει μια ρίζα, ακόμα και αν αυτή υπάρχει. Ο λόγος είναι ότι ενδέχεται οι παράγωγοι που χρησιμοποιούμε να μας απομακρύνουν συνεχώς από τη ρίζα. Υπάρχει επίσης και η ειδική περίπτωση σε κάποιο σημείο η παράγωγος να είναι μηδενική, οπότε η εφαπτόμενη να μην τέμνει τον άξονα των x. Παρ' όλα αυτά, μπορεί να αποδειχτεί ότι



Σχήμα 6.3: Η επανάληψη της Μεθόδου του Νεύτωνα, που δίνει το σημείο x_{n+1} χρησιμοποιώντας το σημείο x_n.

1. Αν το σημείο x_0 είναι αρκετά κοντά στη ρίζα x^*, τότε σίγουρα η μέθοδος θα συγκλίνει σε αυτό.
2. Σε κάθε επανάληψη το σφάλμα, δηλαδή η απόσταση $x_{n+1} - x^*$, θα είναι το τετράγωνο του προηγούμενου, δηλαδή του $x_n - x^*$. Αν, για παράδειγμα, το σφάλμα σε μια επανάληψη είναι 10^{-3}, σε δύο μόνο επαναλήψεις ακόμα το σφάλμα είναι 10^{-12}!

Επομένως, η μέθοδος δεν συγκλίνει πάντα, και αυτό λαμβάνεται υπόψη στον ακόλουθο ψευδοκώδικα, αφού ορίζουμε ένα μέγιστο πλήθος επαναλήψεων και ελέγχουμε για την περίπτωση να είναι μηδενική η παράγωγος. Όταν όμως συγκλίνει, είναι πολύ γρήγορη.

```
/* NEWTON-RAPHSON METHOD */

function newton_raphson(f,f',x(0),M,Ef)

/* INPUT:                                                */
/* f: A function whose root is needed                    */
/* f': The derivative of the function                    */
/* x(0): Starting location of the iterations             */
/* M: Maximum number of iterations                       */
/* Ef:  Maximum tolerance in value of function           */
/*      at estimated root location.                      */

1:  n=0;   /* Initialization */

2:  WHILE n<M & |f(x(n))|>Ef & |f'(x(n))|>0,
```

```
    x(n+1)=x(n)-f(x(n))/f'(x(n));
    n=n+1;
END;

RETURN x(n) /* x(n) is an estimate of the root */
```

Δείτε τα επόμενα παραδείγματα. Παντού τα αποτελέσματα αναφέρονται με ακρίβεια 4 δεκαδικών ψηφίων, και έχουν προσδιοριστεί υλοποιώντας τον παραπάνω ψευδοκώδικα.

Παράδειγμα 6.10. **($f(x) = x - \cos x$, $x_0 = 0$)** Σαν ένα πρώτο παράδειγμα της μεθόδου, εξετάζουμε τη μέθοδο στην περίπτωση της συνάρτησης $f(x) = x - \cos x$, με αρχικό σημείο το $x_0 = 0$. Η μέθοδος ανακαλύπτει διαδοχικά τα ακόλουθα σημεία:

n	x_n	$f(x_n)$	$f'(x_n)$
0	0	-1.0000	1.0000
1	1.0000	0.4597	1.8415
2	0.7504	0.0189	1.6819
3	0.7391	0.0000	1.6736
4	0.7391	0.0000	1.6736
5	0.7391	0	1.6736

Μετά από μόλις 3 επαναλήψεις, η ρίζα έχει εντοπιστεί με ακρίβεια 4 δεκαδικών ψηφίων, ενώ μετά από συνολικά 5 επαναλήψεις η ρίζα έχει εντοπιστεί επακριβώς και με την ακρίβεια που χρησιμοποιεί ο υπολογιστής. Μπορείτε να συγκρίνετε αυτή την επίδοση με την αντίστοιχη επίδοση της Μεθόδου της Διχοτόμησης, που χρησιμοποιείται στο Παράδειγμα 4.9 για την ίδια συνάρτηση.

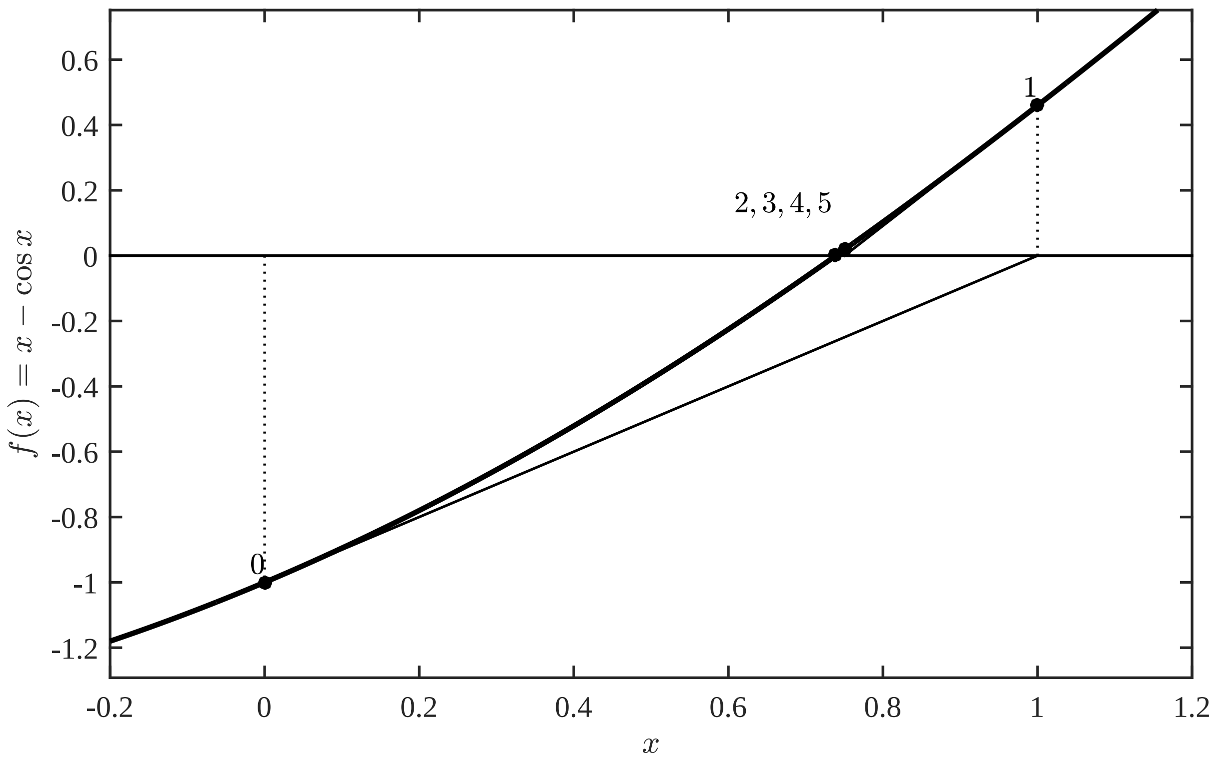
Στο Σχήμα 6.4 έχουμε απεικονίσει γραφικά την εξέλιξη της μεθόδου.

Παράδειγμα 6.11. **($f(x) = x - \cos x$, $x_0 = 4$)** Σαν ένα δεύτερο παράδειγμα, εξετάζουμε τη μέθοδο στην περίπτωση της συνάρτησης $f(x) = x - \cos x$, αλλά τώρα με αρχικό σημείο το $x_0 = 4$. Η μέθοδος ανακαλύπτει διαδοχικά τα ακόλουθα σημεία:

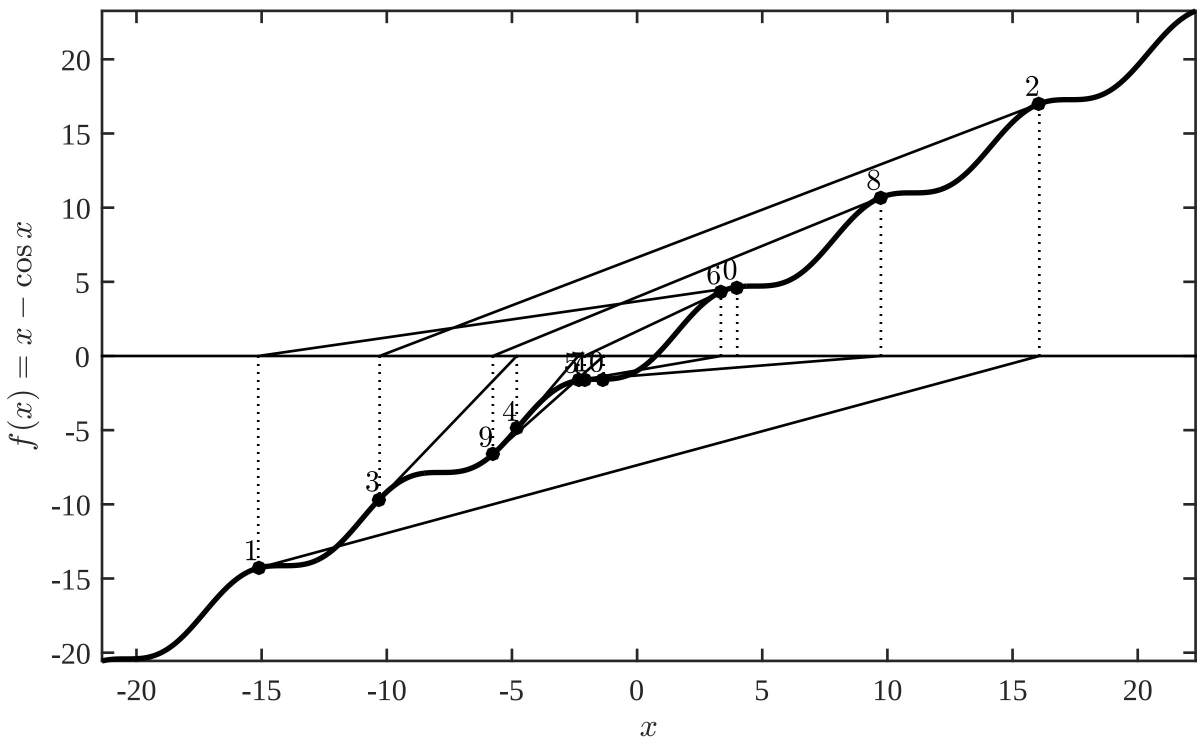
n	x_n	$f(x_n)$	$f'(x_n)$
0	4.0000	4.6536	0.2432
1	-15.1352	-14.2948	0.4581
2	16.0706	17.0056	0.6452
3	-10.2856	-9.6338	1.7584
4	-4.8068	-4.9011	1.9955
5	-2.3508	-1.6475	0.2891
6	3.3483	4.3270	0.7948
7	-2.0959	-1.5946	0.1347
8	9.7403	10.6909	0.6897
9	-5.7613	-6.6282	1.4985
10	-1.3381	-1.5687	0.0269

Η κατάσταση τώρα είναι πολύ διαφορετική απ' ότι στο προηγούμενο παράδειγμα. Επειδή το αρχικό μας σημείο ήταν αρκετά μακριά από τη ρίζα, η μέθοδος αδυνατεί να την προσεγγίσει, και πραγματοποιεί ταλαντώσεις γύρω από αυτή. Αν και οι λεπτομέρειες παραλείπονται, τελικά η μέθοδος συγκλίνει μετά από περίπου 30 επαναλήψεις καθώς, περίπου τυχαία, σε μια από αυτές επιλέγεται ένα σημείο αρκετά κοντά στη λύση.

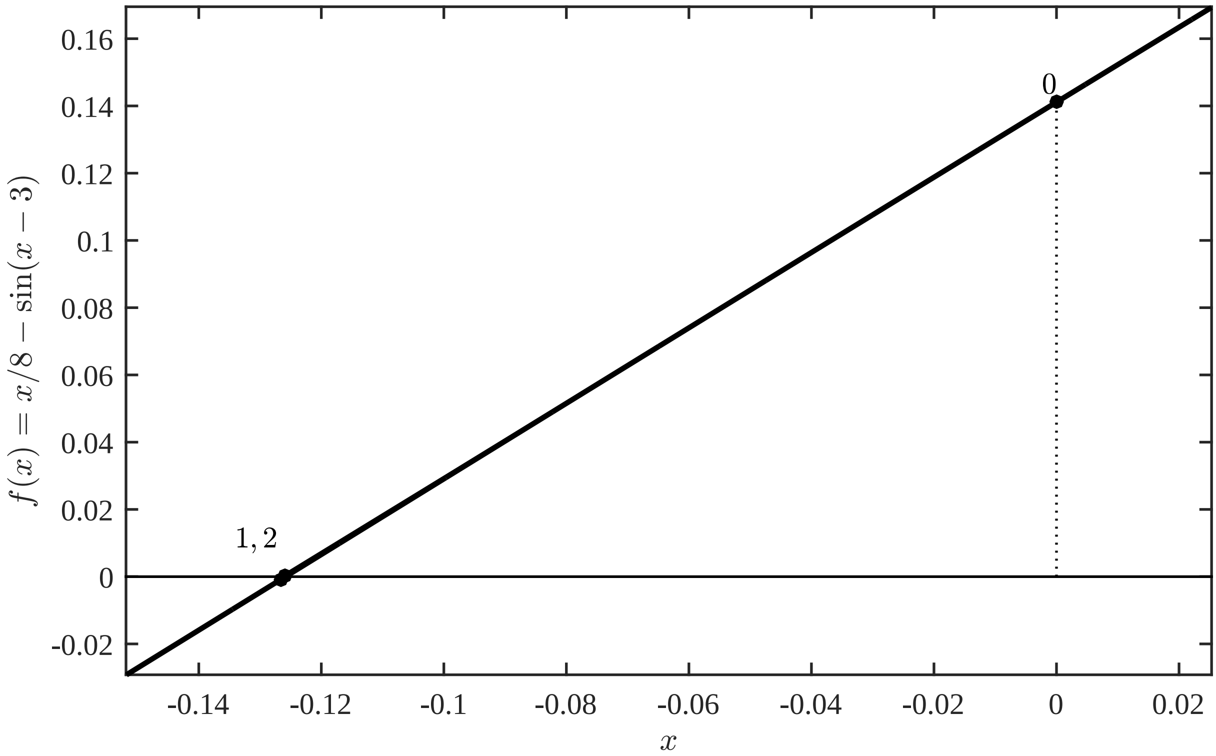
Στο Σχήμα 6.5 έχουμε απεικονίσει γραφικά την εξέλιξη της μεθόδου.



Σχήμα 6.4: Επαναλήψεις της Μεθόδου του Νεύτωνα για τη συνάρτηση $f(x) = x - \cos x$ με $x_0 = 0$ του Παραδείγματος 6.10. Η μέθοδος πρακτικά συγκλίνει μετά από 3 επαναλήψεις.



Σχήμα 6.5: Επαναλήψεις της Μεθόδου του Νεύτωνα για τη συνάρτηση $f(x) = x - \cos x$ με $x_0 = 4$ του Παραδείγματος 6.11. Η ρίζα δεν εντοπίζεται ακόμα και μετά από 10 επαναλήψεις.



Σχήμα 6.6: Επαναλήψεις της Μεθόδου του Νεύτωνα για τη συνάρτηση $f(x) = x/8 - \sin(x-3)$ με $x_0 = 0$ του Παραδείγματος 6.12. Η ρίζα εντοπίζεται πρακτικά μετά από 3 επαναλήψεις.

Παράδειγμα 6.12. ($f(x) = x/8 - \sin(x-3)$, $x_0 = 0$) Σαν ένα τελευταίο παράδειγμα, εξετάζουμε τη μέθοδο στην περίπτωση της συνάρτησης $f(x) = x/8 - \sin(x-3)$, αλλά τώρα με αρχικό σημείο το $x_0 = 0$. Η μέθοδος ανακαλύπτει διαδοχικά τα ακόλουθα σημεία:

n	x_n	$f(x_n)$	$f'(x_n)$
0	0	0.1411	1.1150
1	-0.1266	-0.0008	1.1249
2	-0.1259	-0.0000	1.1249
3	-0.1259	0.0000	1.1249

Η σύγκλιση επιτυγχάνεται πρακτικά μετά από 3 επαναλήψεις. Μπορείτε να συγκρίνετε την ταχύτητα σύγκλισης εδώ με την ταχύτητα σύγκλισης της Μεθόδου Διχοτόμησης, που εφαρμόζεται για την ίδια συνάρτηση στο Παράδειγμα 4.9. Παρατηρήστε ότι η μέθοδος εντοπίζει μια ρίζα της συνάρτησης διαφορετική από αυτές που εντόπισε η Μέθοδος της Διχοτόμησης.

Στο Σχήμα 6.6 έχουμε απεικονίσει γραφικά την εξέλιξη της μεθόδου.

Ασκήσεις

6.15. [$\star\star$] (Υλοποίηση Μεθόδου Νεύτωνα) Υλοποιήστε τη Μέθοδο του Νεύτωνα σε μια γλώσσα προγραμματισμού της προτίμησής σας. Η ρουτίνα που θα δημιουργήσετε πρέπει, κατ' ελάχιστον, να επιστρέφει αριθμητικά δεδομένα για κάθε επανάληψη της μεθόδου και την τελική εκτίμηση για τη ρίζα της δοσμένης συνάρτησης, δεν είναι όμως απαραίτητο να παράγει κάποιο σχήμα. Τα ορίσματα εισόδου πρέπει να περιλαμβάνουν τουλάχιστον το αρχικό σημείο όπου υπολογίζεται η τιμή της συνάρτησης, το μέγιστο πλήθος επαναλήψεων, και την ανοχή

E_f στην τιμή της συνάρτησης. (Επομένως, η ρουτίνα θα διακόπτεται όταν βρεθεί x_n με $|f(x_n) \leq E_f$.) Δεν είναι απαραίτητο να δίνεται ως όρισμα η συνάρτηση της οποία ζητείται η ρίζα και η παράγωγός της. Επομένως, μπορείτε για κάθε συνάρτηση της οποίας τη ρίζα θέλετε να υπολογίσετε να τροποποιείτε ανάλογα τον κώδικά σας.

6.16. (Αριθμητικό παράδειγμα) Χρησιμοποιήστε την υλοποίηση της Άσκησης 6.15 για να εντοπίσετε μια ρίζα της συνάρτησης $f(x) = 2 - \cos x + x^3$ με ακρίβεια τεσσάρων δεκαδικών ψηφίων.

6.17. [$\star$] (Υπολογισμός τόξου εφαπτόμενης) Χρησιμοποιήστε την υλοποίηση της Άσκησης 6.15 για να εντοπίσετε την τιμή του τόξου εφαπτόμενης arctan 1 με ακρίβεια τεσσάρων δεκαδικών ψηφίων. Μην χρησιμοποιήσετε στην υλοποίησή σας έτοιμη ρουτίνα που να υπολογίζει το τόξο εφαπτομένης. (Η άσκηση δεν έχει νόημα αν υποθέσουμε ότι υπάρχει τέτοια ρουτίνα στη διάθεσή μας.) Παρατηρήστε ότι δεν χρειάζεται να εκτελέσουμε τη μέθοδο του Νεύτωνα για να βρούμε την απάντηση. Πράγματι, $\tan \pi/4 = 1 \Rightarrow \arctan 1 = \pi/4$. Ο ουσιαστικός στόχος της άσκησης είναι να επιβεβαιώσουμε ότι η σύγκλιση γίνεται στο σημείο που αναμένουμε!

6.18. [$\star\star$] (Αποτυχία της Μεθόδου του Νεύτωνα) Σχεδιάστε μια συνάρτηση τέτοια ώστε αν εκτελέσουμε για αυτήν τη συνάρτηση τη Μέθοδο του Νεύτωνα με κατάλληλα επιλεγμένο αρχικό σημείο x_0, η μέθοδος θα ταλαντώνεται επ' άπειρον μεταξύ αυτού του σημείου και ενός άλλου, x_1, παρότι υπάρχει ρίζα ανάμεσά τους.

6.5 Περαιτέρω Μελέτη

Κυρτότητα Η έννοια της κυρτότητας είναι κεφαλαιώδους σημασίας στα εφαρμοσμένα μαθηματικά και αποτελεί αντικείμενο τρέχουσας έρευνας. Ο λόγος είναι ότι αν θέλουμε να βρούμε το σημείο στο οποίο ελαχιστοποιείται μια συνάρτηση πολλών μεταβλητών, τότε το πρόβλημά μας είναι κατά βάση εύκολο αν η συνάρτηση είναι κυρτή, και δύσκολο αν η συνάρτηση δεν είναι κυρτή. Είναι απλό να καταλάβουμε το λόγο που συμβαίνει αυτό στην ειδική περίπτωση των συναρτήσεων μίας μεταβλητής. Πράγματι, όπως είναι διαισθητικά αναμενόμενο, μια κυρτή συνάρτηση θα ελαχιστοποιείται σε ένα μόνο σημείο, ή έστω σε ένα διάστημα. Επομένως, κάθε τοπικό ελάχιστο είναι και ολικό. Επιπλέον, εξετάζοντας τη συμπεριφορά της συνάρτησης γύρω από ένα σημείο μπορούμε να μαντέψουμε προς τα πού πρέπει να μετακινηθούμε για να βρούμε το σημείο (ή διάστημα) ελάχιστου. Αυτές οι ιδιότητες δεν ισχύουν στην περίπτωση που η συνάρτηση δεν είναι κυρτή. Οι παραπάνω συλλογισμοί επεκτείνονται και στην περίπτωση συναρτήσεων πολλών μεταβλητών.

Προβλήματα ελαχιστοποίησης συναρτήσεων εμφανίζονται πολύ συχνά στις εφαρμοσμένες επιστήμες, καθώς πολλά προβλήματα μπορούν να αναχθούν σε αυτά, ακόμα και κάποια που είναι φαινομενικά άσχετα με την ελαχιστοποίηση μιας συνάρτησης, όπως η βελτιστοποίηση του πρωτοκόλλου TCP/IP, η επιλογή των ισχύων με τις οποίες εκπέμπουν τα κινητά τηλέφωνα σε ένα δίκτυο κινητής τηλεφωνίας, η επιλογή των παραμέτρων της λειτουργίας ενός δικτύου διανομής ηλεκτρικής ενέργειας, και η επιλογή των τιμών των στοιχείων σε ένα ηλεκτρικό κύκλωμα (για παράδειγμα, οι νόμοι του Kirchhoff μπορούν να προκύψουν ως λύση ενός προβλήματος βελτιστοποίησης μιας συνάρτησης, που συχνά είναι κυρτή). Επομένως, αν η συνάρτηση που πρέπει να ελαχιστοποιηθεί είναι κυρτή, ή μπορεί να γίνει κυρτή με κάποια κατάλληλη προσέγγιση ή κάποιο μετασχηματισμό, τότε το πρόβλημα είναι πολύ πιο εύκολο να λυθεί από τη γενική περίπτωση.

Λόγω αυτής της σπουδαιότητας της έννοιας της κυρτότητας, κάναμε μια μικρή εισαγωγή σε αυτήν στην Παράγραφο 6.3. Παρατηρήστε ότι τυπικά η σύνδεση της παραγράφου με το Κεφάλαιο 5 είναι σχετικά μικρή, και βασίζεται στην επίκληση του Θεωρήματος Μέσης Τιμής στην απόδειξη του Λήμματος 6.2.

Περισσότερα για προβλήματα κυρτής βελτιστοποίησης μπορείτε να μελετήσετε στα κλασικά βιβλία των Boyd και Vandenberghe [BOYD] (που διατίθεται στο διαδίκτυο) και του Μπερτσεκά [BERT].

Μέθοδος του Νεύτωνα Στην πράξη, η Μέθοδος του Νεύτωνα μπορεί να συνδυαστεί με άλλες μεθόδους (όπως η Μέθοδος της Διχοτόμησης) που πλησιάζουν αρκετά κοντά στη ρίζα, την οποία στη συνέχεια αναλαμβάνει να εντοπίσει ακριβώς και με μεγάλη ταχύτητα η Μέθοδος του Νεύτωνα. Μπορούμε έτσι να συνδυάσουμε τα καλύτερα στοιχεία της κάθε μεθόδου.

Η ιδέα να χρησιμοποιήσουμε την παράγωγο για να προσεγγίσουμε τη ρίζα της συνάρτησης μπορεί να επεκταθεί, με τη χρησιμοποίηση παραγώγων μεγαλύτερης τάξης. Σε αυτή την περίπτωση η προσέγγιση είναι ακόμα ταχύτερη.

Η Μέθοδος του Νεύτωνα, πάντως, δεν είναι η μόνη με την οποία μπορούμε να εντοπίσουμε τις ρίζες μιας συνάρτησης. Το πρόβλημα είναι πολύ παλιό, και υπάρχει μια μεγάλη σειρά από πολύ όμορφες μεθόδους, καθεμία εκ των οποίων κάνει διαφορετικές παραδοχές για τη συνάρτηση της οποίας ψάχνουμε τις ρίζες. Μπορείτε να δείτε περισσότερα στα βιβλία Αριθμητικής Ανάλυσης που αναφέρουμε στη βιβλιογραφία [FORE], [FORG], [PRES], [AKPI].