Updates and notes for: High-Dimensional Probability, by Roman Vershynin

After this book was published, many colleagues offered further suggestions and noticed many typos, gaps and inaccuracies. I am grateful to Karthik Abinav, Diego Armentano, Aaron Berk, Soham De, Hu Fu, Adam Jozefiak, Nick Harvey, Harrie Hendriks, Chris Liaw, Hengrui Luo, Mark Meckes, Sikander Randhawa, Karthik Sankararaman, Sasha Sodin, and especially to Aryeh Kontorovich, Jake Knigge, and Abbas Mehrabian who offered substantial feedback that lead to significant improvements.

Section 1.4: Add the following paragraph in the end of this section: "Both Proposition 1.2.4 and Corollary 1.2.5 are due to Chebyshev. However, following established tradition, we call Proposition 1.2.4 Markov's inequality."

Section 2.1, discussion after Theorem 2.1.3: Add the following footnote after "using Stirling's approximation":

Our somewhat informal notation $f \asymp g$ stands for the equivalence of functions (functions of N in this particular example) up to constant factors. Precisely, $f \asymp g$ means that there exist positive constants c, C such that the inequality $cf(x) \le g(x) \le Cf(x)$ holds for all x, or sometimes for all sufficiently large x. For similar one-sided inequalities that hold up to constant factors, we use notation $f \lesssim g$ and $f \gtrsim g$.

Theorem 3.4.6: At the end of the proof, replace "This completes the proof" by "This completes the proof by the characterization of subgaussian distributions (recall Proposition 2.5.2 and Remark 2.5.3)".

In Section 3.8, first paragraph, it is stated that "It is unknown if the quadratic dependence on K in Theorem 3.1.1 is optimal." This problem has been solved in the paper https://arxiv.org/abs/2001.10631, where the quadratic dependence on K in Theorem 3.1.1 has been improved to $O(K \log K)$.

Section 4.8, above the second to last paragraph: Add the following paragraph: "Davis-Kahan's Theorem 4.5.5, originally proved in Davis-Kahan (1970), has become an invaluable tool in numerical analysis and statistics. There are numerous extensions, variants, and alternative proofs of this theorem; see in particular Wedin (1972), Yu-Wang-Samworth (2015), Vu (2011), Bhatia (1997, Section VII.3), Stewart-Sun (1990, Chapter V)."

The references are:

R. Bhatia, *Matrix analysis.* Graduate Texts in Mathematics, 169. Springer-Verlag, New York, 1997.

C. Davis, W. M. Kahan, *The rotation of eigenvectors by a pertubation. III.* SIAM J. Numer. Anal. 7 (1970), 1–46.

G. W. Stewart, Ji G. Sun, *Matrix perturbation theory.* Computer Science and Scientific Computing. Academic Press, Inc., Boston, MA, 1990.

P.-A. Wedin, *Perturbation bounds in connection with singular value decomposition,* BIT Numerical Mathematics 12 (1972), 99–111.

V. Vu, *Singular vectors under random perturbation,* Random Structures & Algorithms 39 (2011), 526–538.

Y. Yu, T. Wang, R. J. Samworth, *A useful variant of the Davis-Kahan theorem for statisticians,* Biometrika 102 (2015), 315–323.

Section 5.4.2: Replace the first paragraph by the following: “So far, our attempts to extend scalar concepts for matrices have not met a considerable resistance. But this does not always go so smoothly. We already saw in Exercise 5.4.5 how the noncommutativity of the matrix product ($AB \neq BA$) may cause scalar properties to fail for matrices. Here is one more such situation: the identity is $e^{x+y} = e^x e^y$ holds for scalars but fails for matrices.”

Section 5.4.2, Right below Theorem 5.4.7, add the following sentence: “Unfortunately, the Golden-Thompson inequality does not hold for three or more matrices: in general, the inequality $\mathrm{tr}(e^{A+B+C}) \leq \mathrm{tr}(e^A e^B e^C)$ may fail.”

Section 7.5.4:
(a) second paragraph: after “the cube”, add B_∞^n
(b) last paragraph, “hyperbolic” picture of the B_1^n” – delete “the”.

Exercise 8.2.8 should have 3 coffee cups.

Proof of Lemma 9.1.6, second sentence: Change the beginning to “By definition of the sub-gaussian norm,[1] the conclusion...”

Beginning of Section 10.1:
(a) Remove the sentence “It is natural...” from the second paragraph.
(b) Instead, create the following paragraph: ”To make the signal recovery problem amenable to the methods of high-dimensional probability, we assume the following probabilistic model. We suppose that the measurement matrix A in (10.1) is a realization of a random matrix drawn from some distribution. More specifically, we assume that the rows A_i are independent random vectors, which makes the observations y_i independent, too.”

[1] Recall (2.14) and Remark 2.5.3.