Solution to the equation $y = Ax$.

A is an $n \times m$ matrix. It means that we look for x such that

$$x = \arg\min \|Ax - y\|^2$$

Over determined case

If A is skinny ($n > m$), the set is called overdetermined. There are more equations that unknown therefore there is not an exact solution for $y = Ax$ and we just minimize the difference as given in the above equation. The solution is

$$x = (A^T A)^{-1} A^T y = A^+ y$$

Where A^+ is pseudo-inverse of A.

If X is a matrix, the solution is the same A^+. In fact if we had $y = Ax + n$ with n as noise and y the observation, this solution minimizes $\|n\|^2$.

Under determined case

If A is fat ($n < m$), there are many solutions because there are more unknowns than equations. We are looking for the solution that minimizes $\|x\|^2$.

$$x = A^T (AA^T)^{-1} y$$

Minimization of two metrics

If we are to minimize $J_1 = \|Ax - y\|^2$ and at the same time $J_2 = \|Fx - g\|^2$ where $x \in R^n$, one can minimize $J_1 + \mu J_2 = \|Ax - y\|^2 + \mu \|Fx - g\|^2$. We can write:

$$J_1 + \mu J_2 = \left\| \begin{bmatrix} A \\ \sqrt{\mu} F \end{bmatrix} x - \begin{bmatrix} y \\ \sqrt{\mu} g \end{bmatrix} \right\|^2 = \left\| \tilde{A}^T x - \tilde{y} \right\|^2$$

The solution will be

$$x = \left(\tilde{A}^T \tilde{A} \right)^{-1} \tilde{A}^T y = (A^T A + \mu F^T F)^{-1} (A^T y + \mu F^T g)$$

As an example if $F = I$ and $g = 0$, $J_2 = \|x\|^2$ and then

$$x = (A^T A + \mu I)^{-1} A^T y$$

Note that this solution works for all A because the square matrix $A^T A + \mu I$ is always invertible.

Example for MMSE receiver

$$y = Hx + n$$

We select $J_1 = \|Hx - y\|^2$ and $J_2 = E\|n\|^2$. With this assumption, $F = 0$ and $g = n$. Therefore we minimize the following metric:

$$\min \|Hx - y\|^2 + \sigma^2 I$$

The solution is

$$x = (H^*H + \sigma^2 I)^{-1} H^* y$$

Note that in the zero forcing method the noise is ignored and the solution is just pseudo-inverse $x = (H^*H)^{-1}H^*y$. So when the channel is weak the $(H^*H)^{-1}$ gives large values for x. There is also the problem of the rank of A, if it is not full rank, the inversion cannot take place.