Errata for Algorithms for Convex Optimization

Nisheeth K. Vishnoi

This errata refers to the printed edition of the book, available here and here. The author would like to thank the contributors to this errata. Further corrections, feedback, and comments are welcome and should be emailed to the author.

List of contributors to errata: Sergio Cabello, Anay Mehrotra

First print (2021)

- Chapter 1 - Bridging continuous and discrete optimization

- Chapter 2 - Preliminaries
  - Definition 2.7 (Page 20)
    Old text: $\sum_{k\geq 3} \nabla^k f(a)[x-a, \ldots, x-a]$
    New text: $\sum_{k\geq 3} \frac{1}{k!} \nabla^k f(a)[x-a, \ldots, x-a]$
  - Page 20:
    Old text: $\left\{\alpha_0 + \sum_{i=1}^{k} \alpha_i v_i : \alpha_0, \alpha_1, \ldots, \alpha_k \in \mathbb{R}\right\}$
    New text: $\left\{v_0 + \sum_{i=1}^{k} \alpha_i v_i : v_0 \in \mathbb{R}^n, \;\; \alpha_1, \ldots, \alpha_k \in \mathbb{R}\right\}$
  - Exercise 2.1 (a) (Page 29)
    Old text: $a_1, \ldots, a_m \in \mathbb{Q}^m$
    New text: $a_1, \ldots, a_m \in \mathbb{Q}^n$

- Exercise 2.1 (c) (Page 29)
  Old text: A is a symmetric $n \times n$ matrix and X runs over symmetric matrices
  New text: A is a real-symmetric $n \times n$ matrix and X runs over real-symmetric matrices

- Chapter 3 - Convexity
  - Theorem 3.6 (Page 39)
    Old text: $f : K \to R$
    New text: $f : K \to \mathbb{R}$

- Chapter 4 - Convex Optimization and Efficiency
  - Section 4.1.1 (Page 50)
    Old text: $b \in \mathbb{R}^n$
    New text: $b \in \mathbb{R}^m$
  - Exercise 4.9 (c) (Page 66)
    Old text: vertex of P
    New text: vertex of K
  - Exercise 4.11 (c) (Page 67)
    Old text: polynomial time
    New text: polynomial time algorithm

- Chapter 5 - Duality and Optimality
  - Theorem 5.5 (Page 72)
    Old text: Suppose that the functions $f, f_1, f_2, ..., f_m$ and ...
    New text: Suppose that the functions $f, f_1, f_2, ..., f_m$ are convex and ...
  - Exercise 5.11 (Page 80)
    Old text: consider the primal problem where
    New text: consider the primal problem from Equation (5.2) where
  - Exercise 5.11 (Page 80)
    Old text: Prove that the dual is

New text: Prove that the objective of the dual is

- Exercise 5.12 (Page 80)

  Old text: primal convex optimization problem over

  New text: primal convex optimization problem from Equation (5.2) over

- Exercise 5.12 (Page 80) (Page 81)

  Old text: $a^\top Xa \leq 0$

  New text: $a^\top Xa \leq 1$

- Exercise 5.14 (c) (Page 81)

  Old text: positive orthant

  New text: non-negative orthant

- Chapter 6 - Gradient Descent

  - Page 91

    The last inequality on this page is an equality

  - Last equation on Page 92

    Old text: $R_t \leq$

    New text: $R_t =$

  - Exercise 6.5 (Page 102)

    Old text: which satisfies for every $x \in \mathbb{R}^n$, one has $mI \preceq \nabla^2 f(x) \preceq MI$

    New text: which satisfies for every $x \in \mathbb{R}^n$, $mI \preceq \nabla^2 f(x) \preceq MI$

- Chapter 7 - Mirror Descent and Multiplicative Weights Update

  - Page 112

    Old text: we can ignore terms that depend only on x

    New text: we can ignore terms that depend only on x^t

  - Page 112

    Old text: $f\left(x^t\right)$ may be more than $f\left(x^{t+1}\right)$

    New text: $f\left(x^{t+1}\right)$ may be more than $f\left(x^t\right)$

  - Lemma 7.4 (Page 115)

    In item 2, drop "for all $i = 1, 2, \ldots, n$" at the end

- Page 130

  Old text: $\alpha_e := \sum_{e \in N(v)} w_v^t$

  New text: $\alpha_e := \sum_{v:\, e \in N(v)} w_v^t$

- Page 131

  In the first two inequalities replace x_e by x_e^t

- Page 131

  Old text: $\frac{1}{T} \cdot \frac{1}{n}\left(\sum_{e:\, v \in e} x_e^t - 1\right) \leq \frac{1}{T} \cdot T \cdot 0 + \delta$

  New text: $\frac{1}{T} \cdot \sum_{t=0}^{T-1} \frac{1}{n}\left(\sum_{e:\, v \in e} x_e^t - 1\right) \leq \frac{1}{T} \cdot T \cdot 0 + \delta$

- Page 134

  Old text: $x := \frac{1}{T} \sum_{t=0}^{T-1} x^i$

  New text: $x := \frac{1}{T} \sum_{t=0}^{T-1} x^t$

- Exercise 7.11 (Page 134)

  Old text: Assume that $||\nabla f(x)||_2 \leq G$

  New text: Assume that $\mathbb{E}\left[||g(x)||_2^2\right] \leq G^2$

- Exercise 7.17 (Page 139)

  Old text: oracle satisfy $\left|\left|g^t\right|\right| \leq 1.$

  New text: oracle satisfy $0 \leq g_i^t \leq 1$ for all $i \in [n]$

- Exercise 7.17 (d) (Page 139)

  Old text: the following always holds:

  $$-\varepsilon \sum_{t=0}^{T-1} \hat{g}_i^t - \ln n \leq -\frac{\varepsilon}{1 - n\varepsilon} \sum_{t=0}^{T-1} \langle p^t, \hat{g}^t \rangle + \frac{2\varepsilon^2}{1 - n\varepsilon} nT$$

  New text: the following holds for any fixed i:

$$-\varepsilon \sum_{t=0}^{T-1} g_i^t - \ln n \leq -\frac{\varepsilon}{1-n\varepsilon} \sum_{t=0}^{T-1} \langle p^t, g^t \rangle + \frac{2\varepsilon^2}{1-n\varepsilon} nT$$

- Chapter 8 - Accelerated Gradient Descent
  - Equation 8.15 (Page 147)

    In this equation, also define $\lambda_t := (1-\gamma_t)\lambda_{t-1}$
  - Exercise 8.3 (Page 157)

    Old text: $f(x) := \min_{x \in \mathbb{R}^n} \frac{1}{2}(x-x^\star)^\top A(x-x^\star)$

    New text: $\min_{x \in \mathbb{R}^n} \frac{1}{2}(x-x^\star)^\top A(x-x^\star)$
  - Exercise 8.3 (Page 157)

    Set $x_0 = x_1$.
  - Exercise 8.3 (b) (Page 158)

    Old text: $\theta := \max\left\{\left|1-\sqrt{\eta\lambda_1}\right|, \left|1-\sqrt{\eta\lambda_n}\right|\right\}$

    New text: $\theta := \max\left\{\left|1-\sqrt{\eta\lambda_1}\right|, \left|1-\sqrt{\eta\lambda_n}\right|\right\}^2$

- Chapter 9 - Newton's Method
  - Figure 9.1 (Page 161)

    Replace f by g and $f(x_0)$ by $g(x_0)$
  - Page 168

    Old text: $D^{(3)}$

    New text: D^3
  - Theorem 9.4 (Page 167)

    Old text: Let x_0 be an arbitrary starting point

    New text: Let x_0 be the starting point specified in the NE condition
  - Proof of 9.4 (Page 167)

    Old text: We can take $M = \frac{L\left|\left|H(x_0)^{-1}\right|\right|}{2} \leq \frac{L}{2h}$

    New text: Since f and x_0 satisfy the NE condition, we can take

    $M = \frac{L\left|\left|H(x_0)^{-1}\right|\right|}{2} \leq \frac{L}{2h}$

- Chapter 10 - An Interior Point Method for Linear Programming
  - Theorem 10.2 (Page 186)

    Old text: or terminates stating that the polyhedron is infeasible. The algorithm runs in poly$\left(L, \log \frac{1}{\varepsilon}\right)$ time.

    New text: The algorithm runs in poly$\left(L, \log \frac{1}{\varepsilon}\right)$ time.
  - Lemma 10.9 (Page 194)

    Let x_T and η_T be the variables defined in Algorithm 8
  - Theorem 10.10 (Page 194)

    Old text: outputs a point $\widehat{x} \in P$ that satisfies

    New text: outputs a point $\widehat{x}$ ∈int(P) that satisfies
  - Proof of Lemma 10.7 (Page 196)

    Old text: $n_{\eta'}(x) = H(x)^{-1} \nabla f_{\eta'}(x)$

    New text: $-n_{\eta'}(x) = H(x)^{-1} \nabla f_{\eta'}(x)$
  - Lemma 10.15 (Page 203)

    Old text: For any $\eta > 0$ denote

    New text: For any $\eta > 0$ and $x \in \mathbb{R}^n$ denote

- Chapter 11 - Variants of the Interior Point Method and Self-Concordance
  - Equation 11.3 (Page 218)

    Old text:

$$\begin{bmatrix} B & 0 \\ I & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b \\ u \end{bmatrix}$$

    New text:

$$\begin{bmatrix} B & 0 \\ I & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} b \\ \rho \end{bmatrix}$$

- Section 11.2.2 (Page 222)

  Old text: (where $E = \ldots$ such that

  New text: (where $E = \ldots$) such that

- Equation 11.8 (Page 225)

  Old text: $\widetilde{n}_\eta(x) := X^2 \left(A^\top \left(AX^{-2} A^\top \right)^{-1} AX^2 - I \right) \left(\eta c + X^{-1} 1 \right)$

  New text: $\widetilde{n}_\eta(x) := X^2 \left(A^\top \left(AX^{2} A^\top \right)^{-1} AX^2 - I \right) \left(\eta c + X^{-1} 1 \right)$

- Lemma 11.10 (Page 228)

  Old text: $s, t \in G$

  New text: $s, t \in V$

- Page 229

  Old text: $\frac{F}{2} \leq x_{\hat{i}} \leq F = u_{\hat{i}} - F$

  New text: $\frac{F}{2} \leq x_{\hat{i}} \leq F = \rho_{\hat{i}} - F$

- Page 229

  Old text: $\min \left\{ \frac{1}{2}, \frac{F}{2} \right\} \leq y_i$

  New text: $\min \left\{ \frac{1}{2}, F \right\} \leq y_i$

- Page 231

  Old text: $\left|\left| H(x)^{-1} g(x) \right|\right|_x = g(x)^\top H(x)^{-1} g(x)$

  New text: $\left|\left| H(x)^{-1} g(x) \right|\right|_x^2 = g(x)^\top H(x)^{-1} g(x)$

- Page 231

  Old text: $\left|\left| H(x)^{-1} g(x) \right|\right|_x = 1^\top \Pi 1 = ||\Pi 1||_2^2 \leq ||1||_2^2 = m$

  New text: $\left|\left| H(x)^{-1} g(x) \right|\right|_x^2 = 1^\top \Pi 1 = ||\Pi 1||_2^2 \leq ||1||_2^2 = m$

- Exercise 11.10 (Page 244)

  Old text: self-concordance implies that for all $x \in K$

  New text: self-concordance implies that for all $x \in$int(K)

- Exercise 11.15 (Page 247)

  Part (d) of this exercise should be its own exercise 11.16

- Chapter 12 - Ellipsoid Method for Linear Programming

- Algorithm 9 (Page 253)

  Old text: A number u such that $u \leq y^\star + \varepsilon$

  New text: A number u such that $y^\star \leq u \leq y^\star + \varepsilon$
- Algorithm 10 (Page 255)

  Before Line 1 add "Initialize $t = 0$" and after Line 9 add "Set $t := t + 1$"
- Proof of Theorem 12.7 (Page 256)

  Replace all occurences of n by m
- Section 12.2.2 (Page 257)

  Old text: the method developed we present in

  New text: the method developed in
- Section 12.4 (Page 262)

  Old text: The proof of Lemma 12.12 is has two parts

  New text: The proof of Lemma 12.12 has two parts
- Proof of Lemma 12.18 (Page 272)

  Old text: $\geq \langle c, x^\star \rangle + (1 - \alpha)$

  New text: $= \langle c, x^\star \rangle + (1 - \alpha)$
- Section 12.5.5 (Page 273)

  Old text: but still manageable The idea

  New text: but still manageable. The idea
- Exercise 12.3 (Page 274)

  Replace all occurences of n by m
- Part (c), Exercise 12.3 (Page 274)

  Old text: a polynomial time linear optimization oracle

  New text: a polynomial time separation oracle
- Exercise 12.6 (Page 276)

  Old text: larger than the lower bound derived above

  New text: larger than the upper bound derived above

- Chapter 13 - Ellipsoid Method for Convex Optimization

  - Section 13.2.2 (Page 287)

    Lovász misspelt as Lovás.

- Theorem 13.12 (Page 287)

  Old text: $l_0 \leq u_0 \in R$

  New text: $l_0 \leq u_0$

- Section 13.3 (below Definition 13.14) (Page 289)

  Old text: Δ_ω is convex

  New text: Δ_Ω is convex

- Definition of $f^-(x)$ in Exercise 13.10 (Page 305)

  Old text: $\sum_{S \subseteq [n]} \alpha_S f(S)$

  New text: $\sum \alpha_S F(S)$