

Regularization of a finite-dimensional problem

As the most of variational methods, the regularization has a finite-dimensional analog. It is applicable to the minimization problem of a convex but not strongly convex function which may have infinitely many solutions. The idea of regularization is to slightly perturb the function by small but a strictly convex term; the perturbed problem has a unique solution to matter how small the perturbation is. The numerical advantage of the regularization is the convergence of minimizing sequences.

Let us illustrate ideas of regularization by studying a finite dimensional problem. Consider a linear system

$$Ax = b \tag{1}$$

where A is a square $n \times n$ matrix and b is a known n -vector.

We know from linear algebra that the Fredholm Alternative holds:

- If $\det A \neq 0$, the problem has a unique solution:

$$x = A^{-1}b \quad \text{if } \det A \neq 0 \tag{2}$$

- If $\det A = 0$ and $Ab \neq 0$, the problem has no solutions.
- If $\det A = 0$ and $Ab = 0$, the problem has infinitely many solutions.

In practice, we also deal with an additional difficulty: The determinant $\det A$ may be a “very small” number and one cannot be sure whether its value is a result of rounding of digits or it has a “physical meaning.” In any case, the errors of using the formula (2) can be arbitrary large and the norm of the solution is not bounded.

To address this difficulties, it is helpful to restate linear problem (1) as an extremal problem:

$$\min_{x \in R^n} (Ax - b)^2 \tag{3}$$

This problem does have at least one solution, no matter what the matrix A is. This solution coincides with the solution of the original problem (1) when this problem has a unique solution; in this case the cost of the minimization problem (3) is zero. Otherwise, the minimization problem provides “the best approximation” of the non-existing solution.

If the problem (1) has infinitely many solutions, so does problem (3). Corresponding minimizing sequences $\{x^s\}$ can be unbounded, $\|x^s\| \rightarrow \infty$ when $s \rightarrow \infty$.

In this case, we may select a solution with minimal norm. We use the *regularization*, passing to the perturbed problem

$$\min_{x \in R^n} (Ax - b)^2 + \epsilon x^2$$

The solution of the last problem exists and is unique. Indeed, we have by differentiation

$$(A^T A + \epsilon I)x - A^T b = 0$$

and

$$x = (A^T A + \epsilon I)^{-1} A^T b$$

We mention several properties of regularization.

1. The inverse $(A^T A + \epsilon I)^{-1}$ exists since the matrix $A^T A$ is nonnegative defined, and ϵ is positively defined. The eigenvalues of the matrix $(A^T A + \epsilon I)^{-1}$ are not smaller than ϵ^{-1}
2. Suppose that we are dealing with a well-posed problem (1), that is the matrix A is not degenerate. If $\epsilon \ll 1$, the solution approximately is $x = A^{-1}b - \epsilon(A^2 A^T)^{-1}b$. When $\epsilon \rightarrow 0$, the solution becomes the solution (2) of the unperturbed problem, $x \rightarrow A^{-1}b$.
3. If the problem (1) is ill-posed, the norm of the solution of the perturbed problem is still bounded:

$$\|x\| = \|(A^T A + \epsilon I)^{-1}\| \|b\| \leq \frac{1}{\epsilon} \|b\|$$

Instead of the regularizing term ϵx^2 , we may use any positively define quadratic $\epsilon(x^T P x + p^T x)$ or where matrix P is positively defined, $P > 0$, or $\epsilon((x-c)^T P(x-c))$ (the attractor to the target point c), or another strongly convex function of x .