Περιοχές ἐμπιστοσύνης.

Ἡ προσέγγιση ποὺ προκύπτει ἀπὸ τὴν ἀνάπτυξη Taylor ὅταν λαμβάνουμε ὑπόψιν μέχρι καὶ τετραγωνικοὺς ὅρους δηλαδή:

$$f(x+h) - f(x) \approx h^T \nabla f(x) + \frac{1}{2} h^T \nabla^2 f(x) h \equiv q(h)$$

Εἶναι ἀκριβὴς σὲ μιὰ μικρὴ περιοχὴ $h \in \Omega$ γύρω ἀπὸ τὸ x.
Ἡ περιοχὴ αὐτὴ ὀνομάζεται περιοχὴ ἐμπιστοσύνης. Τὸ πρόβλημα ποὺ πρέπει νὰ ἀντιμετωπισθεῖ εἶναι πλέον βελτιστοποίηση μὲ περιορισμό.

$$\min_{h \in \Omega} q(h)$$

Συνήθως ὁ περιορισμὸς $h \in \Omega$ διατυπώνεται ὡς $\|h\| \leq R$ καὶ ἀντιστοιχεῖ σὲ μιὰ σφαιρικὴ περιοχὴ ἀκτίνας R.

Γράφοντας $q(h) = \frac{1}{2} h^T B h + g^T h$
ἔχουμε:

$$\min_h q(h) \quad \text{ὅπου} \quad \|h\| \leq R$$

Υπάρχουν (ὅπως καὶ στὴν περίπτωση τῆς γραμμικῆς ἀναζήτησης) δύο τακτικές.

α) Ἀκριβὴς ἐπίλυση τοῦ προβλήματος
β) Προσεγγιστικὴ ἐπίλυση.

Ἐὰν ἡ ἀκρίβεια τῆς προσέγγισης $f(x+h) - f(x) = q(h)$ γιὰ τὸ προκύπτον βῆμα h εἶναι καλή, μεγαλώνουμε τὴν ἀκτίνα R τῆς περιοχῆς ἐμπιστοσύνης, ἄλλως τὴν ἐλαττώνουμε.

Ακριβής ἐπίλυση.

Μποροῦμε νά ἀκολουθήσουμε τήν παρακάτω διαδικασία.
Ἡ ἐπίλυση min $q(h)$ χωρίς περιορισμούς ἐπιλέγεται
ὡς $h = h_N \equiv -B^{-1}g$
· Ἐάν $\|h_N\| \leq R$ τότε τό πρόβλημα ἔχει ἐπιλυθεῖ.
Ἐάν $\|h_N\| > R$ τότε χρησιμοποιοῦμε μιά μέθοδο ποινῆς.

$$Q(h) = q(h) + \frac{1}{2}\lambda \left(\|h\| - R\right)^2 \qquad \lambda > 0$$

Παραγωγίζοντας ἔχουμε: $(\nabla Q = 0)$

$$\left[B + \lambda\left(1 - \frac{R}{\|h\|}\right) I\right] h = -g$$

Στό ὅριο ὅπου τό $\lambda \to \infty$ τότε τό $\|h\| \to R$
ὁπότε τό γινόμενο $\lambda\left(1 - \frac{R}{\|h\|}\right)$ δέν εἶναι καθορισμένο.
Το γινόμενο δέν μπορεῖ νά εἶναι μηδέν
γιατί τότε τό προκύπτον $h = h_N$ δέν ἱκανοποιεῖ
τό $\|h_N\| = R$.
Ἐπίσης τό γινόμενο δέν μπορεῖ νά εἶναι ἄπειρο
γιατί τότε $\|h\| = 0$.
Ἀντιμετωπίζουμε λοιπόν τό γινόμενο αὐτό ὡς παράμετρο καί ἔχουμε νά λύσουμε τό πρόβλημα:

$$(B + aI)\, h = -g \quad \text{ἤ} \quad h = h(a) = -(B + aI)^{-1} g$$

καί ὑπολογίζουμε τό $\underline{a}$ γιά τό ὁποῖο ἰσχύει

$$\|h(\alpha)\|^2 = R^2 \quad \text{ἤ} \quad g^T (B + aI)^{-2} g = R^2$$

Προσεγγιστική ἐπίλυση.

Σημεῖο Newton: Ἡ λύση χωρὶς περιορισμούς, δηλαδὴ τὸ σημεῖο $h_N = -B^{-1}g$ ὀνομάζεται σημεῖο Newton.

Σημεῖο Cauchy: Τὸ σημεῖο ποὺ ἐλαχιστοποιεῖ τὴν $q(h) = \frac{1}{2}h^T B h + g^T h$ κατὰ μῆκος τῆς διεύθυνσης g, δηλαδὴ τὸ

$$h_C = -\left(\frac{g^T g}{g^T B g}\right) g$$

λέγεται σημεῖο Cauchy.

Ὁ ὑπολογισμός του ἔχει ὡς ἑξῆς.
Θέτουμε $h = \lambda g$ ὁπότε:

$$q(\lambda) = \frac{1}{2}\lambda^2 g^T B g + \lambda g^T g$$

Παραγωγίζοντας ὡς πρὸς λ καὶ θέτοντας $\frac{d}{d\lambda} q(\lambda) = 0$ ὑπολογίζουμε τὸ λ:

$$\lambda g^T B g + g^T g = 0 \Rightarrow \lambda = -\frac{g^T g}{g^T B g}$$

καὶ

$$h_C = \lambda g = -\left(\frac{g^T g}{g^T B g}\right) g$$

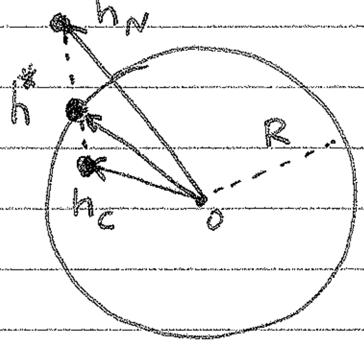
Παρατηροῦμε ὅτι ἡ ἀκριβὴς λύση προέρχεται ἀπὸ τήν:

$$(B + aI)h = -g \quad (\text{μὲ } |h| = R)$$

Γιὰ $\alpha = 0$ $h = h_N$
Γιὰ $\alpha \gg$ τῶν ἰδιοτιμῶν τοῦ B $h = -\frac{1}{\alpha} g$
εἶναι δηλαδὴ στὴν κατεύθυνση τῆς g καὶ ἄρα τὸ καλύτερο σημεῖο θὰ εἶναι τὸ $h = h_C$

Ἐπιβάλλοντας τὸν περιορισμὸ $|h| = R$ τὸ
α θὰ πάρει κάποια ἐνδιάμεση τιμή.

Τὸ βέλτιστο σημεῖο h^* θὰ μπορέσουμε νὰ τὸ προσεγγίσουμε στὸ ἐπίπεδο ποὺ ὁρίζεται ἀπὸ τὰ δύο ἀνύσματα h_c καὶ h_N.



Ἐὰν ἑνώσουμε τὸ σημεῖο Cauchy μὲ τὸ σημεῖο Newton τὸ σημεῖο τομῆς τῆς εὐθείας αὐτῆς μὲ τὴν σφαίρα ἀκτίνας R εἶναι μιὰ προσέγγιση τῆς βέλτιστης λύσης h^*.

Ἐὰν $|h_N| < R$ τότε $h^* = h_N$

Ἐὰν $|h_N| > R$ καὶ $|h_c| > R$ $\quad h^* = -R\dfrac{g}{|g|}$

Ἐὰν $|h_N| > R$ καὶ $|h_c| < R$ τότε διαλέγουμε τὸ h^* μὲ τὴν παραπάνω διαδικασία ὅπως φαίνεται στὸ σχῆμα.

Ἀποδεικνύεται ὅτι ἡ $\varphi(h)$ εἶναι φθίνουσα καθόλη τὴν διαδρομὴ ἀπὸ τὸ $h_c \rightarrow h_N$.

Σημεῖο DOG-LEG:

Ἡ διαδρομὴ $O \rightarrow h_c \rightarrow h_N$ τέμνει τὴν σφαίρα (O, R) στὸ h^* τὸ ὁποῖο φέρει τὸ ὄνομα "dog-leg". (Powell).

Μέθοδος Newton.

Βασίζεται στό τετραγωνικό μοντέλο:

$$f(x+h) \approx q(h) = f(x) + h^T g(x) + \frac{1}{2} h^T B h$$

όπου $g(x) = \nabla f(x)$, $B(x) = \nabla^2 f(x)$

Ἡ λύση τοῦ $\min_h q(h)$ δίδει: $h^* = -B^{-1} g$

Η $\underline{\underline{k}}^{οστή}$ ἐπανάληψη τῆς Newton ἔχει ὡς ἑξῆς:

Δοθέντων τῶν $x^{(k)}$, $g^{(k)}$, $B^{(k)}$

1) Υπολογίζεται τό: $h^{*(k)} = -B^{-1(k)} g^{(k)}$

2) Γραμμική αναζήτηση: $\min_{\lambda > 0} f(x^{(k)} + \lambda h^{(k)})$

πού ἐντοπίζει τήν λύση λ^*

3) $x^{(k+1)} = x^{(k)} + \lambda^* h^{(k)}$

$g^{(k+1)} = \nabla f(x^{(k+1)})$

$B^{(k+1)} = \nabla^2 f(x^{(k+1)})$

Εάν τό $h^{(k)}$ δέν είναι φθίνουσα κατεύθυνση (δηλαδή εάν $g^{(k)T} B^{-(k)} g^{(k)} < 0$) τό βήμα (2) δέν προσφέρει πρόοδο.

Πρέπει νά βρεθεί κάποια άλλη κατεύθυνση $h^{(k)}$ πού νά είναι φθίνουσα. Υπάρχουν διάφοροι τρόποι. Οι προκύπτουσες μέθοδοι ὀνομάζονται παραλλαγμένες μέθοδοι Newton (Modified Newton)

-Ἐὰν $h_N = -B^{-1}g$ εἶναι φθίνουσα διεύθυνση,

δηλαδὴ ἐὰν $g^T h_N < 0$ ἢ $g^T B^{-1} g > 0$

Δὲν συμβαίνει ὅτι ὁ B-πίνακας εἶναι θετικὰ ὁρισμένος.
-Ἐὰν ὅμως ὁ B εἶναι θετικὰ ὁρισμένος τότε ἡ κατεύθυνση $h_N = -B^{-1}g$ εἶναι φθίνουσα.

-Ἐὰν $g^T h_N > 0$, δηλαδὴ ἡ h_N δὲν εἶναι φθίνουσα τότε τροποποιοῦμε τὸν πίνακα B ὡς ἑξῆς:

α) $B \to B' = B + \lambda I$ $\lambda > 0$, μὲ κατάλληλο $\underline{\lambda}$ ἔτσι ὥστε ὁ B' νά γίνει θετικὰ ὁρισμένος.

β) Διαγωνοποιοῦμε τὸν $B = Q^T \Lambda Q$, Λ διαγώνιος, Q ὀρθογώνιος.
-Ἐὰν κάποιες ἰδιοτιμὲς εἶναι ἀρνητικὲς τὶς ἀντικαθιστοῦμε μὲ ἕνα μικρὸ θετικὸ ἀριθμό, σχηματίζοντας ἔτσι ἕναν πίνακα Λ'
$B \to B' = Q^T \Lambda' Q$

γ) Παραγοντοποίηση Cholesky: $B = LDL^T$
ὅπου L κάτω τριγωνικὸς καὶ D διαγώνιος.

$B \to B' = LD'L^T$ ὅπου ὁ διαγώνιος πίνακας D' εἶναι ὁ D διορθωμένος ἔτσι ὥστε τυχὸν ἀρνητικοὶ ἀντικαθίστανται μὲ μικροὺς θετικοὺς ἀριθμούς.