Critical Points

- Given a function $f(\vec{x})$ set the derivatives equal to zero to identify the critical points.
- In 1D, find *all* critical points $x$ that solve $f'(x) = 0$.
- In multiple spatial dimensions, set the gradient to zero, i.e. $\nabla f(\vec{x}) = 0$, and solve for $\vec{x}$.
- This is a system of equations: $\frac{\partial f}{\partial x_1}(\vec{x}) = 0, \ldots, \frac{\partial f}{\partial x_i}(\vec{x}) = 0, \ldots, \frac{\partial f}{\partial x_n}(\vec{x}) = 0$, where there is one equation for each $x_i$.
- Any $\vec{x}$ that simultaneously satisfies all the equations is a critical point.

Classifying Critical Points

- The second derivative is used to classify critical points.
- In 1D, given a critical point $x$:
  - if $f''(x) > 0$, concave up, minimum
  - if $f''(x) < 0$, concave down, maximum
  - otherwise, when the second derivative vanishes (inflection point), neither min/max
- In multiple spatial dimensions, consider the Hessian matrix $H(\vec{x})$ of all second partial derivates, i.e. $H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$
  - $H(\vec{x})$ is symmetric since the order of differentiation doesn’t matter, i.e. $H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} = H_{ji}$
  - if $H$ is positive definite at a critical point, then it’s a local minimum
  - if $H$ is negative definite at a critical point, then it’s a local maximum
  - otherwise $H$ is indefinite and the critical point is a saddle point.
- In 1D $H(x) = [f''(x)]$, a 1x1 matrix, and it is positive/negative definite when $f''(x)$ is positive/negative, respectively (and indefinite when $f''(x) = 0$)
- 2D example with $H$ negative definite, positive definite, and indefinite.

<table>
<thead>
<tr>
<th></th>
<th>Local maxima</th>
<th>Local minima</th>
<th>Saddle</th>
</tr>
</thead>
</table>

![3D plots of local maxima, local minima, and saddle points](image)
**Quadratic Form**

- Given a square matrix $A$, the quadratic form is $f(x) = \frac{1}{2} x^T A x - b^T x + c$
- Minimize $f(x)$ by taking the gradient and setting it equal to zero:
  - $\nabla f(x) = \frac{1}{2} A x + \frac{1}{2} A^T x - b = A x - b = 0$ assuming $A$ is symmetric
  - E.g. in 1D, $f(x) = \frac{1}{2} ax^2 - bx + c$ has a critical point at its line of symmetry at $x=b/a$
  - So, solve $A x = b$ to find the critical point
- Check the second derivative matrix to categorize the critical point
  - Second derivative (Hessian) $H = \frac{1}{2} (A + A^T) = A$ assuming $A$ is symmetric
  - If $A$ is SPD, the solution to $A x = b$ is a minimum
  - If $A$ is Symmetric Negative Definite, the solution to $A x = b$ is a maximum
  - Indefinite $A$
  - E.g. in 1D, $f(x) = \frac{1}{2} ax^2 - bx + c$ has Hessian $H = [a]$, and $a > 0$ indicates concave up with $x = b/a$ representing a minimum

**Least Squares**

- Minimize $r = b - A c$ using the L2 norm, i.e. minimize $\sqrt{r \cdot r}$
- Equivalent to minimizing $c^T A^T A c - 2 b^T A c + b^T b$ or $\frac{1}{2} c^T A^T A c - b^T A c$
- $A^T A$ is obviously symmetric
- Using $A = U \Sigma V^T$ gives $A^T A = \Sigma^T U^T U \Sigma V^T = \Sigma^T \Sigma V^T = \Sigma^2 V^T$ where $\Sigma^2 = \Sigma^T \Sigma$ is a diagonal matrix of positive numbers since $A$ has full column rank
- Solving $(\Sigma^2 V^T) v = \lambda v$ is equivalent to $\Sigma^2 (V^T v) = \lambda (V^T v)$ showing that the eigenvalues of $A^T A$ are the square of the singular values of $A$
  - And the eigenvectors $v$ have $V^T v = e_k$ for a standard basis vectors $e_k$
  - Or $v = V e_k$, that is, the eigenvectors of $A^T A$ are the columns of $V$
- Thus, $A^T A$ is SPD, and the minimum of $\frac{1}{2} c^T A^T A c - b^T A c$ is found by solving $A^T A c = A^T b$
- $A^T A c = A^T b$ are called the normal equations