Homework 12

Self grades are due at 11 PM on December 7, 2023.

1. **Robust Linear Programming**

In this problem we will consider a version of linear programming under uncertainty.

Consider vector $\vec{x} \in \mathbb{R}^n$. Note that $\vec{x}^\top \vec{y} \leq \|\vec{x}\|_1$ for all $\vec{y}$ such that $\|\vec{y}\|_\infty \leq 1$. Further this inequality is tight, since it holds with equality for $\vec{y} = \operatorname{sgn}(\vec{x})$. This observation will be useful in solving the problem.

Let us focus now on a LP in inequality form:

$$\min_{\vec{x}} \quad \vec{c}^\top \vec{x} \tag{1}$$

$$\text{s.t.} \quad \vec{a}_i^\top \vec{x} \leq b_i, \text{ for each } i = 1, ..., m. \tag{2}$$

Consider the set of linear inequalities in (2). Suppose you don't know the vectors $\vec{a}_i$ exactly. Instead you are given nominal values $\vec{\hat{a}}_i$, and you know that the actual vectors satisfy $\left\|\vec{a}_i - \vec{\hat{a}}_i\right\|_\infty \leq \rho$ for a given $\rho > 0$. In other words, the actual components a_{ij} can be anywhere in the intervals $[\hat{a}_{ij} - \rho, \hat{a}_{ij} + \rho]$. Or equivalently, each vector $\vec{a}_i$ can lie anywhere in a hypercube with corners $\vec{\hat{a}}_i + \vec{v}$ where $\vec{v} \in \{-\rho, \rho\}^n$. We desire that the set of inequalities that constrain problem (2) be satisfied for all possible values of $\vec{a}_i$; i.e., we replace these with the constraints

$$\vec{a}_i^\top \vec{x} \leq b_i \text{ for each } \vec{a}_i \in \{\vec{\hat{a}}_i + \vec{v} \mid \|\vec{v}\|_\infty \leq \rho\}, i = 1, ..., m. \tag{3}$$

(a) Argue why for our LP we can replace the infinite set of constraints in (3) by a finite set of $2^n m$ constraints of the form

$$\vec{\hat{a}}_i^\top \vec{x} + \vec{v}^\top \vec{x} \leq b_i \text{ for each } \vec{v} \in \{-\rho, \rho\}^n, i = 1, ..., m. \tag{4}$$

Solution: For each $i \in \{1, \cdots, n\}$, the 2^n vectors $\vec{\hat{a}}_i^\top \vec{x} + \vec{v}^\top \vec{x}$ we get on the LHS of (4) are the vertices of the polyhedron given by all the vectors of the form $\vec{\hat{a}}_i^\top \vec{x} + \vec{v}^\top \vec{x}$ for $\|\vec{v}\|_\infty \leq p$. Hence if all of the inequalities in (4) hold, then all the inequalities in (3) hold. Since the inequalities in (4) are a subset of the inequalities in (3), we have shown what is required.

(b) Show that the constraint set in Equation (3) is in fact equivalent to the much more compact set of m nonlinear inequalities

$$\vec{\hat{a}}_i^\top \vec{x} + \rho \|\vec{x}\|_1 \leq b_i, \quad i = 1, ..., m. \tag{5}$$

Hint: The observation made at the beginning of the problem statement may be useful here.

Solution: The inequalities in Equation (3) can be written as:

$$\vec{\hat{a}}_i^\top \vec{x} + \vec{v}^\top \vec{x} \leq b_i \text{ for each } \vec{v} \in \{v \mid \|\vec{v}\|_\infty \leq \rho\}, i = 1, ..., m. \tag{6}$$

or equivalently as,

$$\vec{\hat{a}}_i^\top \vec{x} + \rho\vec{v}^\top \vec{x} \leq b_i \text{ for each } \vec{v} \in \{\vec{v} \mid \|\vec{v}\|_\infty \leq 1\}, i = 1, ..., m. \tag{7}$$

This is equivalent to

$$\vec{\hat{a}}_i^\top \vec{x} + \rho \max_{\|\vec{v}\|_\infty \leq 1} \vec{v}^\top \vec{x} \leq b_i \text{ for each } i = 1, ..., m. \tag{8}$$

We have $\max_{\|\vec{y}\|_\infty \leq 1} \vec{x}^\top \vec{y} = \|\vec{x}\|_1$, which gives us the equivalent constraint set:

$$\vec{\hat{a}}_i^\top \vec{x} + \rho \|\vec{x}\|_1 \leq b_i, \text{ for each } i = 1, ..., m. \tag{9}$$

We are interested in situations where the vectors $\vec{a}_i$ are uncertain, but satisfy bounds $\left\|\vec{a}_i - \vec{\hat{a}}_i\right\|_\infty \leq \rho$ for given $\vec{\hat{a}}_i$ and ρ. We want to minimize $\vec{c}^\top \vec{x}$ subject to the constraint that the inequalities $\vec{a}_i^\top \vec{x} \leq b_i$ are satisfied for *all* possible values of $\vec{a}_i$.
We call this a *robust LP* :

$$\begin{aligned} \min_{\vec{x}} \quad & \vec{c}^\top \vec{x} && (10) \\ \text{s.t.} \quad & \vec{a}_i^\top \vec{x} \leq b_i, \text{ for each } \vec{a}_i \in \{\vec{\hat{a}}_i + \vec{v} \mid \|\vec{v}\|_\infty \leq \rho\}, \text{ for each } i = 1, ..., m. && (11) \end{aligned}$$

(c) Using the result from part (b), express the above optimization problem as an LP.

Solution: From part (b), we can rewrite the problem as

$$\begin{aligned} \min_{\vec{x}} \quad & \vec{c}^\top \vec{x} && (12) \\ \text{s.t.} \quad & \vec{\hat{a}}_i^\top \vec{x} + \rho \|\vec{x}\|_1 \leq b_i, \text{ for each } i = 1, ..., m. && (13) \end{aligned}$$

We can express this optimization problem as an LP by introducing slack variables t_i, for each $i \in \{1, \cdots, m\}$:

$$\begin{aligned} \min_{\vec{x}, \vec{t}} \quad & \vec{c}^\top \vec{x} && (14) \\ \text{s.t.} \quad & \vec{\hat{a}}_i^\top \vec{x} + \rho \sum_i t_i \leq b_i, \text{ for each } i = 1, ..., m. && (15) \\ & x_i \leq t_i, \text{ for each } i = 1, ..., m. && (16) \\ & -x_i \leq t_i, \text{ for each } i = 1, ..., m. && (17) \end{aligned}$$

2. Formulating Problems as LPs or QPs

This problem explores what kinds of problems can be formulated as LPs or QPs. For each $j \in \{1, 2, 3, 4\}$, either formulate the problem

$$p_j^* := \min_{\vec{x} \in \mathbb{R}^n} f_j(\vec{x}),$$

for the function $f_j : \mathbb{R}^n \to \mathbb{R}$ below as a QP or LP, or explain why it cannot be done. In our formulations, we always use $\vec{x} \in \mathbb{R}^n$ as the variable, and assume that $A \in \mathbb{R}^{m \times n}$, $\vec{y} \in \mathbb{R}^m$.

(a) $f_1(\vec{x}) = \|A\vec{x} - \vec{y}\|_\infty + \|\vec{x}\|_1$.

Solution: We replace the ℓ_∞ norm with constraints on the maximum absolute value of each element of $A\vec{x} - \vec{y}$, and rewrite absolute values as linear constraints. This gives us the LP formulation:

$$\begin{aligned} p_1^* = \quad \min_{\vec{x}, \vec{z} \in \mathbb{R}^n, t \in \mathbb{R}} \quad & t + \vec{\mathbf{1}}^\top \vec{z} \\ \text{s.t.} \quad & -z_i \leq x_i \leq z_i, \text{ for each } i = 1, \cdots, n \\ & -t \leq (A\vec{x} - \vec{y})_i \leq t, \text{ for each } i = 1, \cdots, m. \end{aligned}$$

(b) $f_2(\vec{x}) = \|A\vec{x} - \vec{y}\|_2^2 + \|\vec{x}\|_1$.

Solution: This gives a QP, which we can write as

$$\begin{aligned} p_2^* = \quad \min_{\vec{x}, \vec{z} \in \mathbb{R}^n} \quad & \vec{x}^\top (A^\top A)\vec{x} - 2\vec{y}^\top A\vec{x} + \vec{y}^\top \vec{y} + \vec{\mathbf{1}}^\top \vec{z}, \\ \text{s.t.} \quad & -z_i \leq x_i \leq z_i, \quad \text{for each } i = 1, \cdots, n. \end{aligned}$$

(c) $f_3(\vec{x}) = \|A\vec{x} - \vec{y}\|_2^2 - \|\vec{x}\|_1$.

Solution: The problem is not convex. Consider for instance the special case with $n = 1$, $A = 1$, $y = 0$: plot $f_3(x) = x^2 - |x|$ to verify it is not convex.

(d) $f_4(\vec{x}) = \|A\vec{x} - \vec{y}\|_2^2 + \|\vec{x}\|_1^2$.

Solution: We get the QP

$$\begin{aligned} p_4^* = \quad \min_{\vec{x}, \vec{z} \in \mathbb{R}^n} \quad & \vec{x}^\top (A^\top A)\vec{x} - 2\vec{y}^\top A\vec{x} + \vec{y}^\top \vec{y} + \left(\textstyle\sum_{i=1}^n z_i\right)^2 \\ \text{s.t.} \quad & -z_i \leq x_i \leq z_i, \text{ for each } i = 1, \cdots, n. \end{aligned}$$

Notice that $\left(\sum_{i=1}^n z_i\right)^2 = \vec{z}^\top Q\vec{z}$, where we define Q as an $n \times n$ matrix of all ones. Thus, our problem can also be written as:

$$\begin{aligned} p_4^* = \quad \min_{\vec{x}, \vec{z} \in \mathbb{R}^n} \quad & \vec{x}^\top (A^\top A)\vec{x} - 2\vec{y}^\top A\vec{x} + \vec{y}^\top \vec{y} + \vec{z}^\top Q\vec{z} \\ \text{s.t.} \quad & -z_i \leq x_i \leq z_i, \text{ for each } i = 1, \cdots, n. \end{aligned}$$

3. A Minimum Time Path Problem

This question illustrates how to formulate an optimization problem as an SOCP. The problem studied in this question arises in optics. Consider Figure 1, in which a point in 0 must move to reach point $p = [4 \;\; 2.5]^\top$, crossing three layers of fluids having different densities. In the first layer, the point must travel at speed v_1, while in the second layer and third layers it must travel at lower maximum speeds, respectively $v_2 = v_1/\eta_2$, and $v_3 = v_1/\eta_3$, with $\eta_2, \eta_3 > 1$. Assume $v_1 = 1$, $\eta_2 = 1.5$, $\eta_3 = 1.2$. Formulate a SOCP whose objective is to find the fastest (i.e., minimum time) path from 0 to p. (It is not necessary to solve this SOCP). Denote with x_1, x_2, x_3 the horizontal coordinates of points p_1, p_2 and p, respectively, and with y_1, y_2, y_3 the corresponding vertical coordinates (which are given by $y_1 = 1$, $y_2 = 2$, $y_3 = 2.5$). Define as h_1, h_2, h_3 the lengths of the horizontal projections of the three legs, that is:

$$h_1 = x_1, \quad h_2 = x_2 - x_1, \quad h_3 = x_3 - x_2.$$

Hint: You may use path leg lengths ℓ_1, ℓ_2, ℓ_3 as variables, and observe that, in this problem, equality constraints of the type "something" $= \ell_i$ can be equivalently substituted for by inequality constraints "something" $\leq \ell_i$ (explain why).

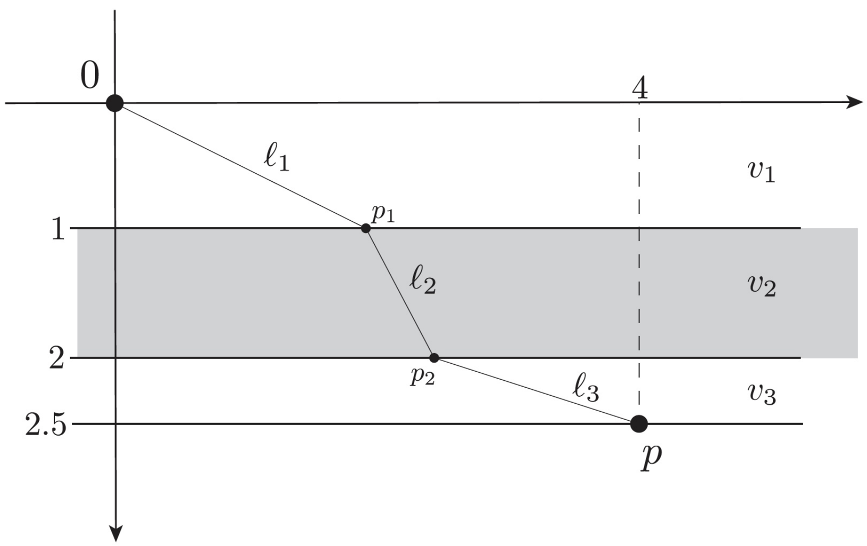

Figure 1: A minimum-time path problem.

Solution: Denote with x_1, x_2, x_3 the horizontal coordinates of points p_1, p_2 and p, respectively, and with y_1, y_2, y_3 the corresponding vertical coordinates (which are given by $y_1 = 1$, $y_2 = 2$, $y_3 = 2.5$). Define as h_1, h_2, h_3 the lengths of the horizontal projections of the three legs, that is:

$$h_1 = x_1, \quad h_2 = x_2 - x_1, \quad h_3 = x_3 - x_2.$$

We shall write the problem using the variables ℓ_1, ℓ_2, ℓ_3, and h_1, h_2, h_3, as variables. Note that by the Pythagoras theorem it must hold that

$$\begin{aligned}
\ell_1 &= \sqrt{h_1^2 + y_1^2} = \sqrt{h_1^2 + 1}, \\
\ell_2 &= \sqrt{h_2^2 + (y_2 - y_1)^2} = \sqrt{h_2^2 + 1}, \\
\ell_3 &= \sqrt{h_3^2 + (y_3 - y_2)^2} = \sqrt{h_3^2 + (0.5)^2}.
\end{aligned}$$

The total travel time is

$$T = \frac{\ell_1}{v_1} + \frac{\ell_2}{v_2} + \frac{\ell_3}{v_3}.$$

We hence set up our optimization problem as follows:

$$\begin{aligned} \min_{h,\ell\in\mathbb{R}} \quad & \tfrac{\ell_1}{v_1} + \tfrac{\ell_2}{v_2} + \tfrac{\ell_3}{v_3} \\ \text{s.t.} \quad & \ell_1 = \sqrt{h_1^2+1}, \\ & \ell_2 = \sqrt{h_2^2+1}, \\ & \ell_3 = \sqrt{h_3^2+(0.5)^2}, \\ & h_1+h_2+h_3 = 4. \end{aligned}$$

where the last constraint ensures that the horizontal coordinate of point p is equal to 4, as indicated in the figure. This formulation of the problem is not convex, due to the presence of nonlinear *equality* constraints. However, since we are minimizing a positive linear combination of the ℓ_i, we can substitute $=$ with $\geq$ in the constraints, obtaining the formulation

$$\begin{aligned} \min_{h,\ell\in\mathbb{R}} \quad & \tfrac{\ell_1}{v_1} + \tfrac{\ell_2}{v_2} + \tfrac{\ell_3}{v_3} \\ \text{s.t.} \quad & \ell_1 \geq \sqrt{h_1^2+1} = \left\| \begin{bmatrix} h_1 \\ 1 \end{bmatrix} \right\|_2, \\ & \ell_2 \geq \sqrt{h_2^2+1} = \left\| \begin{bmatrix} h_2 \\ 1 \end{bmatrix} \right\|_2, \\ & \ell_3 \geq \sqrt{h_3^2+(0.5)^2} = \left\| \begin{bmatrix} h_3 \\ 0.5 \end{bmatrix} \right\|_2, \\ & h_1+h_2+h_3 = 4. \end{aligned}$$

This problem is an SOCP, and the inequalities will be satisfied with equality at an optimum. Hence, this problem is equivalent to the original one.

4. A Slalom Problem

A skier must slide from left to right by going through n parallel gates of known positions (x_i, y_i) and widths c_i, $i = 1, \ldots, n$. The initial position (x_0, y_0) is given, as well as the final one, (x_{n+1}, y_{n+1}). Before reaching the final position, the skier must go through gate i by passing between the points $(x_i, y_i - c_i/2)$ and $(x_i, y_i + c_i/2)$ for each $i \in \{1, \ldots, n\}$. Figure 2 is a representation. Use values for (x_i, y_i, c_i) from Table 1.

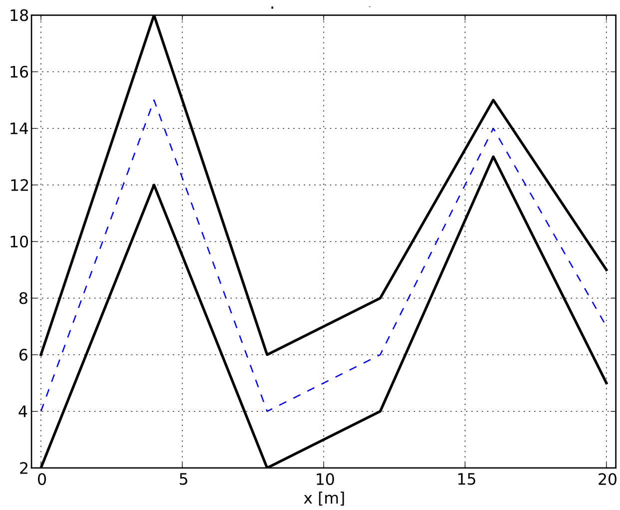

Figure 2: Slalom problem with $n = 5$ gates. The initial and final positions are fixed. The final position is not included in the figure. The skier slides from left to right. The middle path is dashed and connects the center points of gates.

Table 1: Problem data for Problem 2.

i	x_i	y_i	c_i
0	0	4	N/A
1	4	15	6
2	8	4	4
3	12	6	4
4	16	14	2
5	20	7	4
6	24	4	N/A

(a) Given the data $\{(x_i, y_i, c_i)\}_{i=0}^{n+1}$, write an optimization problem that minimizes the total length of the path. Your answer should come in the form of an SOCP.

Solution: Assume that (x_i, z_i) is the crossing point of gate i, the path length minimization problem is thus

$$\min_{\vec{z}} \quad \sum_{i=1}^{n+1} \left\| \begin{bmatrix} x_i \\ z_i \end{bmatrix} - \begin{bmatrix} x_{i-1} \\ z_{i-1} \end{bmatrix} \right\|_2 \tag{18}$$

$$\text{s.t.} \quad y_i - c_i/2 \leq z_i \leq y_i + c_i/2, \text{ for } i = 1, \ldots, n \tag{19}$$

$$z_0 = y_0, z_{n+1} = y_{n+1}, \tag{20}$$

which is equivalent to

$$\min_{\vec{z},\vec{t}} \quad \sum_{i=1}^{n+1} t_i \tag{21}$$

$$\text{s.t.} \quad y_i - c_i/2 \le z_i \le y_i + c_i/2, \text{ for } i = 0, \ldots, n+1 \tag{22}$$

$$\left\| \begin{bmatrix} x_i \\ z_i \end{bmatrix} - \begin{bmatrix} x_{i-1} \\ z_{i-1} \end{bmatrix} \right\|_2 \le t_i, \text{ for } i = 1, \ldots, n+1. \tag{23}$$

with the convention $c_0 = c_{n+1} = 0$. Hence, the problem is an SOCP.

(b) Solve the problem numerically with the data given in Table 1. You may use the starter code provided in the Jupyter notebook accompanying this HW. *HINT: You should be able to use packages such as* `cvxpy` *and* `numpy`*.*

Solution: The code can be found in the corresponding Jupyter notebook.

5. Least Squares with Equality Constraints

Consider the least squares problem with equality constraints

$$\min_{\vec{x}\in\mathbb{R}^n} \|A\vec{x}-\vec{b}\|_2^2 \;:\; G\vec{x}=\vec{h}, \tag{24}$$

where $A \in \mathbb{R}^{m\times n}$, $\vec{b}\in\mathbb{R}^m$, $G\in\mathbb{R}^{p\times n}$ and $\vec{h}\in\mathbb{R}^p$. For simplicity, we will assume that $\text{rank}(A)=n$ and $\text{rank}(G)=p$.

Using the KKT conditions, determine the optimal solution of this optimization problem.

Solution: This is a convex optimization problem with equality constraints. Let $\vec{\nu}\in\mathbb{R}^p$ denote the dual variables, which are unconstrained. If we can find a primal point $\vec{x}^*\in\mathbb{R}^n$ and a dual point $\vec{\nu}^*\in\mathbb{R}^p$ that together satisfy the KKT conditions then, because this is a convex optimization problem, we will learn that $\vec{x}^*$ is a primal optimal point, $\vec{\nu}^*$ is a dual optimal point, and strong duality holds. We will therefore write the KKT equations and attempt to solve them in order to find a primal optimal point.

The Lagrangian is

$$\begin{aligned} L(\vec{x},\vec{\nu}) &= (A\vec{x}-\vec{b})^\top(A\vec{x}-\vec{b}) + \vec{\nu}^\top(G\vec{x}-\vec{h}) &(25)\\ &= \vec{x}^\top A^\top A\vec{x} + (G^\top\vec{\nu}-2A^\top\vec{b})^\top\vec{x} - \vec{\nu}^\top\vec{h} + \vec{b}^\top\vec{b}. &(26)\end{aligned}$$

The KKT conditions are

$$\begin{aligned} G\vec{x}^* &= \vec{h}, &(27)\\ 2A^\top A\vec{x}^* + G^\top\vec{\nu}^* - 2A^\top\vec{b} &= \vec{0}, &(28)\end{aligned}$$

which are the primal feasilibity and the Lagrangian stationarity conditions respectively. Since the dual variables are unconstrained there is no dual feasiblity condition on ν^*, and since there are no inequality constraints there are no complementary slackness conditions. As for the Lagrangian stationarity conditions, they come from computing

$$\nabla_x L(\vec{x},\vec{\nu}^*) = 2A^\top A\vec{x} + G^\top\vec{\nu}^* - 2A^\top\vec{b}.$$

Since $\text{rank}(A)=n$, we can invert $A^\top A$, and from the Lagrangian stationary conditions in (28) we get

$$\vec{x}^* = (A^\top A)^{-1}\left(A^\top\vec{b} - \frac{1}{2}G^\top\vec{\nu}^*\right). \tag{29}$$

Substituting this in the primal feasibility condition in (27), we get

$$G(A^\top A)^{-1}A^\top\vec{b} - \frac{1}{2}G(A^\top A)^{-1}G^\top\vec{\nu}^* = \vec{h},$$

which gives

$$\vec{\nu}^* = -2\left(G(A^\top A)^{-1}G^\top\right)^{-1}\left(\vec{h} - G(A^\top A)^{-1}A^\top\vec{b}\right), \tag{30}$$

where we have used the assumption that $\text{rank}(G)=p$ to conclude that $G(A^\top A)^{-1}G^\top$ is invertible.

At this point we have found a primal point $\vec{x}^*$ and a dual point $\vec{\nu}^*$ that satisfy the KKT conditions, so we can claim that this $\vec{x}^*$ is primal optimal, this $\vec{\nu}^*$ is dual optimal, and strong duality holds. We can get an explicit formula for $\vec{x}^*$ by substituting (30) in (29), which gives

$$\begin{aligned} x^* &= (A^\top A)^{-1}\left(A^\top\vec{b} + G^\top\left(G(A^\top A)^{-1}G^\top\right)^{-1}\left(\vec{h} - G(A^\top A)^{-1}A^\top\vec{b}\right)\right)\\ &= (A^\top A)^{-1}A^\top\vec{b} + (A^\top A)^{-1}G^\top\left(G(A^\top A)^{-1}G^\top\right)^{-1}\left(\vec{h} - G(A^\top A)^{-1}A^\top\vec{b}\right).\end{aligned}$$

At this point we have already solved the problem, but it is instructive to discuss the solution a bit more.

To get some intuition for this solution, note that $\vec{\hat{x}} := (A^\top A)^{-1}A^\top\vec{b}$ would be the unique solution of the unconstrained least squares problem $\min_{\vec{x}\in\mathbb{R}^n}\|A\vec{x}-\vec{b}\|_2^2$ (it is the unique solution because we have assumed that $\text{rank}(A) = n$, so $\vec{\hat{x}}$ is uniquely defined by the condition that $A\vec{\hat{x}}$ equals the orthogonal projection of $\vec{b}$ onto the range space of A, i.e. $R(A)$). Also note that $\vec{h} - G(A^\top A)^{-1}A^\top\vec{b} = \vec{h} - G\vec{\hat{x}}$ can be viewed as an error term represent the amount by which $\vec{\hat{x}}$ fails to satisfy the constraint. To find the solution of the constrained problem we are therefore adjusting the solution to the unconstrained problem, i.e. $\vec{\hat{x}}$, by a vector which is determined by the amount by which this unconstrained solution fails to satisfy the constraint.

As for why the adjustment takes the form that it does, note that the constrained problem can be viewed as the problem of projecting b onto the intersection of $R(A)$ with the affine subspace $\{A\vec{x} : G\vec{x} = \vec{h}\}$, and $\vec{x}^*$ is indeed the unique vector such that $A\vec{x}^*$ equals this projection. Now, since the intersection of $R(A)$ with the affine subspace $\{A\vec{x} : G\vec{x} = \vec{h}\}$ is contained in $R(A)$, we see by the successive projection property of nested subspaces that $A\vec{x}^*$ is characterized by the property that $A\vec{x}^* - A\vec{\hat{x}}$ must be orthogonal to $A\vec{u}$ for every $\vec{u} \in N(G)$. To check that the $\vec{x}^*$ we found about satisfies this condition, note that

$$A\vec{x}^* - A\vec{\hat{x}} = A(A^\top A)^{-1}G^\top\left(G(A^\top A)^{-1}G^\top\right)^{-1}\left(\vec{h} - G(A^\top A)^{-1}A^\top\vec{b}\right),$$

and for $\vec{u} \in N(G)$ we can write

$$\begin{aligned}
&(A\vec{u})^\top A(A^\top A)^{-1}G^\top\left(G(A^\top A)^{-1}G^\top\right)^{-1}\left(\vec{h} - G(A^\top A)^{-1}A^\top\vec{b}\right)\\
&= \vec{u}^\top G^\top\left(G(A^\top A)^{-1}G^\top\right)^{-1}\left(\vec{h} - G(A^\top A)^{-1}A^\top\vec{b}\right)\\
&= 0,
\end{aligned}$$

where the last step comes from $G\vec{u} = 0$. This explains why $\vec{x}^*$ looks the way it does.

