Lecture 27: Multiagent Consensus Equilibrium

Charles A. Bouman

Purdue University

Wednesday, October 28, 2026

Outline

1. Two agents to K: weighted cost, proximal maps, MACE equations
2. Proximal-map agents and general agents
3. The weights μ_k: rebalancing the forward and prior agents
4. Solving MACE equations: fixed point of $\mathbf{T}$, MACE-DR

Reading: Section 10.4.

Today's Big Ideas

- MACE mixes proximal-map agents with cost-free action agents.
- Every agent's output equals the same weighted average.
- Proximal-map agents make MACE minimize the weighted cost.
- Weights change regularization without retraining any agent.
- The averaging is idempotent, so Douglas-Rachford solves MACE.

From two agents to many

We now solve for the CE solution with **multiple heterogeneous agents**.

- Some agents are **proximal maps** of objective functions.
- Some agents are **operations that perform actions** directly.

Definition

This problem is called **multiagent consensus equilibrium (MACE)**.

To understand MACE, we start with the special case of K agents, each with a cost function $f_k(x)$.

The weighted cost function

In this special case, the complete cost function is

$$f(x) = \sum_{k=1}^{K} \mu_k \, f_k(x) \ , \tag{10.23}$$

where the weights μ_k satisfy $1 = \sum_{k=1}^{K} \mu_k$.

- Each $f_k(x)$ may be a soft or hard constraint.

The desired solution is

$$x^* = \arg\min_x \left\{ \sum_{k=1}^{K} \mu_k \, f_k(x) \right\} \ . \tag{10.24}$$

The K proximal maps

Each of the K agents has an associated proximal map:

$$\begin{aligned} F_1(v) &= \arg\min_x \left\{ f_1(x) + \frac{1}{2\sigma^2}\|x-v\|^2 \right\} , \\ &\vdots \\ F_K(v) &= \arg\min_x \left\{ f_K(x) + \frac{1}{2\sigma^2}\|x-v\|^2 \right\} . \end{aligned}$$

- σ^2 is a **user-selectable parameter**.
- All K maps use the same σ^2.
- The choice of σ^2 does not affect the final solution.

Stacked notation

Define the stacked state and the stacked operator:

$$\mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_K \end{bmatrix}, \qquad \mathbf{F}(\mathbf{w}) = \begin{bmatrix} F_1(w_1) \\ \vdots \\ F_K(w_K) \end{bmatrix}.$$

Define the stacked **weighted averaging operator**:

$$\mathbf{G}_\mu(\mathbf{w}) = \begin{bmatrix} \bar{w} \\ \vdots \\ \bar{w} \end{bmatrix}, \qquad \text{where} \quad \bar{w} = \sum_{k=1}^{K} \mu_k \, w_k \ .$$

The MACE equations

The **MACE equations** are

$$\mathbf{F}(\mathbf{w}^*) = \mathbf{G}_\mu(\mathbf{w}^*) \ , \tag{10.25}$$

and the solution is

$$x^* = \sum_{k=1}^{K} \mu_k \, w_k^* \ .$$

- Every agent output equals the same weighted average.

$$F_k(w_k^*) = x^* \quad \text{for } k = 1, \ldots, K$$

- The reconstruction x^* is that weighted average.

Two agents as a special case

Set $K = 2$, $\mu_1 = \mu_2 = 1/2$, $F_1 = F$, and $F_2 = H$.

Then $\mathbf{G}_\mu$ is the plain average

$$\mathbf{G}_\mu(\mathbf{w}) = \begin{bmatrix} \bar{w} \\ \bar{w} \end{bmatrix}, \qquad \bar{w} = \frac{w_1 + w_2}{2},$$

and the MACE equations become

$$\begin{aligned} F(w_1^*) &= \bar{w}^*, \\ H(w_2^*) &= \bar{w}^*. \end{aligned}$$

- These are the two-agent CE equations in parallel form.
- MACE adds more agents and unequal weights.

Proximal-map agents: the zero-gradient condition

When the agents are proximal maps, differentiating the MACE equations shows that

$$\mathbf{0} \in \nabla f(x)|_{x=x^*} \quad , \tag{10.26}$$

where x^* is the solution to the MACE equations.

Property

If every $f_k(x)$ is a **proper closed convex** function, then x^* is a MACE solution **if and only if** it is the minimum of $f(x) = \sum_{k=1}^{K} \mu_k f_k(x)$.

Deriving the zero-gradient condition

The MACE equations say $F_k(w_k^*) = x^*$ for every k.

The zero-gradient condition of each proximal map gives

$$\mathbf{0} \in \nabla f_k(x^*) + \frac{1}{\sigma^2}(x^* - w_k^*) \quad \text{for } k = 1, \ldots, K \ .$$

Multiply by $\sigma^2 \mu_k$ and sum over k:

$$\mathbf{0} \in \sigma^2 \sum_{k=1}^{K} \mu_k \nabla f_k(x^*) + \underbrace{\sum_{k=1}^{K} \mu_k x^*}_{x^*} - \underbrace{\sum_{k=1}^{K} \mu_k w_k^*}_{x^*} = \sigma^2 \nabla f(x^*) \ .$$

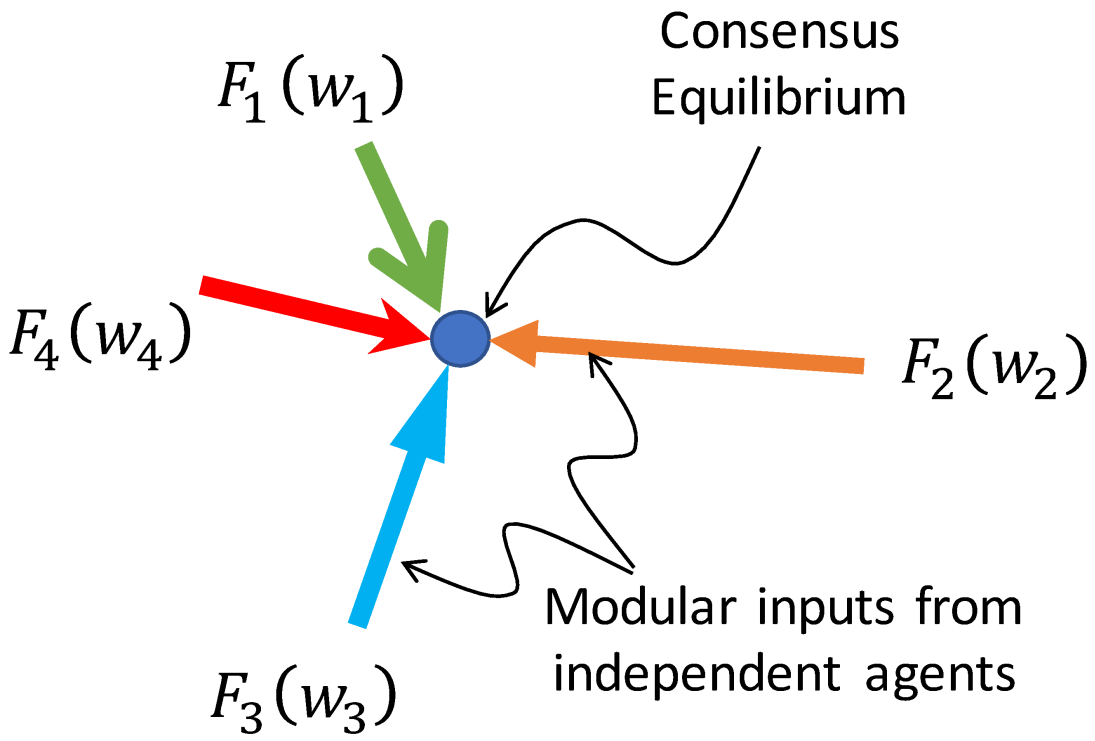
- The two weighted sums cancel because $1 = \sum_k \mu_k$ and $x^* = \sum_k \mu_k w_k^*$.
- σ^2 drops out, so it does not affect the solution.

General agents: equilibrium without a cost function

The power of the MACE equations comes from **plug-in agents F_k that are not proximal maps**.

- The MACE solution then does **not** minimize any cost.
- It is an **equilibrium that balances the agents' actions**.
- Agents may conflict. Each agent moves the state toward its own goal.

A graphical interpretation of the MACE solution



- Each agent moves its state w_k to the common point x^*.
- The solution also satisfies $x^* = \sum_{k=1}^{K} \mu_k F_k(w_k)$.
- The agents' forces balance here; this need not minimize a cost.

The weights μ_k

The weights set the **relative strength** of the agents.

- Changing μ_k changes only the weighted average in $\mathbf{G}_\mu$.
- Forward and prior weights set the **regularization**.

Why this matters

Changing the weights does not change the agents F_k.

- Learned agents may need **retraining** after each change.
- The weights avoid that retraining.

Example: rebalancing the forward and prior agents

Take $K = 2$ proximal-map agents with weights $\mu_1 = \mu$ and $\mu_2 = 1 - \mu$:

$$f_1(x) = \frac{1}{2}\|y - Ax\|^2 \text{ (forward)}, \qquad f_2(x) = \frac{1}{2}\|x\|^2 \text{ (prior)}.$$

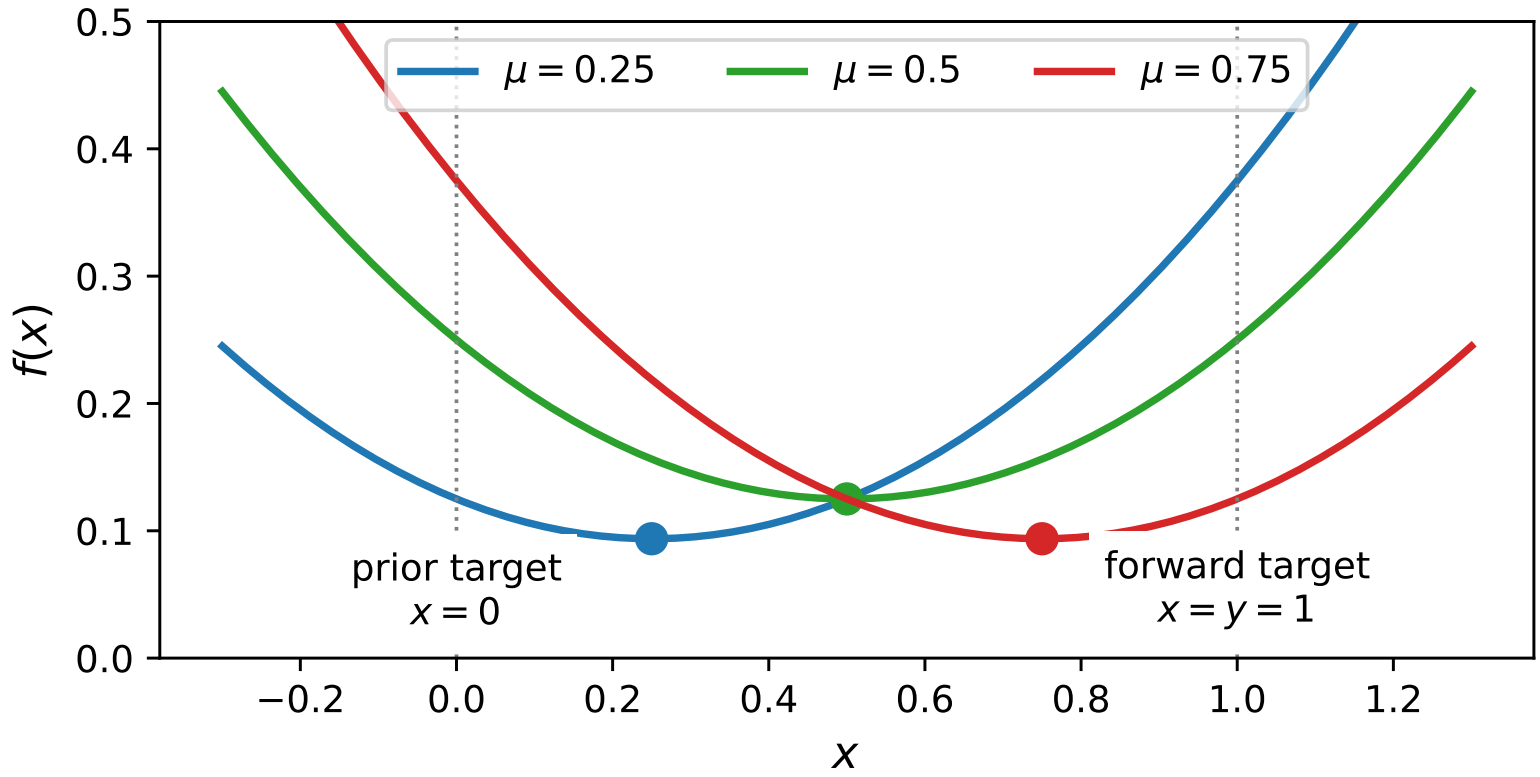
Setting the gradient of $\mu f_1(x) + (1-\mu) f_2(x)$ to zero gives

$$x^* = \left(A^t A + \frac{1-\mu}{\mu} I\right)^{-1} A^t y .$$

- The ratio $(1-\mu)/\mu$ acts as the regularization parameter.
- $\mu \to 1$: forward dominates, and regularization vanishes.
- $\mu \to 0$: the prior agent dominates, and $x^* \to 0$.
- Neither agent changes. Only the weighted average changes.

Example: rebalancing the forward and prior agents

Scalar case with $A = 1$ and $y = 1$: the solution is $x^* = \mu\, y$.



- Each curve is $f(x) = \mu f_1(x) + (1 - \mu) f_2(x)$.
- The dot, $x^* = \mu$, moves from prior toward forward as μ grows.

The weighted averaging operator is idempotent

For any choice of weights, applying $\mathbf{G}_\mu$ twice gives the same result:

$$\mathbf{G}_\mu \mathbf{G}_\mu = \mathbf{G}_\mu \ .$$

- Weights summing to 1 leave $\bar{w}$ unchanged.

Since $\mathbf{G}_\mu$ is linear, it follows that

$$(2\mathbf{G}_\mu - I)^2 = 4\mathbf{G}_\mu \mathbf{G}_\mu - 4\mathbf{G}_\mu + I = I \ ,$$

so $(2\mathbf{G}_\mu - I)^{-1} = (2\mathbf{G}_\mu - I)$.

The MACE solution is a fixed point of $\mathbf{T}$

Using $(2\mathbf{G}_\mu - I)^{-1} = (2\mathbf{G}_\mu - I)$:

$$\begin{aligned} \mathbf{F}\mathbf{w}^* &= \mathbf{G}_\mu \mathbf{w}^* \ , \\ (2\mathbf{F} - I)\mathbf{w}^* &= (2\mathbf{G}_\mu - I)\mathbf{w}^* \ , \\ (2\mathbf{G}_\mu - I)(2\mathbf{F} - I)\mathbf{w}^* &= \mathbf{w}^* \ . \end{aligned}$$

So $\mathbf{w}^*$ solves the MACE equations if and only if

$$\mathbf{T}\mathbf{w}^* = \mathbf{w}^* \ , \qquad \text{where} \quad \mathbf{T} = (2\mathbf{G}_\mu - I)(2\mathbf{F} - I) \ .$$

- Solve for the fixed point with Mann iterations, $\mathbf{w} \leftarrow \rho\mathbf{T}\mathbf{w} + (1-\rho)\mathbf{w}$, with $\rho \in (0, 1)$.
- The iterations converge when $\mathbf{T}$ is nonexpansive.
- For $K = 2$ and equal weights, $\mathbf{T}$ is the two-model operator.

Unequal weights: $\mathbf{G}_\mu$ may be expansive

- When $\mu_k = 1/K$, $\mathbf{G}_\mu$ is a **projection**, and $(2\mathbf{G}_\mu - I)$ is nonexpansive.
- Unequal weights: $\mathbf{G}_\mu$ may be expansive (a chapter exercise).
- Then $\mathbf{T}$ may not be nonexpansive, even with proximal F_k.
- In theory, MACE-DR may then **fail to converge**.

In practice

This failure does not seem to occur.

The MACE-DR algorithm

MACE-DR Algorithm

Initialize Reconstruction: $x^{(0)} \in \mathbb{R}^N$

$$\mathbf{w} \leftarrow \begin{bmatrix} x^{(0)} \\ \vdots \\ x^{(0)} \end{bmatrix}$$

Until Converged $\{$

$$\begin{aligned} \mathbf{x} &\leftarrow \mathbf{F}(\mathbf{w}) \\ \mathbf{z} &\leftarrow \mathbf{G}_\mu(2\mathbf{x} - \mathbf{w}) \\ \mathbf{w} &\leftarrow \mathbf{w} + 2\rho(\mathbf{z} - \mathbf{x}) \end{aligned}$$

$\}$

$x^* \leftarrow \mathbf{z}_1$

$\mathrm{Return}(x^*)$

MACE-DR solves the MACE equations with DR and Mann iterations. It provably converges when $\mathbf{T} = (2\mathbf{G}_\mu - I)(2\mathbf{F} - I)$ is nonexpansive.

MACE-DR is the Mann iteration on $\mathbf{T}$

With $\mathbf{x} = \mathbf{F}(\mathbf{w})$ and $\mathbf{z} = \mathbf{G}_\mu(2\mathbf{x} - \mathbf{w})$:

$$\begin{aligned} \mathbf{T}\mathbf{w} &= (2\mathbf{G}_\mu - I)(2\mathbf{x} - \mathbf{w}) \\ &= 2\mathbf{z} - 2\mathbf{x} + \mathbf{w} \ . \end{aligned}$$

The Mann iteration is then

$$\rho\mathbf{T}\mathbf{w} + (1-\rho)\mathbf{w} = \mathbf{w} + \rho(\mathbf{T}\mathbf{w} - \mathbf{w}) = \mathbf{w} + 2\rho(\mathbf{z} - \mathbf{x}) \ .$$

- Every $\mathbf{z}$ entry is one average; $\mathbf{z}_1$ is the result.
- The K updates in $\mathbf{F}(\mathbf{w})$ can run in parallel.

What the MACE-DR algorithm computes

- With proximal F_k and nonexpansive $\mathbf{T}$, it finds x^*:

$$x^* = \arg\min_x \left\{ \sum_{k=1}^{K} \mu_k \, f_k(x) \right\} .$$

- It works with **any mix** of objective and action agents.
- If $\mathbf{T}$ is nonexpansive, the solution balances all forces.

Summary

- MACE combines K heterogeneous agents F_k with weights $1 = \sum_{k=1}^{K} \mu_k$.
- The MACE equations are $\mathbf{F}(\mathbf{w}^*) = \mathbf{G}_\mu(\mathbf{w}^*)$, with reconstruction $x^* = \sum_{k=1}^{K} \mu_k w_k^*$.
- For proper closed convex f_k, $\mathbf{0} \in \nabla f(x)|_{x=x^*}$, so $x^* = \arg\min_x \sum_{k=1}^{K} \mu_k f_k(x)$.
- Quadratic example: regularization $(1-\mu)/\mu$; no agent retrained.
- $\mathbf{G}_\mu \mathbf{G}_\mu = \mathbf{G}_\mu$, so $\mathbf{w}^*$ is the fixed point of $\mathbf{T} = (2\mathbf{G}_\mu - I)(2\mathbf{F} - I)$, computed by MACE-DR.

Next lecture: generative Plug-and-Play.