Linear and Convex Optimization
Errata

Updated May 27, 2021

Chapter 15

p. 328 line -4 (from the bottom of the page) "there a" should be "there are"

p. 334 Equation (15.14) The entry $\mathbf{S}$ should be $-\mathbf{S}$.

p. 336 Algorithm 15.2 In Step 4, the stopping condition $\|\mathbf{d}\| < \epsilon$ should be $\lambda\|\mathbf{d^x}\| < \epsilon$.

p. 336 **Example 15.9.** The first line of the solution should have $g'(x) = 1$, not 0. This affects the displayed matrix equation, which then does not have a simple solution, the numerical results, and the discussion of the results. Here is a corrected solution.

The derivatives are $f'(x) = 10 - e^x$, $f''(x) = -e^x$, $g'(x) = 1$, and $g''(x) = 0$. Equation (15.14) for the Newton step is

$$\begin{bmatrix} e^x & 1 \\ 1 & \frac{2-x}{y} \end{bmatrix} \begin{bmatrix} d^x \\ d^y \end{bmatrix} = \begin{bmatrix} 10 - e^x - y \\ 2 - x - \frac{\mu}{y} \end{bmatrix}.$$

We apply the algorithm starting at $x = 1$, $y = 0.5$, and $\mu = 2$. However, instead of backtracking, we can just compute the slack in the constraint and include the ratio $(2 - x)/d^x$ in the step size formula (15.15). We use $\alpha = 0.8$, i.e., the final step size is (15.15) multiplied by 0.8, and $\gamma = 0.1$ to decrease μ. The table shows that the algorithm converges to the optimal solution $x = 2$. To understand the step size, first consider the Newton step for the original objective. The maximum of $f(x)$ is at $x = \ln(10) \approx 2.30$. For any $x < \ln 10$, the Newton step moves beyond this to the maximum of the quadratic fit. The Newton step for the barrier problem is smaller. For example, from the initial $x = 1$, it moves to $x + d^x = 1 + 1.641 = 2.641$; the Newton step for $f(x)$ would have moved to $x = 10/e \approx 3.68$. Because $x = 2.641$ is infeasible, $\lambda_{\max}$ is reduced to a step that would preserve feasibility ($x = 2$), then multiplied by $\alpha = 0.8$ to obtain the λ shown in the table. If backtracking had been used, the initial $\lambda = 1$ would have been multiplied repeatedly by β until $x + \lambda d^x \leq 2$, then multiplied by 0.8. At this iteration, the specific value of μ does not affect the step taken (but a larger μ would). In contrast, for iterations $k \geq 2$ the step $x + d^x$ is feasible, so the maximum λ of 0.8 is used. Even though μ is smaller for these iterations, x is closer to the boundary and it is enough of a penalty to keep the Newton step feasible. Using $\gamma = 0.1$, so that μ decreases rapidly, allows for faster convergence of the algorithm than if γ were larger. Note that the constraint $y \geq 0$ does not limit the step size; in fact, y gets larger.

k	μ	x	y	d^x	λ
0	2	1	0.5	1.641	0.487
1	0.2	1.8	1.631	0.208	0.770
2	0.02	1.96	2.450	0.035	0.8
3	0.002	1.988	2.611	0.011	0.8
4	2×10^{-4}	1.997	2.616	0.003	0.8
5	–	1.999	2.612	–	

p. 337 The sentence before Example 15.10 is inaccurate. It should read "The next example demonstrates the matrix notation on a two-variable problem."

p. 337-8 **Example 15.10.** The numerical results are incorrect. Here is a corrected solution, starting at p. 337 line -5 (from the bottom of the page).

The Newton steps for $k = 0, 3, 4, 5$ are feasible, $\lambda_{\max} = 1$ in (15.15), and no backtracking is needed. The parameter $\alpha = 0.8$ was used, so the step size is $\lambda = \alpha = 0.8$. For $k = 1$ and 2, the Newton step would make $y_1 < 0$ so λ is set to $0.8\lambda_{\max} = 0.8(-y1/dx_1)$. Again no backtracking is needed and β is not used. The algorithm converges to the optimal solution $\mathbf{x} = (0.164, 0.066)$ which occurs on the boundary where $g_2(\mathbf{x}) = 0$. As $\mathbf{x}$ approaches constraint 2, its Lagrange multiplier y_2 becomes large relative to y_1, reflecting the larger penalty being applied to constraint 2 in the barrier problem.

k	μ	$\mathbf{x}$	$\mathbf{y}$	$-g_1(\mathbf{x})$	$-g_2(\mathbf{x})$	$f(\mathbf{x})$	$\Vert\lambda d^x\Vert$
0	1	$(0.5, 0.6)$	$(5, 10)$	0.107	0.246	2.700	0.3557
1	0.1	$(0.295, 0.309)$	$(1.645, 8.243)$	0.234	0.149	4.120	0.1569
2	0.01	$(0.219, 0.172)$	$(0.329, 7.155)$	0.296	0.069	4.788	0.0887
3	0.001	$(0.178, 0.093)$	$(0.066, 6.948)$	0.329	0.018	5.166	0.0244
4	10^{-4}	$(0.167, 0.072)$	$(0.014, 6.909)$	0.336	0.004	5.267	0.0050
5	10^{-5}	$(0.164, 0.067)$	$(0.003, 6.900)$	0.338	0.001	5.288	0.0002

p. 341 Exercise 15.12c "Repeat (b)" should read "Repeat (a) and (b)".

p. 341 Exercise 15.13. Add (c) Change the initial $\mathbf{y}$ to $(0.25, 8)$ and repeat (a) and (b).