Optimization and Computational Linear Algebra for Data Science
Hints for the review exercises

1 Last year's review exercises

1. Show that for all $x \in \mathbb{R}^n$, $ABx = BAx$. (You can decompose such x in the given basis)
2. (a) See homework 10. (b) Use (a).
3. Use the definition of eigenvectors/eigenvalues.
4. Using the spectral Theorem there exists an orthonormal basis $(v_1, \dots, v_n)$ of $\mathbb{R}^n$ consisting of eigenvectors of A. Decompose x in such a basis and compute Ax.
5. If $f : \mathbb{R}^2 \to \mathbb{R}$ is convex, and if (α^*, β^*) is a minimizer of f then $\nabla f(\alpha^*, \beta^*) = 0$.
6. Use the definition of $\|x\|_\infty$ and $\|x\|$.
7. By the spectral theorem, you can decompose x in an orthonormal basis of $\mathbb{R}^n$ made of eigenvectors of A.
8. Many possible ways to do this. (a) Show that $\mathrm{Ker}(A^\mathsf{T}) = \mathrm{Ker}(AA^\mathsf{T})$, and then use the rank-nullity theorem and the fact that $\mathrm{rank}(A) = \mathrm{rank}(A^\mathsf{T})$. (b) Compute AA^T using the SVD of A: $A = U\Sigma V^\mathsf{T}$.
9. (a) Use Lagrange multipliers. (b) The set of solution of $Ax = b$ is $A^+b + \mathrm{Ker}(A)$. The result follow from the same arguments than problem 1 of homework 10.
10. False.
11. Show that $\|x + y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2$.
12. Show that $\sum_{i=1}^n \langle x, u_i\rangle^2 = \|x\|^2$.
13. Compute AA^T.
14. Show that if λ is an eigenvalue of A associated with the eigenvector u if and only if Qu is an eigenvector of B with eigenvalue λ.
15. Justify that $x = \sum_{i=1}^m \langle x, v_i\rangle v_i$. Then expand $\| \sum \langle x, v_i\rangle v_i\|^2$ and make simplifications.
16. Use the SVD of A.
17. (a) See Homework 3. (b) Let $V \in \mathbb{R}^{n\times n}$ be the matrix whose columns are $v_1, \dots, v_n$. Show that $\mathrm{Tr}(V^T Av) = \sum_{i=1}^n v_i^\mathsf{T} Av_i$. Then use (a). (c) Use the spectral theorem and (a).
18. Use Problem 1.b from homework 7.
19. Expand the right-hand side.
20. (a). $A^2 = 0$. (b) Take for instance

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

21. (a) convex, (b) not convex (c) not convex (d) convex. One can verify these points by computing the Hessian.

22. (a) ... (b) There is a unique global minima.

23. V is of dimension 2, hence $\dim(V^{\perp}) = 4 - 2 = 2$. $v_1 = (1, -1, 0, 0)$ and $v_2 = (0, 0, 1, -1)$ work.

24. (a) Yes. (b) The derivative of a sum is equal to the sum of the derivaties, and the derivatives of λp (for some $\in \mathbb{R}$) is equal to $\lambda p'$. (c) $\mathrm{Ker}(\mathcal{D})$ is the set of polynomials p that are constant (i.e. there exists $a \in \mathbb{R}$ such that $p(x) = a$ for all $x \in \mathbb{R}$). (d) $\mathrm{Im}(\mathcal{D}) = \mathcal{P}_{d-1}$. (e) (i) check the usual conditions (ii) For polynomial of degree $\leq d$, Taylor formula of order d is exact:
$$T_s(p)(x) = p(x+s) = \sum_{k=0}^{d} \frac{p^{(k)}(x)}{k!} s^k = \sum_{k=0}^{d} \frac{\mathcal{D}^k(p)(x)}{k!} s^k.$$
(iii) The matrix has 0 below the diagonal and for $j \geq i$, $M_{i,j} = \binom{j-1}{i-1}$.

25. (a) Let B be a rank 1 matrix. One can therefore write $B = uv^T$ for some $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$.
$$\|A - B\|_F^2 = \|A\|_F^2 - 2u^T A v + \|u\|^2 \|v\|^2.$$
Now, $u^T A v \leq \|u\| \|v\| \sigma_1$. Hence, writing $r = \|u\| \|v\|$
$$\|A - B\|_F^2 \geq \sum_{i=1}^{\min(n,m)} \sigma_i^2 - 2\sigma_1 r + r^2 = \sum_{i=2}^{\min(n,m)} \sigma_i^2 + (\sigma_1 - r)^2 = \|A - A'\|_F^2 + (\sigma_1 - r)^2$$
(b) Let $B = uv^T$ be a rank 1 matrix. Let v_1, v_2 be the first two right-singular vectors of A. $Span(v)^{\perp}$ has dimension $n - 1$, hence one can find a vector of unit norm z in $Span(v)^{\perp} \cap \mathrm{Span}(v_1, v_2)$. We write $z = \alpha_1 v_1 + \alpha_2 v_2$. Since $\|z\| = 1$ and v_1, v_2 orthogonal, we have $\alpha_1^2 + \alpha_2^2 = 1$. By definition of the spectral norm
$$\|A - B\|_{\mathrm{Sp}} \geq \|(A - B)z\| = \|Az\| = \sqrt{\alpha_1^2 \sigma_1^2 + \alpha_2^2 \sigma_2^2} \geq \sigma_2 = \|A - A'\|_{\mathrm{Sp}}.$$

26. xx^T is rank 1 and has two distinct eigenvalues 0 and 1. Hence H has two distinct eigenvalues -1 and 1.

27. The vector $(1, 1, \dots, 1)$ is an eigenvector associated with the eigenvalue d. By contradiction let x be an eigenvector associated with the eigenvalue $\lambda > d$. Let i such that $|x_i| = \|x\|_\infty > 0$. Then
$$|x_i|\lambda = |\sum_{j=1}^{n} G_{i,j} x_j| \leq \sum_{j=1}^{n} G_{i,j} |x_j| \leq |x_i| \sum_{j=1}^{n} G_{i,j} = d|x_i|.$$
We get a contradiction.

28. Use the matrix product formula.

29. (a) convex but not subspace (b) not convex (c) subspace (hence convex)

30. (a) convex but not strictly convex (b) convex but not strictly convex (c) convex but not strictly convex (d) not convex (e) not convex

31. Cauchy-Schwarz.

32. Apply the spectral theorem to A.