Underdetermined Linear Systems and Regularization

Underdetermined Linear Systems

Consider the linear system

$$Ac = y, \qquad A \in \mathbb{R}^{m \times n},\ c \in \mathbb{R}^n,\ y \in \mathbb{R}^m,$$

with

$$\boxed{m < n.}$$

Underdetermined Linear Systems

Consider the linear system

$$Ac = y, \qquad A \in \mathbb{R}^{m\times n},\ c \in \mathbb{R}^n,\ y \in \mathbb{R}^m,$$

with

$$\boxed{m < n.}$$

Underdetermined system

There are fewer equations than unknowns.

Geometric Interpretation

- Each equation defines a hyperplane in $\mathbb{R}^n$.
- With only $m < n$ hyperplanes, a single point cannot be isolated.
- The intersection is typically an affine subspace.

Geometric Interpretation

- Each equation defines a hyperplane in $\mathbb{R}^n$.
- With only $m < n$ hyperplanes, a single point cannot be isolated.
- The intersection is typically an affine subspace.

$$\boxed{\dim(\text{solution set}) \geq n - m}$$

Solution Set Structure

If the system is consistent ($y \in \mathrm{Im}(A)$), then:

$$\boxed{c = c_0 + z, \qquad z \in \ker(A)}$$

Solution Set Structure

If the system is consistent ($y \in \mathrm{Im}(A)$), then:

$$\boxed{c = c_0 + z, \qquad z \in \ker(A)}$$

- c_0: any particular solution,
- $\ker(A)$ is nontrivial.

Solution Set Structure

If the system is consistent ($y \in \mathrm{Im}(A)$), then:

$$\boxed{c = c_0 + z, \qquad z \in \ker(A)}$$

- c_0: any particular solution,
- $\ker(A)$ is nontrivial.

Key consequence

There are infinitely many solutions.

Ill-Posedness and Instability

If $y \notin \mathrm{Im}(A)$:

- no exact solution exists,
- many solutions fit the data almost equally well.

Ill-Posedness and Instability

If $y \notin \mathrm{Im}(A)$:

- no exact solution exists,
- many solutions fit the data almost equally well.

Practical consequences

- model is **non-identifiable**,
- tiny perturbations in y cause large changes in c,
- solutions are unstable.

Why This Matters

Underdetermined systems are common in modern data analysis:

- biology: gene expression (n genes $\gg m$ samples),
- signal processing: high-dimensional signals,
- machine learning: many features, few observations.

Why This Matters

Underdetermined systems are common in modern data analysis:

- biology: gene expression (n genes $\gg m$ samples),
- signal processing: high-dimensional signals,
- machine learning: many features, few observations.

Conclusion

Underdetermination is the rule, not the exception.

Regularization: General Idea

To recover a meaningful solution, we impose **additional constraints**:

- sparsity,
- smoothness,
- non-negativity,
- structural assumptions.

Regularization: General Idea

To recover a meaningful solution, we impose **additional constraints**:

- sparsity,
- smoothness,
- non-negativity,
- structural assumptions.

Regularization

Modify the optimization problem so that a **unique and stable** solution exists.

Tikhonov (L2) Regularization

Instead of solving $Ac = y$, we solve:

$$\boxed{c_\lambda = \arg\min_c \left\{ \|Ac - y\|_2^2 + \lambda \|c\|_2^2 \right\}}$$

Tikhonov (L2) Regularization

Instead of solving $Ac = y$, we solve:

$$\boxed{c_\lambda = \arg\min_c \left\{ \|Ac - y\|_2^2 + \lambda \|c\|_2^2 \right\}}$$

- data fidelity: $\|Ac - y\|_2^2$,
- regularization: $\lambda \|c\|_2^2$,
- $\lambda > 0$ controls the trade-off.

Interpretation of Tikhonov Regularization

- penalizes large coefficients,
- selects the smallest-norm solution,
- stabilizes ill-posed problems.

Interpretation of Tikhonov Regularization

- penalizes large coefficients,
- selects the smallest-norm solution,
- stabilizes ill-posed problems.

Limit case

As $\lambda \to 0$, the solution converges to the minimum-norm least-squares solution.

Derivation of Tikhonov Normal Equations

Define:

$$J(c) = \|Ac - y\|_2^2 + \lambda \|c\|_2^2 = (Ac - y)^\top (Ac - y) + \lambda c^\top c$$

Derivation of Tikhonov Normal Equations

Define:

$$J(c) = \|Ac - y\|_2^2 + \lambda\|c\|_2^2 = (Ac - y)^\top (Ac - y) + \lambda c^\top c$$

Gradient:

$$\nabla_c J(c) = 2A^\top (Ac - y) + 2\lambda c$$

Derivation of Tikhonov Normal Equations

Define:

$$J(c) = \|Ac - y\|_2^2 + \lambda \|c\|_2^2 = (Ac - y)^\top (Ac - y) + \lambda c^\top c$$

Gradient:

$$\nabla_c J(c) = 2A^\top (Ac - y) + 2\lambda c$$

Setting $\nabla_c J(c) = 0$ yields:

$$A^\top A c + \lambda c = A^\top y$$

Tikhonov Normal Equations

$$\boxed{(A^\top A + \lambda I)c = A^\top y}$$

Tikhonov Normal Equations

$$\boxed{(A^\top A + \lambda I)c = A^\top y}$$

Key property

For $\lambda > 0$, $A^\top A + \lambda I$ is always invertible.

Explicit Tikhonov Solution

$$\boxed{c_\lambda = (A^\top A + \lambda I)^{-1} A^\top y}$$

Explicit Tikhonov Solution

$$\boxed{c_\lambda = (A^\top A + \lambda I)^{-1} A^\top y}$$

- unique solution,
- exists even if A is rank-deficient,
- numerically stable.

Why Tikhonov Works

- $A^\top A$ may be ill-conditioned or singular,
- adding λI shifts eigenvalues.

Why Tikhonov Works

- $A^\top A$ may be ill-conditioned or singular,
- adding λI shifts eigenvalues.

If $\sigma_i(A)$ are singular values of A:

$$\sigma_i(A^\top A + \lambda I) = \sigma_i(A)^2 + \lambda$$

Why Tikhonov Works

- $A^\top A$ may be ill-conditioned or singular,
- adding λI shifts eigenvalues.

If $\sigma_i(A)$ are singular values of A:

$$\sigma_i(A^\top A + \lambda I) = \sigma_i(A)^2 + \lambda$$

Effect

Improves conditioning and suppresses noise amplification.

Other Regularization Techniques

LASSO (L1 regularization)

$$\min_c \|Ac - y\|_2^2 + \lambda \|c\|_1$$

Other Regularization Techniques

LASSO (L1 regularization)

$$\min_c \|Ac - y\|_2^2 + \lambda\|c\|_1$$

Elastic Net

$$\|Ac - y\|_2^2 + \alpha\|c\|_1 + \beta\|c\|_2^2$$

Other Regularization Techniques

LASSO (L1 regularization)

$$\min_c \|Ac - y\|_2^2 + \lambda \|c\|_1$$

Elastic Net

$$\|Ac - y\|_2^2 + \alpha \|c\|_1 + \beta \|c\|_2^2$$

Constraints-based methods

$$c \geq 0, \; \|c\|_0 \leq k, \; \text{monotonicity, group structure}$$

Summary

- Underdetermined systems ($m < n$) admit infinitely many solutions
- Solutions are unstable and non-identifiable
- Regularization restores uniqueness and stability
- Tikhonov regularization leads to a well-posed problem

Summary

- Underdetermined systems ($m < n$) admit infinitely many solutions
- Solutions are unstable and non-identifiable
- Regularization restores uniqueness and stability
- Tikhonov regularization leads to a well-posed problem

Take-home message

Regularization is essential in high-dimensional inverse problems.