MATH 565: Lecture 18 (03/12/2026)

Today: * projective gradient method (PGM) for linear constraints.
* convex quadratic program (CQP)

Recall: $\min\{F(\bar{w}) \mid \bar{w} \in C\}$

Let's consider $C = \{\bar{w} \in \mathbb{R}^d \mid A\bar{w} = \bar{b}\}$, i.e., we're solving

$$\begin{aligned} \min \ & F(\bar{w}) \\ \text{s.t.} \ & A\bar{w} = \bar{b} \end{aligned}$$

For this case, we can use projected gradient method (PGM) effectively.

We assume $F(\bar{w})$ is convex and $A_{m\times d}$ for $m \le d$ (else, $C = \emptyset$ is possible) and that rows of A are LI (i.e., there are no redundant constraints).

Idea

Consider $\bar{w}_t$ that is feasible, i.e., $\bar{w}_t \in C$ ($A\bar{w}_t = \bar{b}$).

Let $\bar{g}_t = \nabla F(\bar{w}_t)$. We set

$$\bar{w}_{t+1} = \bar{w}_t - \alpha \bar{g}_t$$

If $A\bar{g}_t \neq \bar{0}$, then $\bar{w}_{t+1} \notin C$, as

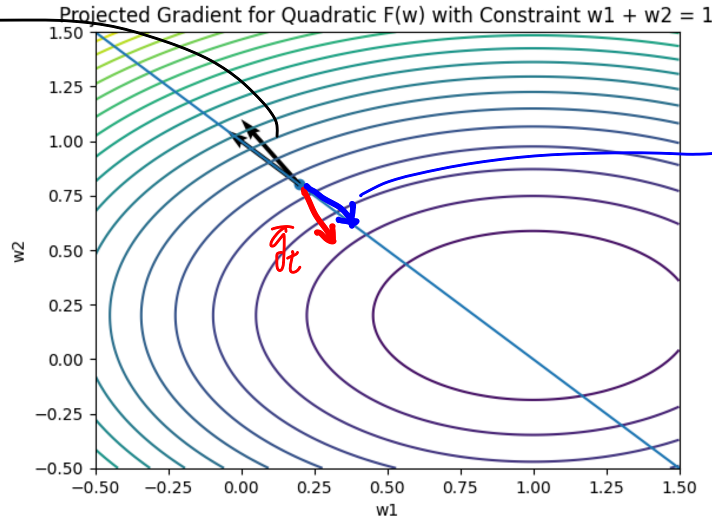
$$A\bar{w}_{t+1} = A(\bar{w}_t - \alpha\bar{g}_t) = \bar{b} - \alpha A\bar{g}_t \neq \bar{b}.$$

To ensure feasibility, we need to project $\bar{g}_t$ onto $A\bar{g}_t = \bar{0}$ (null space of A: $\mathcal{N}(A)$). Equivalently, we want to subtract the component of $\bar{g}_t$ orthogonal to $\mathcal{N}(A)$, or, that we want to subtract the component of A^T along $\bar{g}_t$.

Example $\quad \min F(\bar{w}) = (w_1-1)^2 + 2(w_2-0.2)^2$

s.t. $\quad w_1 + w_2 = 1 \qquad A = [1 \;\; 1], \quad b = 1$

∇J



projection of $\bar{g}_t$ onto C.
descent direction

Aside: How to project columns of A onto $\bar{b}$?

Solve $\quad \min_{\bar{w}} J(\bar{w}) = \frac{1}{2}\|A\bar{w}-\bar{b}\|^2$

$$\nabla J(\bar{w}) = \bar{0} \implies A^T(A\bar{w}-\bar{b}) = \bar{0}$$

$$\implies (A^TA)\bar{w} = A^T\bar{b}$$

$$\implies \bar{w} = (A^TA)^{-1}A^T\bar{b}.$$

$\implies$ The projection of A on $\bar{b}$ is $A\bar{w} = A(A^TA)^{-1}A^T\bar{b}$

Here, we want to find the projection of A^T onto $\bar{g}_t$, which is given by $\bar{g}_{\parallel} = A^T(AA^T)^{-1}A\bar{g}_t$.

Hence, we get the projection step as follows.

$$\Rightarrow \quad \text{Set } \bar{g}_t' = \bar{g}_t - \bar{g}_{\parallel} = \underbrace{\left(I - A^T (AA^T)^{-1} A\right)}_{P} \bar{g}_t$$
$$= P\bar{g}_t$$

where P is the projection matrix.

Note: P is indeed a projection matrix: $P^2 = P$, $P^T = P$. → verify!

But if rows of A are not LI, we can compute $\bar{g}_t'$ using GSO(rows of A) → Gram-Schmidt Orthogonalization.

Projective Gradient Method (PGM)

Step 0. Start with $\bar{w}_0 \in C$, i.e., $A\bar{w}_0 = \bar{b}$. → start with a feasible point

Step t.
$$\bar{g}_t = \nabla F(\bar{w}_t)$$
$$\bar{g}_t' = P\bar{g}_t \longrightarrow \text{stay feasible}$$
$$\bar{w}_{t+1} = \bar{w}_t - \alpha_t \bar{g}_t'$$

Repeat Step t until convergence.

Convex Quadratic Programs (CQPs)

We study the convex quadratic program (CQP) in the form

$$\left.\begin{aligned} \min \quad & J(\bar{w}) = \tfrac{1}{2}\bar{w}^T Q \bar{w} + \bar{p}^T \bar{w} + q \\ \text{s.t.} \quad & A\bar{w} = \bar{b} \end{aligned}\right\} \text{(CQP)}$$

where $Q_{d\times d} \succ 0$ (positive definite), $\bar{p} \in \mathbb{R}^d$, $q \in \mathbb{R}$.

J is strictly convex (as $HJ(\bar{w}) = Q \succ 0$).

Equality constraints (defining C) may be eliminated using Gaussian elimination (variable transformation).

Revisiting our example from earlier today,

$$\begin{aligned} \min \quad & J(\bar{w}) = (w_1 - 1)^2 + 2(w_2 - 0.2)^2 \\ \text{s.t.} \quad & w_1 + w_2 = 1 \end{aligned}$$

We can set $w_2 = 1 - w_1$ in J to get an equivalent unconstrained problem only in w_1:

$$\min \quad J(w_1) = (w_1 - 1)^2 + 2(0.8 - w_1)^2.$$

We now describe how to do this type of variable elimination in general. To this end, we use a variable transformation.

We write $Q = P \Delta P^T$ $\quad$ * Caution! P here is not the projection matrix

for Δ being a diagonal matrix with positive entries, so that $\Delta^{\frac{1}{2}}, \Delta^{-\frac{1}{2}}, \Delta^{-1}$ exist.

$Q = P\Delta P^T$ is the spectral decomposition of Q, where P has the orthonormal eigenvectors of Q, i.e., $P^T P = I \Rightarrow P^T = P^{-1}$, and $\Delta = [\mathrm{diag}(\lambda_i)]$, for eigenvalues $\lambda_1, \dots, \lambda_d > 0$, as $Q \succ 0$. Hence, $P \Delta^{-1} P^T = Q^{-1}$ in particular.

We want to get the form. $\quad$ $\min J(\bar{\omega}') = \frac{1}{2}\|\bar{\omega}'\|^2 + q'$
s.t. $A'\bar{\omega}' = b'$

$$\Rightarrow \quad J(\bar{\omega}) = \frac{1}{2} \bar{\omega}^T [P \Delta P^T] \bar{\omega} + \bar{p}^T \bar{\omega} + q$$

$$= \frac{1}{2} \|\underbrace{\Delta^{\frac{1}{2}} P^T \bar{\omega} + \Delta^{-\frac{1}{2}} P^T \bar{p}}_{\bar{\omega}'}\|^2 + \underbrace{\Big(q - \frac{1}{2} \bar{p}^T \underbrace{[P \Delta^{-1} P^T]}_{Q^{-1}} \bar{p}\Big)}_{q'}$$

With $q' = q - \frac{1}{2} \bar{p}^T [P \Delta^{-1} P^T] \bar{p}$, and setting

$$\bar{\omega}' = \Delta^{\frac{1}{2}} P^T \bar{\omega} + \Delta^{-\frac{1}{2}} P^T \bar{p}$$

the loss function becomes $J(\bar{\omega}') = \frac{1}{2}\|\bar{\omega}'\|^2 + q'$.

Let's double check the reformulation. The first term is

$$\frac{1}{2}\|\Delta^{\frac{1}{2}}P^T\bar{w} + \Delta^{-\frac{1}{2}}P^T\bar{p}\|^2 = \frac{1}{2}\left(\Delta^{\frac{1}{2}}P^T\bar{w} + \Delta^{-\frac{1}{2}}P^T\bar{p}\right)^2$$

$$= \frac{1}{2}\bar{w}^T \underbrace{P\Delta^{\frac{1}{2}}\Delta^{\frac{1}{2}}P^T}_{P\Delta P^T = Q}\bar{w} + \frac{1}{2}\bar{p}^T\underbrace{P\Delta^{-\frac{1}{2}}\Delta^{-\frac{1}{2}}P^T}_{P\Delta^{-1}P^T = Q^{-1}}\bar{p}$$

$$+ \frac{1}{2}\left(\bar{p}^T \underbrace{P\underbrace{\Delta^{-\frac{1}{2}}\Delta^{\frac{1}{2}}}_{I}P^T}_{I}\bar{w} + \bar{w}^T \underbrace{P\underbrace{\Delta^{\frac{1}{2}}\Delta^{-\frac{1}{2}}}_{I}P^T}_{I}\bar{p}\right)$$

$$= \frac{1}{2}\bar{w}^T Q\bar{w} + \bar{p}^T\bar{w} + \frac{1}{2}\bar{p}^T Q^{-1}\bar{p} = J(\bar{w}) - (\underbrace{q - \tfrac{1}{2}\bar{p}^T Q^{-1}\bar{p}}_{q'}) \quad \checkmark$$

Reformulation: we set

$$\bar{w}' = \Delta^{\frac{1}{2}}P^T\bar{w} + \Delta^{-\frac{1}{2}}P^T\bar{p} \quad \text{———} \quad (1)$$

$$\Rightarrow \quad \bar{w} = P\Delta^{-\frac{1}{2}}\bar{w}' - \underbrace{P\Delta^{-1}P^T}_{Q^{-1}}\bar{p}$$

$$\Rightarrow \quad \bar{w} = P\Delta^{-\frac{1}{2}}\bar{w}' - Q^{-1}\bar{p} \quad \text{———} \quad (2)$$

Also, $A\bar{w} = \bar{b}$ with (2) gives

$$A(P\Delta^{-\frac{1}{2}}\bar{w}' - Q^{-1}\bar{p}) = \bar{b}$$

$\Rightarrow \quad AP\Delta^{-\frac{1}{2}}\bar{w}' = \bar{b} + Q^{-1}\bar{p}$, which is $A'\bar{w}' = \bar{b}'$

where $\quad A' = AP\Delta^{-\frac{1}{2}} \quad$ ——— (3)

and $\quad \bar{b}' = \bar{b} + Q^{-1}\bar{p} \quad$ ——— (4)

Hence, the reformulation is written as

$$\min \; J(\bar{w}') = \frac{1}{2}\|\bar{w}'\|^2 + q'$$
$$\text{s.t. } A\bar{w}' = \bar{b}'$$

<u>Aside</u>: How to solve $\left\{ \begin{array}{l} \min \|\bar{w}\|^2 \\ \text{s.t. } A\bar{w} = \bar{b} \end{array} \right\}$?

Any solution should satisfy $A\bar{w} = \bar{b}$, and hence can be decomposed as $\bar{w} = \bar{w}_{row} + \bar{w}_{null}$ where $A\bar{w}_{null} = \bar{0}$, i.e., $\bar{w}_{null} \in \mathcal{N}(A)$, and $\bar{w}_{row} \perp \bar{w}_{null}$. Hence,

orthogonal

$\|\bar{w}\|^2 = \|\bar{w}_{row}\|^2 + \|\bar{w}_{null}\|^2$.

$\Rightarrow$ Hence, to minimize $\|\bar{w}\|^2$, we set $\bar{w}_{null} = \bar{0}$.

$\Rightarrow$ We look for $\bar{w}_{row} = A^T\bar{y}$, a vector in the row space of A.

$\Rightarrow$ We need $A\bar{w}_{row} = AA^T\bar{y} = \bar{b}$

$$\Rightarrow \bar{y} = (AA^T)^{-1}\bar{b}$$

$$\Rightarrow \bar{w} = \bar{w}_{row} = A^T\bar{y} = \underbrace{A^T(AA^T)^{-1}}\bar{b}$$

This is also known as the $\underbrace{\text{right inverse}}$ formula.

$$A\,A^T(AA^T)^{-1} = I$$

One can also derive this result using Lagrangian relaxation.

Applying this result, we get the solution as

$$\bar{w}' = A'^T (A'A'^T)^{-1} \bar{b}'$$

We can rewrite this solution in terms of $\bar{w}$. First,

(3) gives $$A'^T = (AP\Delta^{-1/2})^T = \Delta^{-1/2} P^T A^T$$

$$\Rightarrow A'A'^T = A\underbrace{P\Delta^{-1/2}\Delta^{-1/2}P^T}_{Q^{-1}}A^T = AP\Delta^{-1}P^TA^T$$

$$= AQ^{-1}A^T$$

Hence,

$$(2) \Rightarrow \bar{w} = P\Delta^{-1/2}\bar{w}' - Q^{-1}\bar{p}$$

$$= -Q^{-1}\bar{p} + \underbrace{P\Delta^{-1/2}\Delta^{-1/2}P^T}_{Q^{-1}}A^T[AQ^{-1}A^T]^{-1}\bar{b}'$$

$$= \underbrace{-Q^{-1}\bar{p}}_{\text{solution to unconstrained minimization of } J(\bar{w})} + \underbrace{Q^{-1}\left(A^T[AQ^{-1}A^T]^{-1}\overbrace{(\bar{b} + Q^{-1}\bar{p})}^{\text{by (4)}}\right)}_{\text{adjustment terms for constraints } A\bar{w} = \bar{b}}$$