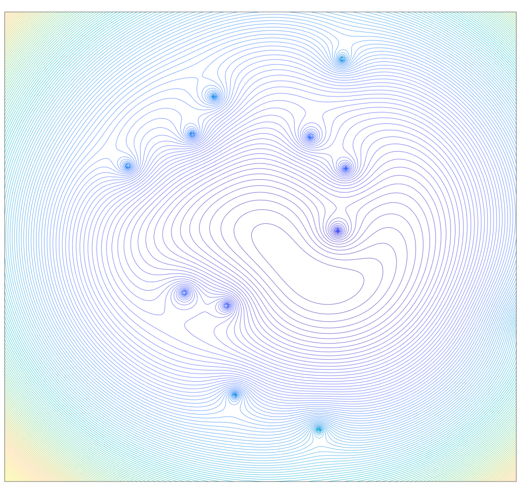
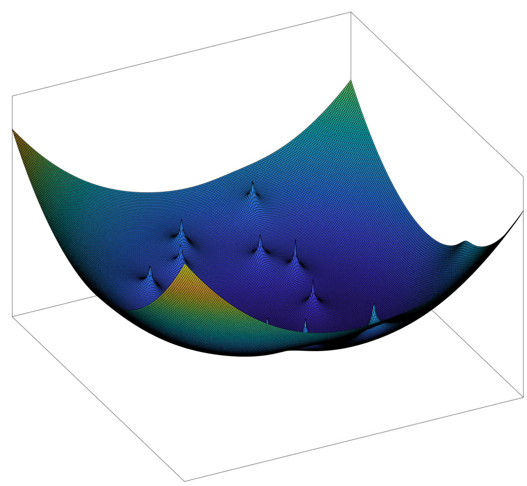
Representative and Sufficiently Distinct Locations

Suppose we have m transmitters located across some area of land and we need to place a single receiver at some location to listen to all transmitters. If the transmitters are weak, then we might wish to have a central location, for example placing the receiver at the mean position of the transmitters. However, if the receiver is too close to a transmitter then this transmitter can interfere with signal reception from other transmitters. So, the mean position may not be optimal. We can formalize this idea as follows.

Let $\{p_k\}_{k=1}^m \subset \mathbb{R}^2$ be m transmitter locations (Each p_k is a position vector in $\mathbb{R}^2$). The mean transmitter position is $p = \frac{1}{m}\sum_{k=1}^m p_k$. We can then seek a receiver location as the solution to the optimization problem

$$\min_{x\in\mathbb{R}^2} \; f(x) = \frac{1}{2}\|x-p\|^2 - \frac{\mu}{2m}\sum_{k=1}^{m} \ln \|x-p_k\|^2 \tag{12}$$

The first term is minimized at the mean position $x = p$. The second term diverges when $x = p_k$ for any k. The fixed parameter $\mu > 0$ governs the relative emphasis on these two terms. Thus f is minimized at some compromise location that is near the mean position but not too close to any one location p_k. Shown below are a contour plot and a surface plot of the objective function for twelve randomly chosen transmitter locations and $\mu = 18$.

Task 1. *Determine, with justification, whether or not the objective function given in Equation 12 is bounded below, continuous, continuously differentiable, convex, coercive. Which, if any, of your conclusions change if $x \in \mathbb{R}^n$ with $n > 2$?*

In this project, we will consider the case of finding an optimal location in $\mathbb{R}^n$ with the additional generalization that uses values $\mu_k > 0$ to weight each term in the second sum individually. That is, we seek to solve the problem

$$\min_{x\in\mathbb{R}^n} \; f(x) = \frac{1}{2}\|x-p\|^2 - \frac{1}{2m}\sum_{k=1}^{m} \mu_k \ln \|x-p_k\|^2 \tag{13}$$

where p_k are m given vectors in $\mathbb{R}^n$, $\mu > 0$ is a weight vector in $\mathbb{R}^m$, and p is the mean position as defined previously.

Task 2. *Code an objective function that returns $f(x)$ and $\nabla f(x)$ for a given decision variable input $x \in \mathbb{R}^n$ and parameters p_k and μ. Use analytically computed gradient computations (not finite difference approximations). Your code should return $f(x) = \infty$ and $\nabla f(x) = \infty$ if $x = p_k$ for any k.*

One possible application of problem 13 is iterate selection in expensive derivative free optimization. Suppose we have an objective function $F(x)$ which takes a long time to evaluate (hours, days, weeks!) and we have no derivative information. Let's say we have evaluated m decision variable vectors $\{p_k\}$ and we wish to select a new point for evaluation that is likely to provide an improved point. We would like to select a point in the neighborhood of our current point set (because all of our objective function information is in this neighborhood) and we also do not want to choose a new point too close to a point we have already evaluated (because this would not be a very good global strategy). But we also want to use the existing objective value information, so we might use large weights associated with points of larger objective value (avoid these points more strongly). The optimization problem 13 is very easy and quick to solve and provides a *candidate* decision variable vector for the larger expensive problem. If the weights are chosen large enough to strongly avoid all points, then the optimal point can lie either inside or outside of the convex hull of the point group $\{p_k\}$.

Task 3. *Consider using our course methods (line search and trust region) on optimization problem 13. Carefully argue that if f is differentiable at the j^{th} iterate $x^{(j)}$ then f will also be differentiable at iterate $x^{(j+1)}$. Then, using the fact that there are a finite number of points of non-differentiability and your results from Task 1, argue that our methods should converge to a local minimizer of f. The convergence theorems we discussed early in the course may be helpful.*

Task 4. *Create a smooth (at least twice continuously differentiable) function that might be used to replace the second term of the objective in problem 13. Can your modified objective function be used to reasonably solve the current problem?*

Task 5. *Solve optimization problem 13 (as shown, not modified) using the point set and weight parameters provided by the instructor and an optimization method of your choice. Use a simply strategy to attempt to find all local minima. Discuss your methods and results.*