Submitted By: Henna Umar (s0453772)

Regularization

According to *Hadamard, 1915*: Given mapping A: X-->Y, equation

$$Ax = y$$

is *well-posed* provided

- a solution exists for each $y \in Y$, $\exists x \in X$ such that $Ax = y$
- the solution is unique i.e. $Ax_1 = Ax_2 \Rightarrow x_1 = x_2$
- the solution is stable i.e. A^{-1} is continuous

An equation is *ill-posed* if it is not *well-posed*

Informally, regularization is defined as it *'Im poses stability on an ill-posed problem in a manner that yields accurate approximate solutions, often by incorporating prior information"*[1]. More formal definition is that it is a parametric family of approximate inverse operators $R_\alpha : Y \to X$ with the following property. If $y_\eta = Ax_{\text{true}} + \eta_\eta$, and $\eta_\eta \to 0$ we can pick parameters α_η such that

$$x_{\alpha_\eta} \stackrel{\text{def}}{=} R_{\alpha_\eta} y_\eta \to x_{\text{true}}$$

Tikhonov's regularization

In simplest case, assume X, Y are Hilbert spaces. To obtain regularized solution to $Ax = y$, choose x to fit data y in least-squares sense, but penalize solutions of large norm. Solve minimization problem

$$x_\alpha = \arg\min_{x \in X} \|Ax - y\|_Y^2 + \alpha \|x\|_X^2$$

$$= \underbrace{(A^* A + \alpha I)^{-1} A^*}_{R_\alpha} y$$

$\alpha > 0$ is called the *regularization parameter.*

Singular Value Decomposition

It is an important tool for analysis and computation as it gives bi-orthogonal diagonalization of linear operator, A = USV*

In an nxn matrix case, $U = [u_1,\ldots,u_n]$, $S = \text{diag}(s_1,\ldots,s_n)$, and $V = [v_1,\ldots,v_n]$ with

$$s_1 \geq s_2 \geq \ldots \geq s_n \geq 0,$$

$$Av_i = su_i, \quad A^* u_i = s_i v_i,$$

$$(u_i, u_j) = \delta_{ij}, \quad (v_i, v_j) = \delta_{ij} \Rightarrow U^* U = I, \quad V^* V = I$$

Tikhonov Filtering

In the case of Tikhonov regularization, using the SVD, A = USV* (and assuming n x n matrix with $s_i > 0$ for simplicity),

$$R_\alpha = (A^* A + \alpha I)^{-1} A^*$$

$$= (VS^* U^* USV^* + \alpha V I V^*)^{-1} VS^* U^*$$

$$= V (S^* S + \alpha I)^{-1} S^* U^*$$

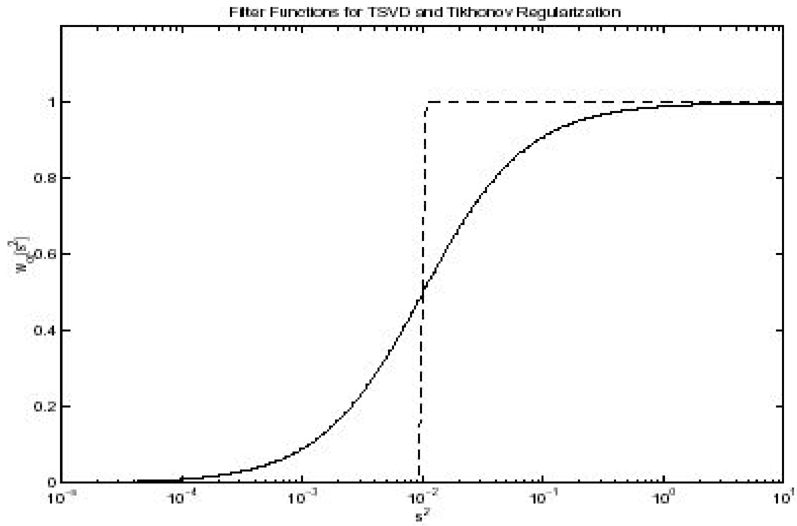
1 "Regularization Methods" by Curt Vogel

$$= \text{V diag}\Big(\underbrace{\frac{s_i^2}{s_i^2+\alpha}}_{\omega_\alpha(s_i^2)}\frac{1}{s_i}\Big)U^*$$

If $\alpha \to 0$, then $\omega_\alpha(s_i^2) \to 1$, so

$$R_\alpha \to \text{V diag}(1/s_i)U^* \stackrel{\text{def}}{=} A^\dagger \quad \text{as } \alpha \to 0$$

Plot of Tikhonov filter function $\omega_\alpha^{\text{Tikh}}(s^2) = \dfrac{s^2}{s^2+\alpha}$ shows that Tikhonov regularization filters out singular components that are small (relative to α) while retaining components that are large.



Truncated SVD (TSVD) Regularization

TSVD filtering function is

$$\omega_\alpha^{TSVD}(s_i^2) = \begin{cases} 0, & s_i^2 \le \alpha, \\ 1, & s_i^2 > \alpha. \end{cases}$$

has "sharp cut-off" behavior instead of "smooth roll-off" behavior of Tikhonov filter.

Iterative Regularization

Certain iterative methods, e.g., steepest descent, conjugate gradients, and Richardson-Lucy (EM), have regularizing effects with the regularization parameter equal to the number of iterations. These are useful in applications, like 3-D imaging, with many unknowns. An example is Landweber iteration, a variant of steepest descent.