How to solve simple QPs

The problem we discussed in class had the following general structure:

Minimize $\lambda \left(\sum_i \sigma_i^2 x_i^2 \; + \; 2 \sum_{i<j} \sigma_{ij} x_i x_j \right) \; - \; \sum_i \mu_i x_i$, (*)

Subject to the constraints:

$\sum_i x_i \; = \; 1$, and

$l_i \; \leq \; x_i \; \leq \; u_i$, for all i.

The method we discussed in class consisted of two steps. For simplicity of notation, we will assume that the variables are numbered $1, 2, \ldots, n$ (as opposed to 0 through $n-1$, which is the way we would probably end up doing it in a program).

Step 1. Achieve feasibility

This works as follows: we *start* with $\bar{x}_i = l_i$ for all i. If it is already the case that $\sum_i \bar{x}_i > 1$ then we can quit because the problem is infeasible (no solutions exist). If $\sum_i \bar{x}_i = 1$ (or very close to it) then our $\bar{x}$ is feasible. In the remaining case, $\sum_i \bar{x}_i < 1$, we simply increase each $\bar{x}_i$, *one at a time*, starting with $\bar{x}_1$, then $\bar{x}_2$, and so on, until we achieve $\sum_i \bar{x}_i = 1$. We stop the increase of any $\bar{x}_i$ when it reaches its upper bound, u_i. If it is the case that each $\bar{x}_i = u_i$ and we *still* have $\sum_i \bar{x}_i < 1$ then again we quit since the problem is infeasible.
Assume we have now a feasible $\bar{x}$, we go on to the next step.

Step 2. Improvement phase

In the improvement phase, we repeatedly try to improve the solution by shifting weight from one variable to another. We will, in this fashion, look at many pairs of variables, one pair at a time.
When considering a pair h, k of different variables, we try to shift some weight ϵ from $\bar{x}_h$ to $\bar{x}_k$ (if $\epsilon < 0$ are actually shifting weight in the other direction). We will pick the value ϵ that gives us the *largest* improvement in (*).
When we reset $\bar{x}_h \leftarrow \bar{x}_h - \epsilon$, and $\bar{x}_k \leftarrow \bar{x}_k + \epsilon$, the objective value (*) becomes:

$$\lambda \left(\sigma_h^2 (\bar{x}_h - \epsilon)^2 + \sigma_k^2 (\bar{x}_k + \epsilon)^2 + \sum_{i \neq h,k} \sigma_i^2 \bar{x}_i^2 \right)$$
$$+ 2\lambda \left(\sigma_{hk} (\bar{x}_h - \epsilon)(\bar{x}_k + \epsilon) + \sum_{i \neq h,k} \sigma_{ih} \bar{x}_i (\bar{x}_h - \epsilon) + \sum_{i \neq h,k} \sigma_{ik} \bar{x}_i (\bar{x}_k + \epsilon) + \sum_{i<j, i,j \neq h,k} \sigma_{ij} x_i x_j \right)$$
$$- \mu_h (\bar{x}_h - \epsilon) - \mu_k (\bar{x}_k + \epsilon) - \sum_{i \neq h,k} \mu_i x_i$$

If we collect in this expression the terms that depend on ϵ we get:

$$\lambda (\sigma_h^2 + \sigma_k^2 - 2\sigma_{hk}) \epsilon^2 \; + \; \left\{ 2\lambda \left[-\sigma_h^2 \bar{x}_h + \sigma_k^2 \bar{x}_k + \sigma_{hk} (\bar{x}_h - \bar{x}_k) + \sum_{i \neq h,k} (\sigma_{ik} - \sigma_{ih}) \bar{x}_i \right] \; + \; (\mu_h - \mu_k) \right\} \epsilon$$

which we can summarize as

$$a\epsilon^2 \; + \; b\epsilon, \qquad (**)$$

for appropriate a, b. We want to choose ϵ so as to minimize the expression (**). We can assume $a \geq 0$ (because $\lambda > 0$ and the covariance matrix is positive-semidefinite).
Suppose that we actually have $a > 0$. Then how do we choose ϵ? What we can do is compute the first derivative of (**) and set it to zero. This happens at

$$\epsilon^* = -\tfrac{b}{2a}.$$

However: this choice of ϵ might result in an infeasible vector: x_k might exceed u_k, for example. To avoid this situation, write

$$\epsilon^+ \;=\; \min\{x_h - l_h, u_k - x_k\} \geq 0, \quad \epsilon^- \;=\; \min\{u_h - x_h, x_k - l_k\} \geq 0,$$

Then, in order to guarantee feasibility, we need to insure that

$$-\epsilon^- \;\leq\; \epsilon \;\leq\; \epsilon^+.$$

So we apply the rule:

(1) If $-\epsilon^- \leq \epsilon^* \leq \epsilon^+$, we set $\epsilon = \epsilon^*$.

(2) If $\epsilon^* < -\epsilon^-$, we set $\epsilon = -\epsilon^-$, and if $\epsilon^* > \epsilon^+$, we set $\epsilon = \epsilon^+$.

Question: how do we minimize (**) when $a = 0$?

What we just described is how to analyze a given pair h, k of variables. We need to build an outer loop around this analysis. This outer loop chooses which pairs to look at, and decides when to stop looking at pairs of variables.
A simple procedure for this is to look at all pairs in round-robin sequence: first the pair $1, 2$, then $1, 3$, and so on, until we look at all pairs, and then we restart from $1, 2$.
In order to decide when to stop, we could either put a limit on the number of pairs that we look at, or on the amount of time we spend doing this, or we could simply stop when the decrease in (*) starts to slow down.