This set covers chapters 16–19 of the book *Numerical Optimization* by Nocedal and Wright, 2nd ed.

From the book, the following exercises: 16.1–3, 16.6, 16.10, 16.15–16, 16.20–21, 17.4, 17.11, 18.5, 19.12–13. In addition, the exercises below. There are no Matlab programming exercises. However, you may find it useful to plot the objective function, constraints and auxiliary functions (e.g. the quadratic-penalty function) with `fcontour`.

IV.1. Quadratic programming. Consider the Markowitz model of portfolio optimization (example 16.1 in the book):

$$\max_{\mathbf{x}\in\mathbb{R}^n} q(\mathbf{x}) = \mathbf{x}^T\boldsymbol{\mu} - \kappa\mathbf{x}^T\mathbf{G}\mathbf{x} \qquad \text{subject to} \qquad \mathbf{e}^T\mathbf{x} = 1,\ \mathbf{x} \geq \mathbf{0}$$

where $\kappa \geq 0$, $\boldsymbol{\mu} > \mathbf{0}$, $\mathbf{G}$ is symmetric positive definite and $\mathbf{e} = (1,\dots,1)^T$. W.l.o.g. assume that the largest component of $\boldsymbol{\mu}$ is the first ($\mu_1 \geq \mu_i$, $i = 1,\dots,n$).

(i) Suppose that $g_{ii} > g_{ij}$, $i = 1,\dots,n$. Show that, for a solution to be at one corner of the feasible polytope, namely $\mathbf{x}^* = (1,0,\dots,0)^T$, the following condition must hold:

$$\kappa \leq \kappa_a \text{ with } \kappa_a = \min_{i=2,\dots,n} \frac{\mu_1 - \mu_i}{2(g_{11} - g_{1i})}.$$

Interpret this solution. (Hint: apply the KKT and 2nd-order conditions to $\mathbf{x}^*$.)

(ii) Suppose that the sum of each column of $\mathbf{G}^{-1}$ is positive. Show that, for a solution to have only positive components ($\mathbf{x} > \mathbf{0}$), the following condition must hold:

$$\kappa > \kappa_c \text{ with } \kappa_c = \frac{1}{2}\mathbf{e}^T\mathbf{G}^{-1}\boldsymbol{\mu} - \frac{1}{2}\mathbf{e}^T\mathbf{G}^{-1}\mathbf{e}\left(\min_{i=1,\dots,n} \frac{(\mathbf{G}^{-1}\boldsymbol{\mu})_i}{(\mathbf{G}^{-1}\mathbf{e})_i}\right).$$

What is the solution $\mathbf{x}^*$? What is the solution for $\kappa \to \infty$?

(iii) Consider a different portfolio optimization problem:

$$\min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{x}^T\mathbf{G}\mathbf{x} \qquad \text{subject to} \qquad \boldsymbol{\mu}^T\mathbf{x} \geq \kappa,\ \mathbf{e}^T\mathbf{x} = 1,\ \mathbf{x} \geq \mathbf{0}$$

where $\kappa \geq 0$ as before. What is the largest κ for which this problem is feasible? (Hint: formulate as an LP $\kappa = \max \boldsymbol{\mu}^T\mathbf{x}$ s.t. $\mathbf{e}^T\mathbf{x} = 1,\ \mathbf{x} \geq \mathbf{0}$; guess its solution and prove it is a solution using the KKT conditions.)

IV.2. Quadratic-penalty, augmented-Lagrangian, log-barrier and interior-point methods.

(i) Consider the constrained optimization problem $\min_{\mathbf{x}} x_1^2 + x_2^2$ s.t. $x_1 + x_2 = 1$. Find the solution $(\mathbf{x}^*, \lambda^*)$ of this problem using the KKT and second-order conditions. Write the quadratic-penalty function $Q(\mathbf{x};\mu)$ and its gradient $\nabla_{\mathbf{x}} Q(\mathbf{x};\mu)$ and Hessian $\nabla^2_{\mathbf{xx}} Q(\mathbf{x};\mu)$. Show that: $Q(\mathbf{x};\mu)$ has a single minimiser $\mathbf{x}_k$ for each $\mu_k > 0$; this minimiser tends to the solution $\mathbf{x}^*$ of the problem as $\mu_k \to \infty$; the Lagrange multiplier estimate $\lambda_k \approx -\mu_k c(\mathbf{x}_k)$ (eq. (17.10) in the book) tends to the Lagrange multiplier λ^* at the solution as $\mu_k \to \infty$; the Hessian of the penalty function at the minimiser becomes progressively more ill-conditioned as $\mu_k \to \infty$, i.e., $\operatorname{cond}\left(\nabla^2_{\mathbf{xx}} Q(\mathbf{x}_k;\mu_k)\right) \to \infty$.

(ii) As in (i) but using the augmented-Lagrangian function $\mathcal{L}_A(\mathbf{x},\lambda;\mu)$ where the variable λ_k is updated as $\lambda_{k+1} \leftarrow \lambda_k - \mu_k c(\mathbf{x}_k)$ (eq. (17.39) in the book).
Also, prove that the sequence of iterates $(\mathbf{x}_k, \lambda_k)$ converges to the solution even if μ_k is kept at a constant value $\mu > 0$ (i.e., we do not drive $\mu_k \to \infty$), and give the convergence order. How does it depend on μ?

(iii) As in (i) but for the problem $\min_{\mathbf{x}} x_1^2 + x_2^2$ s.t. $x_1 \geq 1$ using the log-barrier function $P(\mathbf{x};\mu)$ and where the Lagrange multiplier estimate is $\lambda_k \approx \frac{\mu_k}{c(\mathbf{x}_k)}$ (eq. (19.47) in the book).

(iv) For the problem in (iii), write the system of perturbed KKT equations $\mathbf{F}(\dots) = \mathbf{0}$ for an interior-point method; solve it and find the primal-dual central path; compute the Jacobian $\mathbf{J}$ of $\mathbf{F}$ and indicate how to obtain the Newton step; show that $\mathbf{J}$ does not become progressively more ill-conditioned as we approach the solution.

IV.3. Convex equality-constrained quadratic programming using quadratic-penalty and augmented-Lagrangian methods. Consider the strictly convex equality-constrained QP:

$$\min_{\mathbf{x}\in\mathbb{R}^n} \frac{1}{2}\mathbf{x}^T\mathbf{G}\mathbf{x} + \mathbf{c}^T\mathbf{x} \qquad \text{subject to} \qquad \mathbf{A}\mathbf{x} = \mathbf{b}$$

where $\mathbf{G}$ is symmetric pd and $\mathbf{A}_{m\times n}$ has full rank with $m < n$.

(i) Solve the QP, giving an explicit expression for the solution, i.e., the KKT point $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$.

(ii) Write the quadratic-penalty function $Q(\mathbf{x};\mu)$, prove it is strictly convex and find its minimizer $\mathbf{x}(\mu)$ and its Lagrange multiplier estimates $\boldsymbol{\lambda}(\mu) = -\mu\,(\mathbf{A}\mathbf{x}(\mu) - \mathbf{b})$ explicitly. Prove $\mathbf{x}(\mu) \xrightarrow[\mu\to\infty]{} \mathbf{x}^*$ and $\boldsymbol{\lambda}(\mu) \xrightarrow[\mu\to\infty]{} \boldsymbol{\lambda}^*$. Hint: use the Sherman-Morrison-Woodbury formula.

(iii) Write the augmented-Lagrangian function $\mathcal{L}_A(\mathbf{x},\boldsymbol{\lambda};\mu)$, prove it is strictly convex and find its minimizer $\mathbf{x}(\boldsymbol{\lambda};\mu)$ and the update for $\boldsymbol{\lambda}$ explicitly. Prove $\mathbf{x}(\boldsymbol{\lambda};\mu) \xrightarrow[\mu\to\infty]{} \mathbf{x}^*$ and $\boldsymbol{\lambda}(\mu) \xrightarrow[\mu\to\infty]{} \boldsymbol{\lambda}^*$.

Now we keep the penalty parameter fixed to a value $\mu > 0$. Prove that iterating the update for $\boldsymbol{\lambda}$ also converges to the solution, and determine the convergence order. Hint: $\|\mathbf{M}\mathbf{u}\| \le \|\mathbf{M}\|\,\|\mathbf{u}\|$, using the L_2 norm for vectors and matrices, i.e., $\|\mathbf{u}\| = \sqrt{u_1^2 + \cdots + u_n^2}$ and $\|\mathbf{M}\|$ = largest singular value of $\mathbf{M}$.

IV.4. Generalized spectral problem. Prove the following theorem. Consider the optimisation problem

$$\max_{\mathbf{X}} \operatorname{tr}\left(\mathbf{X}\mathbf{A}\mathbf{X}^T\right) \quad \text{s.t.} \quad \mathbf{X}\mathbf{B}\mathbf{X}^T = \mathbf{I} \tag{1}$$

where $\mathbf{X} \in \mathbb{R}^{L\times N}$, $L < N$, $\mathbf{A},\mathbf{B} \in \mathbb{R}^{N\times N}$, $\mathbf{A}$ is symmetric and $\mathbf{B}$ is symmetric positive definite. Let $\mathbf{C} = \mathbf{B}^{-\frac{1}{2}}\mathbf{A}\mathbf{B}^{-\frac{1}{2}}$ have spectral representation $\mathbf{C} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$ and assume its eigenvalues are distinct though not necessarily positive. Then, the solution of (1) is unique and given by $\mathbf{X} = \mathbf{U}_L^T\mathbf{B}^{-\frac{1}{2}}$, where $\mathbf{U}_L = (\mathbf{u}_1,\dots,\mathbf{u}_L)$ are the leading L eigenvectors of $\mathbf{C}$.