Corrections to *Numerical Optimization*, Second Edition  
Published August 2006  
(Last updated May 17, 2008)

1. p. 5, line −11. “from a a finite” → “from a finite”
2. p. 9, line 18. “n the 1940s” → “in the 1940s”
3. p. 23, line −5. “∇f” → “∇²f”
4. p. 25, line 1. “...,is” → “is, respectively, (6.25) and”
5. p. 26, line 8. “positive definite pₖ” → “positive definite”
6. p. 32, line 8. “k = 0, 1, . . .” → “k = 1, 2, . . .”
7. pp. 34-35, Figures 3.4 and 3.5. “desired slope” → “minimum acceptable slope”
8. p. 40, line −9. “will be able” → “will not be able”
9. p. 49, line 15. “For a proof this result” → “For a proof of this result”
10. p. 49, line 15. “For problems in which ∇f* is close to singular” → “For problems in which ∇²f(x*) is close to singular”
11. p. 55, Example 3.2. Replace formula (3.52) by

\[
L = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
2 & 4 & 1 & 0 \\
3 & 5 & 1 & 1
\end{bmatrix}, \quad B = \begin{bmatrix}
0 & 3 & 0 & 0 \\
3 & 4 & 0 & 0 \\
0 & 0 & 7 & 0 \\
0 & 0 & 45 & 63
\end{bmatrix},
\]

Also, make the replacement “Note that both diagonal blocks in B are 2 × 2” → “Note that the leading diagonal block in B is 2 × 2”
12. p. 75, line −3. “In the latter case, we compute the appropriate” → “When \(\|p^U\| \leq \Delta\), the appropriate value of \(\tau\) is obtained from

\[
\tau = \frac{\Delta}{\|p^U\|}.
\]

Otherwise, when \(\|p^U\| < \Delta < \|p^B\|\), we compute the appropriate”
13. p. 80, line −3. Delete “for some \(t \in (0, 1)\).”.
14. p. 81, line 1. “to denote the Lipschitz” → “to denote half the Lipschitz”
15. p. 84, line −5. “λ ≠ λ_j” → “λ ≠ −λ_j”
16. p. 85, on the line after (4.40). “whcih” → “which”
17. p. 90, line 9. “global minimum” → “global minimizer”.
19. p. 99, line 1. “the sequence \{||g||\}” → “the sequence \{||g_k||\}”
21. p. 145, lines 14-15. Item 2 should read “If \(y_k = B_k s_k\), then the trivial updating formula \(B_{k+1} = B_k\) satisfies the secant condition.”
22. p. 158, formula (6.57) should be
\[
\bar{M}_k = \frac{||\bar{y}_k||^2}{\bar{y}_k^T \bar{s}_k} \leq \frac{(1 + \bar{c} \epsilon_k)^2}{1 - \bar{c} \epsilon_k},
\]
23. p. 158, formula (6.58) should be
\[
\bar{M}_k \leq 1 + \frac{3\bar{c} + \bar{c}^2 \epsilon_k}{1 - \bar{c} \epsilon_k} \epsilon_k \leq 1 + c \epsilon_k.
\]
24. p. 162, exercise 6.5 should read “Prove that if \(y_k \neq B_k s_k\) and \((y_k - B_k s_k)^T s_k = 0\), then there is no symmetric rank-one updating formula that satisfies the secant condition.
25. p. 167, line 9. The first line of this displayed multiline formula should be
\[
\nabla f_{k+1} = \nabla f_k + \nabla^2 f_k p_k + \int_0^1 [\nabla^2 f(x_k + tp_k) - \nabla^2 f(x_k)] p_k \, dt
\]
(The quantities in the integral should be Hessians, not gradients.)
26. p. 171, line 8 of Algorithm 7.2. Remove “in (4.5)”.
27. p. 176, eq (7.14). “Q_j” → “Q_j^T”.
28. p. 192, line 5. “its area is \(q^2\)” → “its area is \(q^{-2}\)”
29. p. 232, line 6. “\(k = 1, 2, \ldots\)” → “\(k = 0, 1, 2, \ldots\)”
30. p. 238, line 18. “toward this value” → “toward the best vertex”.
31. p. 238, line 19. “after some defining some notation” → “after defining some notation”.
32. p. 238, line −9. Should be
\[
\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.
\]
33. p. 239, line 12. “\( f_{-1/2} = \bar{x}(-1/2) \) → “\( f_{-1/2} = f(\bar{x}(-1/2)) \)”

34. p. 239, line 17. “\( f_{1/2} = \bar{x}(1/2) \) → “\( f_{1/2} = f(\bar{x}(1/2)) \)”

35. p. 239, line 23. “three-dimensional” → “two-dimensional”

36. p. 240, caption of Figure 9.4. “simplex method in \( \mathbb{R}^3 \) → “simplex method in \( \mathbb{R}^2 \)”

37. p. 253, line -11. “less sentitive to” → “less sensitive to”

38. p. 255, line -3. “can applied to study” → “can be applied to study”

39. p. 260, lines 2 and 4. “\( \lambda I \)” → “\( \sqrt{\lambda} I \)”.

40. p. 269, Exercise 10.1. Delete the phrase “, and let \( y \in \mathbb{R}^m \) be a vector”

41. p. 269, Exercise 10.5. “Assume also that the \( r_j \) are bounded on \( D \), that is, there exists \( M > 0 \) such that \( |r_j(x)| \leq M \) for all \( j = 1, 2, \ldots, m \) and all \( x \in D \)” → “Assume also that the \( r_j \) and \( \nabla r_j \) are bounded on \( D \), that is, there exists \( M > 0 \) such that \( |r_j(x)| \leq M \) and \( \|\nabla r_j(x)\| \leq M \) for all \( j = 1, 2, \ldots, m \) and all \( x \in D \)”.

42. p. 276, formula (11.11) should be

\[
w(x_k, x^*) = \int_0^1 \left[ J(x_k + t(x^* - x_k)) - J(x_k) \right] (x_k - x^*) \, dt. \tag{1}
\]

43. p. 279, line 11. “at most 1/2” → “at most 3/4”.

44. p. 294, line –7. “\( \int_0^1 \beta L \|p_k\|^2 \, dt \)” → “\( \int_0^1 t \beta L \|p_k\|^2 \, dt \)”.

45. p. 295, line –2. “not be increased” → “not be decreased”.

46. p. 303, line 1. “decreasing in \( \lambda \)” → “decreasing in \( \lambda > 0 \)”

47. p. 314, line –7. “it s easy to identify vectors \( d \) that satisfies” → “it is easy to identify vectors \( d \) that satisfy”


49. p. 317, line –2. “sequence are \( (d = (0, \alpha)^T) \)” → “sequence are \( (d = (0, \alpha)^T \) with \( \alpha \geq 0 \)”

50. p. 324, line 14. “positive scalars such” → “positive scalars such that”

51. p. 324, line –8. “At \( t = 0, z = x^* \), and the Jacobian of \( R \) at this point is” → “At \( t = 0, we have \( z = x^* \), and the Jacobian of \( R \) with respect to \( z \) at this point is”
In fact, the solution \( z \) of (12.40) is an implicit function of \( t \); we can write it as \( z(t) \), and note that \( z_k = z(t_k) \). The implicit function theorem states that \( z \) is a \textit{continuously differentiable} function of \( t \), with

\[
z'(0) = -\nabla_z R(x^*, 0)^{-1} \nabla_t R(x^*, 0),
\]

and we can use (12.40) and (12.41) to deduce that \( z'(0) = d \). Since \( z(0) = x^* \), we have that

\[
\frac{z_k - x^*}{t_k} = \frac{z(0) + t_k z'(0) + o(t_k) - x^*}{t_k} = d + \frac{o(t_k)}{t_k},
\]

from which it follows that (12.29) is satisfied (for \( x = x^* \)), Hence, \( d \in T_\Omega(x^*) \) for an arbitrary \( d \in F(x^*) \), so the proof of (ii) is complete.

52. p. 325, Replace the paragraph starting on line 1 and ending on line 10 (that is, “It remains to verify....” through “proof of (ii) is complete”) with the following paragraph:

53. p. 325, line −11. “at which all feasible sequences” → “at which all feasible sequences approaching \( x \)”

54. p. 328, line 6. “2t” → “2α” in the second equation of this line.

55. p. 329, formula (12.51). “\( A(x^*)^T \lambda^* \)” → “\( A(x^*)^T \lambda \)”

56. p. 333, formula (12.63). replace the term

\[
\frac{1}{2} t_k^2 w^T \nabla^2 \mathcal{L}(x^*, \lambda^*)
\]

by

\[
\frac{1}{2} t_k^2 w^T \nabla^2 \mathcal{L}(x^*, \lambda^*) w
\]

57. p. 333, line −7. “condition (12.65) by” → “condition (12.65) can be replaced by”

58. p. 336, line −3. The matrix in the formula should be

\[
\begin{bmatrix}
-0.8 & 0 \\
0 & 1.4
\end{bmatrix}
\]

59. p. 337, add after line 12: “where \( |A(x^*)| \) denotes the cardinality of \( A(x^*) \).”

60. p. 341, statement of Lemma 12.9. “Then t the normal cone” → “Then the normal cone”

61. p. 341, lines 16 and 19. In these two displayed formulae, replace \( \Rightarrow \) by \( \Leftrightarrow \).

62. p. 344, line 3. “\( q : \mathbb{R}^n \to \mathbb{R} \)” → “\( q : \mathbb{R}^m \to \mathbb{R} \)”

63. p. 344, formula (12.84) should be

\[
\max_{\lambda \in \mathbb{R}^m} q(\lambda) \quad \text{subject to} \quad \lambda \geq 0.
\]
64. p. 351, in formula (12.96), replace $x^6 \sin(1/x) = 0$ by $x^6 \sin(1/x)$.

65. p. 443, line 15. “from from” → “from”.

66. p. 444, line 14. “if does not” → “if it does not”.

67. p. 455, line 15. “to obtain $\hat{y}$” → “to obtain $\hat{z}$”.

68. p. 461, line 15. “the scaled $n \times n$ projection matrix” → “the $n \times n$ matrix”.

69. p. 468, line –6. “positive definite” → “positive semidefinite”.

70. p. 488, line –13. “else (ii) $\Delta t^*$” → “else (ii) if $\Delta t^*$”.

71. p. 600, line –6. “is a nonnegative multiple” → “is a multiple”

72. p. 602, line 16. “(i) the whole space $\mathbb{R}^n$” → “the whole space $\mathbb{R}^2$”

73. p. 609, line 14. “set $x = P^T z$” → “set $x = Pz$”

74. p. 615. lines –12 to –9. Replace this sentence by the following: “By combining these expressions, we find that the difference between this result and the true value $x - y$ may be as large as a quantity that is bounded by $u(|x| + |y| + |x - y|)$ (ignoring terms of order $u^2$).”

75. p. 616, displayed formula on line –4. “≈” → “≤”

76. p. 617, formula (A.32). “≈” → “≤”

77. p. 618, line 7. This displayed formula should be

$$\|x_k - \hat{x}\| \leq \epsilon, \quad \text{for some } k \geq K.$$

78. p. 620, line 16. “have $(1 + (0.5)^k) - 1 = (0.5)^k$” → “have $|(1 + (0.5)^k) - 1| = (0.5)^k$”

79. p. 629, line –1. “$1/\sqrt{13}$” → “$1/\sqrt{3}$”

Thanks to Carlos Henao, Jorgen Sand, Marc Steinbach.