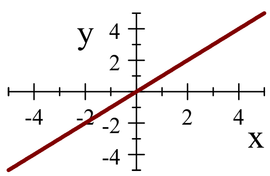
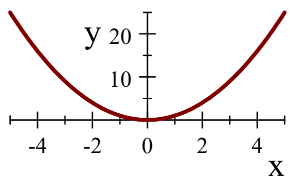
Instructor: Jodin Morey moreyjc@umn.edu

5.2 Minimization of Quadratic Functions

Simplest type of function to minimize is a nonlinear quadratic function (linear have no extrema).



linear



quadratic

Later, we will minimize general quadratic functions of n variables. But first:

n = 1: $p(x) = ax^2 + 2bx + c.$

Calculus Method

Note if $a > 0$, $p(x)$ opens up, one global minimum. If $a < 0$, opens down, no minimum.

Calculus: take derivative, set it to zero, solve to find local extrema $p(x^*)$.

$p'(x) = 2ax + 2b \quad \Rightarrow \quad x^* = -\frac{b}{a}$ and $p(x^*) = c - \frac{b^2}{a}$.

$p''(x) = 2a > 0$. So, if $a > 0$, then $p(x^*)$ is **local minimum**.

General Method

Rewrite as: $p(x) = a(x + \frac{b}{a})^2 + \left(c - \frac{b^2}{a}\right)$.

Observe $a > 0 \quad \Rightarrow \quad$ first-term ≥ 0. Moreover, min attained at $x^* = -\frac{b}{a}$.

Second term is constant, and so unaffected by x.

Thus, **global minimum** $(c - \frac{b^2}{a})$ (not just extrema) is attained at $x^* = -\frac{b}{a}$.

Now generalize to any # of vars!

n ≥ 1: $p(\vec{x}) = p(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} k_{ij} x_i x_j - 2 \sum_{i=1}^{n} f_i x_i + c$

Without loss of generality, we can assume $k_{ij} = k_{ji}$. (!?! see exercise 3.4.15)

Example: $\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 4 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = x^2 + 6xy + y^2$ and $\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = x^2 + 6xy + y^2$.

$p(\vec{x})$ is more general than quadratic form: has linear & constant terms.

Rewrite as: $p(\vec{x}) = \vec{x}^T\mathbf{K}\vec{x} - 2\vec{x}^T\vec{f} + c$, where $\mathbf{K} = (k_{ij})$ and $\vec{f} = (f_i)$.

Example: $p(\vec{x}) = 4x_1^2 - 2x_1x_2 + 3x_2^2 + 3x_1 - 2x_2 + 1$.

Rewriten: $p(\vec{x}) = \vec{x}^T\mathbf{K}\vec{x} - 2\vec{x}^T\vec{f} + c = \begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - 2\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} -\frac{3}{2} \\ 1 \end{bmatrix} + 1.$

In multi-var version, instead of requiring $a > 0$ (as we had in single-var case), we need $\mathbf{K} > 0$ in order to obtain a *unique* minimum.

Theorem: If $\mathbf{K} > 0$ (and hence symmetric), then quadratic function $p(\vec{x}) = \vec{x}^T\mathbf{K}\vec{x} - 2\vec{x}^T\vec{f} + c$ has a unique minimizer, which is the solution to the linear system $\mathbf{K}\vec{x} = \vec{f}$, namely $\vec{x}^* = \mathbf{K}^{-1}\vec{f}$.
The minimum value of $p(\vec{x})$ is equal to any of the following expressions:

$$p(\vec{x}^*) = p\left(\mathbf{K}^{-1}\vec{f}\right) = c - \vec{f}^T\mathbf{K}^{-1}\vec{f} = c - \vec{f}^T\vec{x}^* = c - (\vec{x}^*)^T\mathbf{K}\vec{x}^*. \qquad (*)$$

Proof: Recall positive definiteness implies $\mathbf{K}$ is nonsingular, hence $\mathbf{K}\vec{x} = \vec{f}$ has unique solution $\vec{x}^* = \mathbf{K}^{-1}\vec{f}$.

But is it the unique minimizer of $p(\vec{x})$?

Well, $\forall \vec{x} \in \mathbb{R}^n$, if $\vec{f} = \mathbf{K}\vec{x}^*$, it follows that

$p(\vec{x}) = \vec{x}^T\mathbf{K}\vec{x} - 2\vec{x}^T\vec{f} + c$

$= \vec{x}^T\mathbf{K}\vec{x} - 2\vec{x}^T\mathbf{K}\vec{x}^* + c$

$= \vec{x}^T\mathbf{K}\vec{x} - \vec{x}^T\mathbf{K}\vec{x}^* - \left(\vec{x}^T\mathbf{K}\vec{x}^*\right)^T + c$ (since the transpose of a constant is itself)

$= \vec{x}^T\mathbf{K}\vec{x} - \vec{x}^T\mathbf{K}\vec{x}^* - (\vec{x}^*)^T\mathbf{K}\vec{x} + (\vec{x}^*)^T\mathbf{K}\vec{x}^* + \left[c - (\vec{x}^*)^T\mathbf{K}\vec{x}^*\right]$ (distributed transpose, and added/subtracted same term)

$= \vec{x}^T\mathbf{K}(\vec{x} - \vec{x}^*) - (\vec{x}^*)^T\mathbf{K}(\vec{x} - \vec{x}^*) + \left[c - (\vec{x}^*)^T\mathbf{K}\vec{x}^*\right]$ (factoring out)

$$= (\vec{x}-\vec{x}^*)^T \mathbf{K}(\vec{x}-\vec{x}^*) + \left[c - (\vec{x}^*)^T \mathbf{K}\vec{x}^*\right], \qquad \text{(factoring out again)} \qquad (**)$$

The first term in $(**)$ has form $\vec{y}^T \mathbf{K}\vec{y}$, where $\vec{y} = \vec{x} - \vec{x}^*$.

Since we assumed **K** is positive definite, we know $\vec{y}^T \mathbf{K}\vec{y} > 0$ for all $\vec{y} \neq \vec{0}$.

Thus, the first term achieves its minimum value (namely zero) **iff** $\vec{0} = \vec{y} = \vec{x} - \vec{x}^*$.

Since $\vec{x}^*$ is fixed, the second, bracketed term in $(**)$ doesn't depend on $\vec{x}$, and hence the minimizer of $p(\vec{x})$ coincides with the minimizer of the first term, namely $\vec{x} = \vec{x}^*$.

Moreover, the minimum value of $p(\vec{x})$ is equal to the constant term: $p(\vec{x}^*) = c - (\vec{x}^*)^T \mathbf{K}\vec{x}^*$.

The alternative expressions in $(*)$ follow from simple substitutions. ■

Continued Example: Find minimizer of

$$p(\vec{x}) = \vec{x}^T \mathbf{K}\vec{x} - 2\vec{x}^T \vec{f} + c = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - 2\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} -\frac{3}{2} \\ 1 \end{bmatrix} + 1.$$

According to the thm, we must solve: $\mathbf{K}\vec{x} = \vec{f}$, which from above is:

$$\begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} \\ 1 \end{bmatrix}.$$

$$\left[\begin{array}{cc|c} 4 & -1 & -\frac{3}{2} \\ -1 & 3 & 1 \end{array}\right] \to \left[\begin{array}{cc|c} 4 & -1 & -\frac{3}{2} \\ 0 & \frac{11}{4} & \frac{5}{8} \end{array}\right].$$

K appears regular, with positive pivots. Therefore: $\mathbf{K} > 0$ and $p(\vec{x})$ does have a global minimum.

Solving yields minimizer is: $\vec{x}^* = \begin{bmatrix} -\frac{7}{22} \\ \frac{5}{22} \end{bmatrix}$.

Therefore the minimum (most easily acquired from $p(\vec{x}^*) = c - \vec{f}^T \vec{x}^*$ from the thm) is

$$p(\vec{x}^*) = p\left(-\tfrac{7}{22}, \tfrac{5}{22}\right) = 1 - \begin{bmatrix} -\frac{3}{2} & 1 \end{bmatrix} \begin{bmatrix} -\frac{7}{22} \\ \frac{5}{22} \end{bmatrix} = \frac{13}{44}.$$

Example: For $\mathbf{K} = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 3 \end{bmatrix}$, $\vec{f} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$, and $c = -3$, write out the quadratic function $p(\vec{x}) = \vec{x}^T\mathbf{K}\vec{x} - 2\vec{x}^T\vec{f} + c$, where $\vec{x} \in \mathbb{R}^n$. Then either find the minimizer $\vec{x}^*$ and minimum value $p(\vec{x}^*)$, or explain why there is none.

According to the thm, we must solve: $\mathbf{K}\vec{x} = \vec{f}$, which from above is:

$$\begin{bmatrix} 3 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}.$$

$$\left[\begin{array}{ccc|c} 3 & -1 & 1 & 1 \\ -1 & 2 & -1 & 0 \\ 1 & -1 & 3 & -2 \end{array}\right] \to \left[\begin{array}{ccc|c} 3 & -1 & 1 & 1 \\ 0 & \frac{5}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & 1 & 2 & -2 \end{array}\right] \to \left[\begin{array}{ccc|c} 3 & -1 & 1 & 1 \\ 0 & \frac{5}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 12 & -7 \end{array}\right].$$

$\mathbf{K}$ was regular, with positive pivots. Therefore: $\mathbf{K} > 0$ and $p(\vec{x})$ does have a global minimum.

Solving for minimizer: $\vec{x}^* = \begin{bmatrix} \frac{7}{12} \\ -\frac{1}{6} \\ -\frac{11}{12} \end{bmatrix}$.

Therefore the minimum (most easily acquired from $p(\vec{x}^*) = c - \vec{f}^T\vec{x}^*$ from the thm) is

$$p(\vec{x}^*) = p\left(-\tfrac{7}{22}, \tfrac{5}{22}\right) = -3 - \begin{bmatrix} 1 & 0 & -2 \end{bmatrix}\begin{bmatrix} \frac{7}{12} \\ -\frac{1}{6} \\ -\frac{11}{12} \end{bmatrix} = -\frac{65}{12}.$$