Sample 2-d linear system of the form $Ax = b$:

$$
\begin{bmatrix}
3 & 2 \\
2 & 6
\end{bmatrix} x = \begin{bmatrix}
2 \\
-8
\end{bmatrix}.
$$
Graph of quadratic form $f(x) = \frac{1}{2}x^T Ax - b^T x + c$. The minimum point of this surface is the solution to $Ax = b$.

Contours of the quadratic form. Each ellipsoidal curve has constant $f(x)$. 
Gradient $f'(x)$ of the quadratic form. For every $x$, the gradient points in the direction of steepest increase of $f(x)$, and is orthogonal to the contour lines.
(a) Quadratic form for a positive-definite matrix.
(b) For a negative-definite matrix.
(c) For a singular (and positive-indefinite) matrix. A line that runs through the bottom of the valley is the set of solutions.
(d) For an indefinite matrix.
The method of Steepest Descent.
Steepest Descent converges to the exact solution on the first iteration if the error term is an eigenvector.
Steepest Descent converges to the exact solution on the first iteration if the eigenvalues are all equal.
(a) Large $\kappa$, small $\mu$.

(b) An example of poor convergence. $\kappa$ and $\mu$ are both large.

(c) Small $\kappa$ and $\mu$.

(d) Small $\kappa$, large $\mu$. 
Convergence of Steepest Descent (per iteration) worsens as the condition number of the matrix increases.
The method of Conjugate Gradients.
Convergence of Conjugate Gradients (per iteration) as a function of condition number.

Number of iterations of Steepest Descent required to match one iteration of CG.